Sampling with positive definite kernels and an associated dichotomy

PALLE JORGENSEN AND JAMES TIAN

We study classes of reproducing kernels K on general domains; these are kernels which arise commonly in machine learning models; models based on certain families of reproducing kernel Hilbert spaces. They are the positive definite kernels K with the property that there are countable discrete sample-subsets S; i.e., proper subsets S having the property that every function in $\mathscr{H}(K)$ admits an S-sample representation. We give a characterizations of kernels which admit such non-trivial countable discrete sample-sets. A number of applications and concrete kernels are given in the second half of the paper.

1. Introduction

In the theory of non-uniform sampling, one studies Hilbert spaces consisting of signals, understood in a very general sense. One then develops analytic tools and algorithms, allowing one to draw inference for an "entire" (or global) signal from partial information obtained from carefully chosen distributions of sample points. While the better known and classical sampling algorithms (Shannon and others) are based on interpolation, modern theories go beyond this. An early motivation is the work of Henry Landau, see e.g., [Lan60, LP61, Lan64, Lan67, LLSB84, HLS85]. In this setting, it is

possible to make precise the notion of "average sampling rates" in general configurations of sample points. (See also [ACH+10, AACM11].)

When a positive definite kernel K is given, we denote by $\mathscr{H}(K)$ the associated reproducing kernel Hilbert space (RKHS). In the present paper we study classes of reproducing kernels K on general domains, such kernels arise commonly in machine learning models based on reproducing kernel Hilbert space (see e.g., [JT15b]) 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.

We offer an operator theoretic condition which explains, in a general context, this dichotomy. Our study continues our earlier papers on reproducing kernels and their restrictions to countable discrete subsets; see e.g., [JT16, AJ15, JT15b, JT15a], and also [NSW11, SRB+10, SZ09, SY06, SZ05].

A reproducing kernel Hilbert space (RKHS) is a Hilbert space $\mathscr{H}$ of functions on a prescribed set, say T, with the property that point-evaluation for functions $f \in \mathscr{H}$ is continuous with respect to the $\mathscr{H}$-norm. They are called kernel spaces, because, for every $t \in T$, the point-evaluation for functions $f \in \mathscr{H}$, $f(t)$ must then be given as a $\mathscr{H}$-inner product of f and a vector K_t, in $\mathscr{H}$; called the kernel.

The RKHSs have been studied extensively since the pioneering papers by Aronszajn [Aro43, Aro48]. They further play an important role in the theory of partial differential operators (PDO); for example as Green's functions of second order elliptic PDOs [Nel57, HKL+14]. Other applications include engineering, physics, machine-learning theory [KH11, SZ09, CS02], stochastic processes [AD93, ABDdS93, AD92, AJSV13, AJV14], numerical analysis, and more [LB04, HQKL10, ZXZ12, LP11, Vul13, SS13, HN14, STC04, SS01].

An illustration from *neural networks*: An Extreme Learning Machine (ELM) is a neural network configuration in which a hidden layer of weights are randomly sampled [RW06], and the object is then to determine analytically resulting output layer weights. Hence ELM may be thought of as an approximation to a network with infinite number of hidden units.

Given a positive definite kernel $K : T \times T \to \mathbb{C}$ (or $\mathbb{R}$ for simplification), there are several notions and approaches to sampling (i.e., an algorithmic

reconstruction of suitable functions from values at a fixed and pre-selected set of sample-points):

Definition 1.1. We say that K has non-trivial sampling property, if there exists a countable subset $S \subset T$, and $a, b \in \mathbb{R}_+$, such that

$$(1.1) \qquad a \sum_{s \in S} |f(s)|^2 \leq \|f\|^2_{\mathscr{H}(K)} \leq b \sum_{s \in S} |f(s)|^2, \quad \forall f \in \mathscr{H}(K),$$

where $\mathscr{H}(K)$ is the reproducing kernel Hilbert space (RKHS) of K, see [Aro43] and Remark 1.2 below.

Suppose equality holds in (1.1) with $a = b = 1$; then we say that $\{K(\cdot, s)\}_{s \in S}$ is a *Parseval frame*.

It follows that sampling holds in the form

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

if and only if $\{K(\cdot, s)\}_{s \in S}$ is a Parseval frame; see also Theorem 2.10.

As is well known, when a vector f in a Hilbert space $\mathscr{H}$ is expanded in an orthonormal basis (ONB) B, there is then automatically an associated Parseval identity. In physical terms, this identity typically reflects a *stability* feature of a decomposition based on the chosen ONB B. Specifically, Parseval's identity reflects a conserved quantity for a problem at hand, for example, energy conservation in quantum mechanics.

The theory of frames begins with the observation that there are useful vector systems which are in fact not ONBs but for which a Parseval formula still holds. In fact, in applications it is important to go beyond ONBs. While this viewpoint originated in signal processing (in connection with frequency bands, aliasing, and filters), the subject of frames appears now to be of independent interest in mathematics. See, e.g., [BCL11, Chr14, HKLW07, FJKO05], and also [CD93, BDP05, Dut06].

Remark 1.2. To make the discussion self-contained, we add the following (for the benefit of the readers.)

(i) A given $K : T \times T \to \mathbb{C}$ is positive definite (p.d.) if and only if for all $n \in \mathbb{N}$, $\{\xi\}_{j=1}^n \subset \mathbb{C}$, and all $\{t_j\}_{j=1}^n \subset T$, we have:

$$\sum_i \sum_j \bar{\xi}_i \xi_j K(t_i, t_j) \geq 0.$$

(ii) A function f on T is in $\mathscr{H}(K)$ if and only if there is a constant $C = C(f)$ such that for all n, $(\xi_j)_1^n$, $(t_j)_1^n$, as above, we have

(1.2)
$$\left| \sum_1^n \xi_j f(t_j) \right|^2 \le C \sum_i \sum_j \overline{\xi}_i \xi_j K(t_i, t_j).$$

It follows from the above that reproducing kernel Hilbert spaces (RKHS) arise from a given positive definite kernel K, a corresponding pre-Hilbert form; and then a Hilbert-completion. The question arises: "What are the functions in the completion?" The *a priori* estimate (1.2) in (ii) above is an answer to the question. We will return to this issue in the application section 3 below. By contrast, the Hilbert space completions are subtle; they are classical Hilbert spaces of functions, not always transparent from the naked kernel K itself. Examples of classical RKHSs: Hardy spaces or Bergman spaces (for complex domains), Sobolev spaces and Dirichlet spaces [OST13, ST12, Str10] (for real domains, or for fractals), band-limited L^2 functions (from signal analysis), and Cameron-Martin Hilbert spaces (see Lemma 2.3) from Gaussian processes (in continuous time domain).

Lemma 1.3. *Suppose K, T, a, b, and S satisfy the condition in (1.1), then 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,

$$f(t) = \sum_{s \in S} f(s) B(K(\cdot, s))(t), \text{ for all } t \in T.$$

Proof. Define $A : \mathscr{H}(K) \to l^2(S)$ by $(Af)(s) = f(s)$, $s \in S$; or

$$Af := (f(s))_{s \in S} \in l^2(S).$$

Then the adjoint operator $A^* : l^2(S) \to \mathscr{H}(K)$ is given by

$$A^*\xi = \sum_{s \in S} \xi_s K(\cdot, s), \; \forall \xi \in l^2(S),$$

and

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

holds in $\mathscr{H}(K)$, with $\mathscr{H}(K)$-norm convergence. This is immediate from (1.1).

Now set $B = (A^*A)^{-1}$. Note that

$$\|B\|_{\mathscr{H}(K) \to \mathscr{H}(K)} \leq a^{-1}$$

where a is in the lower bound in (1.1). □

Lemma 1.4. *Suppose K, T, a, b, and S satisfy (1.1), then the linear span of $\{K(\cdot, s)\}_{s \in S}$ is dense in $\mathscr{H}(K)$.*

Proof. Let $f \in \mathscr{H}(K)$, then

$$f \perp \{K(\cdot, s)\}_{s \in S}$$
$$\Updownarrow$$
$$f(s) = \langle K(\cdot, s), f \rangle_{\mathscr{H}(K)} = 0, \ \forall s \in S,$$

by the reproducing property in $\mathscr{H}(K)$. But by (1.1), $b < \infty$, this implies that $f = 0$ in $\mathscr{H}(K)$. Hence the family $\{K(\cdot, s)\}_{s \in S}$ has dense span. □

2. The dichotomy

While questions of sampling have, traditionally, been restricted to settings of signals when some form of time-frequency duality is available, we take here a more general viewpoint, that of general reproducing kernels. It, in turn, encompasses the more traditional Shannon point of view, but our present scope is much wider. Indeed, the more familiar and classical approaches typically do entail special positive definite (p.d.) kernels, but only a very specialized class of p.d. kernels. In detail, in the traditional context of time frequency duality, there are already a variety of approaches to selecting discrete point-configurations which produce a faithful representation of a continuous time signal. In fact this viewpoint includes cases where "signal" encompasses images in suitable digital representations. While the literature here is large, we mention here just [ACH+10, BCL11, Lan60, Lan67] and the papers cited there. Even in these more restricted contexts, the search for discrete point-configurations which produce a faithful representations is highly non-trivial. By contrast, our present focus is that of general positive definite (p.d.) kernels.

While the traditional contexts of various classes of "band-limited signals" typically narrow the search for kernels down to the special cases of

positive definite (p.d.) kernels which arise in multivariable Fourier analysis, our present context widens the focus to a much broader class of p.d. kernels, kernels arising naturally in a number of infinite-dimensional contexts from mathematical physics, and from neighboring areas; including, for example, new classes of p.d. kernels arising directly from diverse contexts of features, and feature space, analysis; more precisely, p.d. kernels arising as part of kernel-algorithms in machine learning. In these wider (non-Fourier) approaches, the question of dichotomy is then especially natural: If a particular p.d. kernel arises from a feature-space analysis, from quantum observables; or in the analysis of fractals (see e.g., [Str10, SY06, SZ09, ZXZ12]), then what is the variety of discrete point-configurations which produce a faithful representations. Dichotomy: Given a p.d. kernel, when is it possible to select associated discrete point-configurations with faithful sampling features, and when not?

We now turn to dichotomy: (i) Existence of countably discrete sampling sets vs (ii) non-existence. To help readers appreciate the nature of the two classes, we begin with two examples, (i) Shannon's kernel for band-limited functions, Example 2.1; and (ii) the covariance kernel for standard Brownian motion, Theorem 2.5.

Question.

  (*i*) *Given a positive definite kernel $K : T \times T \to \mathbb{R}$, how to determine when there exist $S \subset T$, and $a, b \in \mathbb{R}_+$ such that (1.1) holds?*

 (*ii*) *Given K, T as above, how to determine if there is a countable discrete subset $S \subset T$ such that*

$$(2.1) \qquad\qquad \{K(\cdot, s)\}_{s \in S}$$

  has dense span in $\mathscr{H}(K)$?

Example 2.1. Let $T = \mathbb{R}$, and let $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be the Shannon kernel, where

$$(2.2) \qquad K(s,t) := \operatorname{sinc} \pi(s-t)$$
$$= \frac{\sin \pi(s-t)}{\pi(s-t)}, \quad \forall s, t \in \mathbb{R}.$$

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 (1.1) holds for finite $a, b \in \mathbb{R}_+$.

The RKHS of K in (2.2) is the Hilbert space $\subset L^2(\mathbb{R})$ consisting of all $f \in L^2(\mathbb{R})$ such that $suppt(\hat{f}) \subset [-\pi, \pi]$, 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 $\mathbb{C}$; see [Kat04, Sei04, LP11, Pau02]. Using the above observations, we get

$$f(t) = \sum_{n \in \mathbb{Z}} f(n) K(t, n)$$

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

Example 2.2. Let K be the covariant kernel of standard Brownian motion, with $T := [0, \infty)$, or $T := [0, 1)$. Then

$$(2.3) \qquad K(s, t) := s \wedge t = \min(s, t), \ \forall (s, t) \in T \times T.$$

Lemma 2.3. *Let K, T be as in (2.3). Then $\mathscr{H}(K)$ consists of functions f on T such that f has distribution derivative $f' \in L^2(T, \lambda)$, i.e., L^2 with respect to Lebesgue measure λ on T, and*

$$(2.4) \qquad \|f\|^2_{\mathscr{H}(K)} = \int_T |f'(x)|^2 \, dx.$$

Proof. This is well-known, see e.g., [JT15b, AJ15, Hid80]. $\qquad\square$

Remark 2.4 (see also Section 3.1 below). The significance of (2.4) for Brownian motion is as follows:

Fix T, and set $L^2(T) =$ the L^2-space from the restriction to T of Lebesgue measure on $\mathbb{R}$. Pick an ONB $\{\psi_k\}$ in $L^2(T)$, for example a Haar-Walsh orthonormal basis in $L^2(T)$. Let $\{Z_k\}$ be an i.i.d. (independent identically distributed) $N(0, 1)$ system, i.e., standard Gaussian copies. Then

$$(2.5) \qquad B_t(\cdot) = \sum_k \left(\int_0^t \psi_k(s) \, ds \right) Z_k(\cdot)$$

is a realization of standard Brownian motion on T; in particular we have

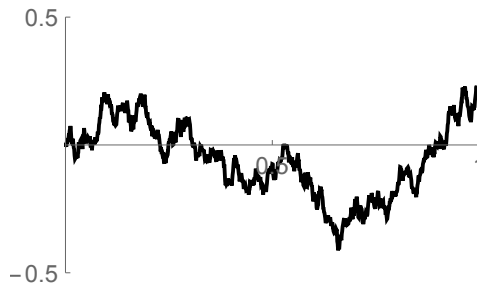$$\mathbb{E}(B_s B_t) = s \wedge t = K(s, t), \ \forall (s, t) \in T \times T.$$

See Figure 2.1.

Figure 2.1: Brownian motion; see (2.5).

Theorem 2.5. *Let K, T be as in (2.3); then there is no countable discrete subset $S \subset T$ such that $\{K(\cdot, s)\}_{s \in S}$ is dense in $\mathscr{H}(K)$.*

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

$$(2.6) \qquad\qquad 0 < x_1 < x_2 < \cdots < x_n < x_{n+1} < \cdots ;$$

then consider the following function

$$(2.7)$$

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

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

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

$$f\left(\frac{x_n + x_{n+1}}{2}\right) = c_n \frac{x_{n+1} - x_n}{2},$$
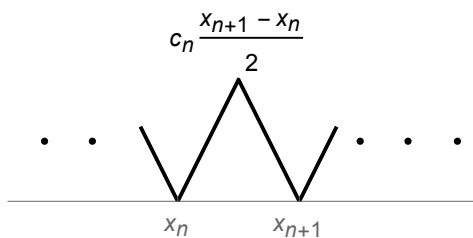
see Figure 2.2.
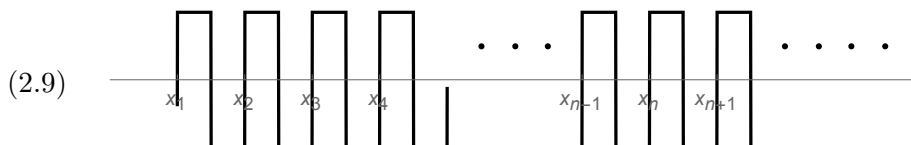
Figure 2.2: The saw-tooth function.

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

(2.8)
$$\sum_{n\in\mathbb{N}} |c_n|^2 (x_{n+1} - x_n) < \infty.$$

Admissible choices for the slope-values c_n include

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

We will now show that $f \in \mathscr{H}(K)$. To do this, use (2.4). For the distribution derivative computed from (2.7), we get

(2.9)

$$\int_0^\infty |f'(x)|^2 \, dx = \sum_{n\in\mathbb{N}} |c_n|^2 (x_{n+1} - x_n) < \infty$$

which is the desired conclusion, see (2.7). □

Corollary 2.6. *For the kernel $K(s,t) = s \wedge t$ in (2.3), $T = [0,\infty)$, 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

(2.10) $$f(x_j) = y_j, \ f \in \mathscr{H}(K)$$

is solvable if

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

Proof. Let f be the piecewise linear spline (see Figure 2.3) for the problem (2.10), see Figure 2.3; then the $\mathscr{H}(K)$-norm is as follows:

$$\int_0^\infty |f'(x)|^2 \, dx = \sum_{j\in\mathbb{N}} \left(\frac{y_{j+1} - y_j}{x_{j+1} - x_j}\right)^2 (x_{j+1} - x_j) < \infty$$
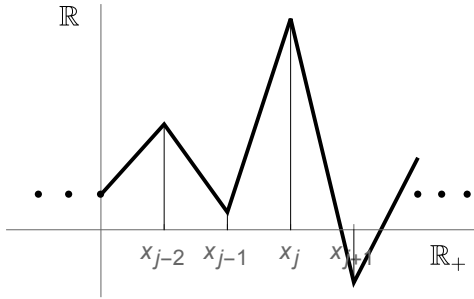
when (2.11) holds. $\square$
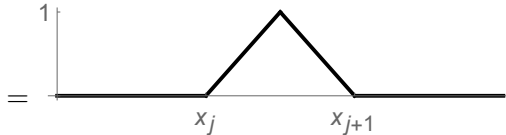

Figure 2.3: Piecewise linear spline.

Remark 2.7. Let K be as in (2.3), where

$$K(s,t) = s \wedge t, \quad s,t \in [0,\infty).$$

For all $0 \le x_j < x_{j+1} < \infty$, let

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

Assuming (2.8) holds, then

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

Remark 2.8. Let $K(s,t) = s \wedge t$, $(s,t) \in [0,\infty) \times [0,\infty)$, extend to $\widetilde{K}(s,t) = |s| \wedge |t|$, $(s,t) \in \mathbb{R} \times \mathbb{R}$, and $\mathscr{H}(\widetilde{K}) = $ all f on $\mathbb{R}$ such that the distribution derivative f' exists on $\mathbb{R}$, and

$$\|f\|_{\mathscr{H}(\widetilde{K})}^2 = \int_{\mathbb{R}} |f'(x)|^2 \, dx.$$

Theorem 2.9. *Let T be a set of cardinality c of the continuum, and let $K : T \times T \to \mathbb{R}$ be a positive definite kernel. Let $S = \{x_j\}_{j \in \mathbb{N}}$ be a discrete subset of T. Suppose there are weights $\{w_j\}_{j \in \mathbb{N}}$, $w_j \in \mathbb{R}_+$, such that*

(2.12)
$$(f(x_j)) \in l^2(\mathbb{N}, w)$$

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

(2.13)
$$\inf_{f \in \mathscr{H}(K)} \left\{ \sum_j w_j |f(x_j)|^2 + |f(t_0) - y_0|^2 + \alpha \|f\|_{\mathscr{H}(K)}^2 \right\}$$

is strictly positive.

Then S is <u>not</u> a interpolation set for (K, T).

Proof. Let L denote the analysis operator defined from condition (2.12) in the statement of the theorem; see also the beginning in the proof of Lemma 1.3 above, and let L^* denote the corresponding adjoint operator, the synthesis operator. Using now [SY06, JT15b], we conclude that the function f which minimizes the problem (2.13) is unique, and in fact

(2.14)
$$f = (\alpha I + L^* L)^{-1} L^* ((y_j) \cup (t_0)).$$

So, by the hypothesis in the theorem, we get $f \in \mathscr{H}(K) \backslash \{0\}$, and $f(x_j) = 0$, for all $j \in \mathbb{N}$. Then it follows that the closed span of $\{K(\cdot, x_j)\}_{j \in \mathbb{N}}$ is not $\mathscr{H}(K)$; specifically, $0 \neq f \in \{K(\cdot, x_j)\}_{j \in \mathbb{N}}^\perp$. See also Lemma 1.4 and Figure 2.4. $\square$
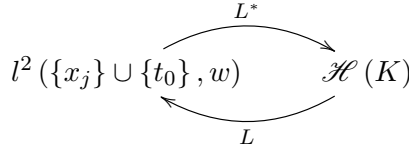
$$l^2\left(\{x_j\} \cup \{t_0\}, w\right) \quad\overset{L^*}{\underset{L}{\rightleftarrows}}\quad \mathscr{H}(K)$$

Figure 2.4: Analysis and synthesis operators.

Theorem 2.10. *Let $K : T \times T \to \mathbb{R}$ be a positive definite kernel, and let $S \subset T$ be a countable discrete subset. The RKHS $\mathscr{H}(K)$ refers to the pair (K, T). For all $s \in S$, set $K_s(\cdot) = K(\cdot, s)$. The the following four conditions are equivalent:*

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

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

(*iii*)
$$K(t,t) = \sum_{s \in S} |K(t,s)|^2, \ \forall t \in T;$$

(*iv*)
$$f(t) = \sum_{s \in S} f(s)K(t,s), \ \forall f \in \mathscr{H}(K), \ \forall t \in T,$$

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

Proof. (i) $\Rightarrow$ (ii). Assume (i), and note that

$$(2.15) \qquad\qquad \langle K_s, f\rangle_{\mathscr{H}(K)} = f(s);$$

and (ii) is immediate from the definition of Parseval-frame.

(ii) $\Rightarrow$ (iii). Assume (ii), and set $f = K_t$. Note that $\|K_t\|^2_{\mathscr{H}(K)} = K(t,t)$, and $\langle K_s, K_t\rangle_{\mathscr{H}(K)} = K(s,t)$.

(iii) $\Rightarrow$ (iv). It is enough to prove that

$$(2.16) \qquad\qquad K_t = \sum_{s \in S} K(t,s)K_s, \ \forall t \in T;$$

then (iv) follows from an application of the reproducing property of the Hilbert space $\mathscr{H}(K)$. Now (2.16) follows from

$$(2.17) \qquad \left\| K_t - \sum\nolimits_{s \in S} K(t,s) K_s \right\|_{\mathscr{H}(K)}^2 = 0.$$

Finally, (2.17) follows from (iii) and multiple application of the kernel property:

$$\text{LHS}_{(2.17)} = K(t,t) + \sum_{(s,s') \in S \times S} \sum K(t,s) K(t,s') K(s',s)$$
$$- 2 \sum_{s \in S} |K(t,s)|^2 = 0.$$

(iv) $\Rightarrow$ (i). It is clear that (i) $\Leftrightarrow$ (ii), and that (iv) $\Rightarrow$ (ii). $\qquad\square$

Remark 2.11 (Stationary kernels). Suppose $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a continuous positive definite kernel, and $K(s,t) = k(s-t)$, i.e., stationary. Set $K_t(\cdot) := K(\cdot,t) = k(\cdot - t)$. By Bochner's theorem,

$$k(t) = \int_{\mathbb{R}} e^{itx} d\mu(x),$$

where μ is a finite positive Borel measure on $\mathbb{R}$. Thus,

$$V : K_t \longmapsto e^{-itx} \in L^2(\mu)$$

extends to an isometry from $\mathscr{H}(K)$ into $L^2(\mu)$.

Let $S \subset \mathbb{R}$ be a countable discrete subset, then for $f \in \mathscr{H}(K)$, we have

$$\langle K_s, f \rangle_{\mathscr{H}(K)} = 0, \ \forall s \in S$$
$$\Updownarrow$$
$$\langle V K_s, V f \rangle_{L^2(\mu)} = 0, \ \forall s \in S$$
$$\Updownarrow$$
$$\int_{\mathbb{R}} e^{isx} (Vf)(x) \, d\mu(x) = 0, \ \forall s \in S.$$

So S has the sampling property if and only if

$$\left[((Vf) \, d\mu)^{\wedge}(s) = 0, \ \forall s \in S \right] \Longrightarrow \left[Vf = 0, \ i.e., \ f = 0, \ \mu - \text{a.e.} \right]$$

3. Discrete RKHSs

A closely related question from the above discussion is the dichotomy of *discrete* vs *continuous* RKHSs. Our focus in the present section is on the discrete case, i.e., RKHSs of functions defined on a prescribed countable infinite discrete set V.

Definition 3.1 ([JT15b]). 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 $x \in V$. Here, $\delta_x(y)$ is the Dirac mass at $x \in V$.

Question 3.2. *Let $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be positive definite, and let $V \subset \mathbb{R}^d$ be a countable discrete subset. When does $K\big|_{V \times V}$ have the discrete mass property?*

Of the examples and applications where this question plays an important role, we emphasize three: (i) discrete Brownian motion-Hilbert spaces, i.e., discrete versions of the Cameron-Martin Hilbert space; (ii) energy-Hilbert spaces corresponding to graph-Laplacians; and finally (iii) RKHSs generated by binomial coefficients. We show that the point-masses have finite $\mathscr{H}$-norm in cases (i) and (ii), but not in case (iii).

Definition 3.3. Let V be a countably infinite set, and let $\mathscr{F}(V)$ denote the set of all *finite* subsets of V.

  (i) For all $x \in V$, set

$$(3.1) \qquad\qquad K_x := K(\cdot, x) : V \to \mathbb{C}.$$

  (ii) Let $\mathscr{H} := \mathscr{H}(K)$ be the Hilbert-completion of the *span* $\{K_x : x \in V\}$, with respect to the inner product

$$(3.2) \qquad \left\langle \sum c_x K_x, \sum d_y K_y \right\rangle_{\mathscr{H}} := \sum \sum \overline{c_x} d_y K(x, y)$$

  $\mathscr{H}$ is then a reproducing kernel Hilbert space (RKHS), with the reproducing property:

$$(3.3) \qquad\qquad \varphi(x) = \langle K_x, \varphi \rangle_{\mathscr{H}}, \ \forall x \in V, \forall \varphi \in \mathscr{H}.$$

  (iii) If $F \in \mathscr{F}(V)$, set $\mathscr{H}_F = span\{K_x\}_{x \in F} \subset \mathscr{H}$, and let

$$(3.4) \qquad\qquad P_F := \text{the orthogonal projection onto } \mathscr{H}_F.$$

(iv) For $F \in \mathscr{F}(V)$, let K_F be the matrix given by

$$(3.5) \qquad K_F := (K(x,y))_{(x,y) \in F \times F} \, .$$

Lemma 3.4. *Let $F \in \mathscr{F}(V) =$ all finite subsets of V, $x_1 \in F$. Assume $\delta_{x_1} \in \mathscr{H}$. Then*

$$(3.6) \qquad P_F(\delta_{x_1})(\cdot) = \sum_{y \in F} \left(K_F^{-1} \delta_{x_1}\right)(y) \, K_y(\cdot) \, .$$

Proof. Show that

$$(3.7) \qquad \delta_{x_1} - \sum_{y \in F} \left(K_F^{-1} \delta_{x_1}\right)(y) \, K_y(\cdot) \in \mathscr{H}_F^{\perp} \, .$$

The remaining part follows easily from this. $\qquad \square$

Theorem 3.5. *Let V and K as above, i.e., we assume that V is countably infinite, and K is a p.d. kernel on $V \times V$. Let $\mathscr{H} = \mathscr{H}(K)$ be the corresponding RKHS. Fix $x_1 \in V$. Then the following three conditions are equivalent:*

(i) $\delta_{x_1} \in \mathscr{H}$;

(ii) $\exists C_{x_1} < \infty$ *such that for all $F \in \mathscr{F}(V)$, we have*

$$(3.8) \qquad |\xi(x_1)|^2 \leq C_{x_1} \sum_{F \times F} \sum \overline{\xi(x)} \xi(y) \, K(x,y) \, .$$

(iii) For $F \in \mathscr{F}(V)$, set

$$(3.9) \qquad K_F = (K(x,y))_{(x,y) \in F \times F}$$

as a $\#F \times \#F$ matrix. Then

$$(3.10) \qquad \sup_{F \in \mathscr{F}(V)} \left(K_F^{-1} \delta_{x_1}\right)(x_1) < \infty.$$

Proof. This is an application of Remark 1.2. Also see [JT15b] for details. $\quad \square$

Let D be an open domain in $\mathbb{R}^d$, and assume $V \subset D$ is countable and discrete subset of D. In this case, we shall consider two positive definite kernels: the

original kernel K on $D \times D$, and $K_V := K|_{V \times V}$ on $V \times V$ by restriction. Thus if $x \in V$, then

$$K_x^{(V)}(\cdot) = K(\cdot, x) : V \longrightarrow \mathbb{R}$$

is a function on V, while

$$K_x(\cdot) := K(\cdot, x) : D \longrightarrow \mathbb{R}$$

is a function on D. Further, let $\mathscr{H}$ and $\mathscr{H}_V$ be the associated RKHSs respectively.

Lemma 3.6. *$\mathscr{H}_V$ is isometrically embedded into $\mathscr{H}$ via the mapping*

$$J^{(V)} : K_x^{(V)} \longmapsto K_x, \; x \in V.$$

Proof. Assume $F \in \mathscr{F}(V)$, i.e., F is a finite subset of V. Let $\xi = \xi_F$ is a function on F, then

$$\left\| \sum\nolimits_{x \in F} \xi(x) K_x^{(V)} \right\|_{\mathscr{H}_V} = \left\| \sum\nolimits_{x \in F} \xi(x) K_x \right\|_{\mathscr{H}}.$$

Note that, by definition, the linear span of $\{ K_x^{(V)} \; ; \; x \in V \}$ is dense in $\mathscr{H}_V$, and the span of $\{ K_x \; ; \; x \in D \}$ is dense in $\mathscr{H}$. We conclude that $J^{(V)}$ extends uniquely to an isometry from $\mathscr{H}_V$ into $\mathscr{H}$. The desired result follows from this. □

In the examples below, we are concerned with cases of kernels $K : D \times D \to \mathbb{R}$ with restriction $K_V : V \times V \to \mathbb{R}$, where V is a countable discrete subset of D. Typically, for $x \in V$, we may have the restriction $\delta_x|_V$ contained in $\mathscr{H}_V$, but δ_x in not in $\mathscr{H}$.

3.1. Brownian Motion

Consider the covariance function of standard Brownian motion B_t, $t \in [0, \infty)$, i.e., a Gaussian process $\{ B_t \}$ with mean zero and covariance function

$$(3.11) \qquad\qquad K(s, t) := \mathbb{E}(B_s B_t) = s \wedge t.$$

Restrict to $V := \{ 0 \} \cup \mathbb{Z}_+ \subset D$, i.e., consider

$$K^{(V)} = K|_{V \times V}.$$

$\mathscr{H}(K)$: Cameron-Martin Hilbert space, consisting of functions $f \in L^2(\mathbb{R})$ s.t.

$$\int_0^\infty |f'(x)|^2 \, dx < \infty, \quad f(0) = 0.$$

$\mathscr{H}_V := \mathscr{H}(K_V)$. Note that

$$f \in \mathscr{H}(K_V) \Longleftrightarrow \sum_n |f(n) - f(n+1)|^2 < \infty.$$

We now show that the restriction of (3.11) to $V \times V$ for an ordered subset (we fix such a set V):

(3.12) $$V: \ 0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots$$

has the discrete mass property (Definition 3.1).

Set $\mathscr{H}_V = RKHS(K\big|_{V \times V})$,

(3.13) $$K_V(x_i, x_j) = x_i \wedge x_j.$$

We consider the set $F_n = \{x_1, x_2, \ldots, x_n\}$ of finite subsets of V, and

(3.14) $$K_n = K^{(F_n)} = \begin{bmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & x_2 & \cdots & x_2 \\ x_1 & x_2 & x_3 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} = (x_i \wedge x_j)_{i,j=1}^n.$$

We will show that condition (iii) in Theorem 3.5 holds for K_V.

Lemma 3.7.

(3.15) $$D_n = \det\left((x_i \wedge x_j)_{i,j=1}^n\right) = x_1 (x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_{n-1}).$$

Proof. Induction. In fact,

$$\begin{bmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & x_2 & \cdots & x_2 \\ x_1 & x_2 & x_3 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} \sim \begin{bmatrix} x_1 & 0 & 0 & \cdots & 0 \\ 0 & x_2 - x_1 & 0 & \cdots & 0 \\ 0 & 0 & x_3 - x_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & x_n - x_{n-1} \end{bmatrix},$$

unitary equivalence in finite dimensions. $\square$

Lemma 3.8. *Let*

(3.16) $$\zeta_{(n)} := K_n^{-1} (\delta_{x_1}) (\cdot)$$

so that

(3.17) $$\|P_{F_n} (\delta_{x_1})\|_{\mathscr{H}_V}^2 = \zeta_{(n)} (x_1).$$

Then,

$$\zeta_{(1)} (x_1) = \frac{1}{x_1}$$

$$\zeta_{(n)} (x_1) = \frac{x_2}{x_1 (x_2 - x_1)}, \quad for\ n = 2, 3, \dots,$$

and

$$\|\delta_{x_1}\|_{\mathscr{H}_V}^2 = \frac{x_2}{x_1 (x_2 - x_1)}.$$

Proof. A direct computation shows the $(1, 1)$ minor of the matrix K_n^{-1} is

(3.18) $$D'_{n-1} = \det \left((x_i \wedge x_j)_{i,j=2}^n \right)$$
$$= x_2 (x_3 - x_2) (x_4 - x_3) \cdots (x_n - x_{n-1})$$

and so

$$\zeta_{(1)} (x_1) = \frac{1}{x_1}, \quad \text{and}$$

$$\zeta_{(2)} (x_1) = \frac{x_2}{x_1 (x_2 - x_1)}$$

$$\zeta_{(3)} (x_1) = \frac{x_2 (x_3 - x_2)}{x_1 (x_2 - x_1) (x_3 - x_2)} = \frac{x_2}{x_1 (x_2 - x_1)}$$

$$\zeta_{(4)} (x_1) = \frac{x_2 (x_3 - x_2) (x_4 - x_3)}{x_1 (x_2 - x_1) (x_3 - x_2) (x_4 - x_3)} = \frac{x_2}{x_1 (x_2 - x_1)}$$

$$\vdots$$

The result follows from this. □

Corollary 3.9. $P_{F_n}(\delta_{x_1}) = P_{F_2}(\delta_{x_1})$, $\forall n \geq 2$. Therefore,

(3.19) $$\delta_{x_1} \in \mathcal{H}_V^{(F_2)} := span\{K_{x_1}^{(V)}, K_{x_2}^{(V)}\}$$

and

(3.20) $$\delta_{x_1} = \zeta_{(2)}(x_1) K_{x_1}^{(V)} + \zeta_{(2)}(x_2) K_{x_2}^{(V)}$$

where

$$\zeta_{(2)}(x_i) = K_2^{-1}(\delta_{x_1})(x_i), \quad i = 1, 2.$$

Specifically,

(3.21) $$\zeta_{(2)}(x_1) = \frac{x_2}{x_1(x_2 - x_1)}$$

(3.22) $$\zeta_{(2)}(x_2) = \frac{-1}{x_2 - x_1};$$

and

(3.23) $$\|\delta_{x_1}\|_{\mathcal{H}_V}^2 = \frac{x_2}{x_1(x_2 - x_1)}.$$

Proof. Note that

$$\zeta_n(x_1) = \|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2$$

and $\zeta_{(1)}(x_1) \leq \zeta_{(2)}(x_1) \leq \cdots$, since $F_n = \{x_1, x_2, \ldots, x_n\}$. In particular, $\frac{1}{x_1} \leq \frac{x_2}{x_1(x_2-x_1)}$, which yields (3.23). $\qquad\square$

Remark 3.10. We showed that $\delta_{x_1} \in \mathcal{H}_V$, $V = \{x_1 < x_2 < \cdots\} \subset \mathbb{R}_+$, with the restriction of $s \wedge t =$ the covariance kernel of Brownian motion. The same argument also shows that $\delta_{x_i} \in \mathcal{H}_V$ when $i > 1$.

Conclusions:

(3.24) $$\delta_{x_i} \in span\left\{K_{x_{i-1}}^{(V)}, K_{x_i}^{(V)}, K_{x_{i+1}}^{(V)}\right\}, \quad \text{and}$$

(3.25) $$\|\delta_{x_i}\|_{\mathcal{H}}^2 = \frac{x_{i+1} - x_{i-1}}{(x_i - x_{i-1})(x_{i+1} - x_i)}.$$

Details are left for the interested readers.

Corollary 3.11. *Let $V \subset \mathbb{R}_+$ be countable. If $x_a \in V$ is an accumulation point (from V), then $\|\delta_a\|_{\mathcal{H}_V} = \infty$.*

Example 3.12 (Sparse sample-points). Let $V = \{x_i\}_{i=1}^\infty$, where

$$x_i = \frac{i\,(i-1)}{2}, \quad i \in \mathbb{N}.$$

It follows that $x_{i+1} - x_i = i$, and so

$$\|\delta_{x_i}\|_{\mathscr{H}}^2 = \frac{x_{i+1} - x_i}{(x_i - x_{i-1})\,(x_{i+1} - x_i)} = \frac{2i - 1}{(i-1)\,i} \xrightarrow[i\to\infty]{} 0.$$

We conclude that $\|\delta_{x_i}\|_{\mathscr{H}} \xrightarrow[i\to\infty]{} 0$ if the set $V = \{x_i\}_{i=1}^\infty \subset \mathbb{R}_+$ is sparse.

Now, some general facts:

Lemma 3.13. *Let $K : V \times V \to \mathbb{C}$ be p.d., and let $\mathscr{H}$ be the corresponding RKHS. If $x_1 \in V$, and if δ_{x_1} has a representation as follows:*

$$(3.26) \qquad\qquad \delta_{x_1} = \sum_{y \in V} \zeta^{(x_1)}(y)\, K_y\,,$$

then

$$(3.27) \qquad\qquad \|\delta_{x_1}\|_{\mathscr{H}}^2 = \zeta^{(x_1)}(x_1)\,.$$

Proof. Substitute both sides of (3.26) into $\langle \delta_{x_1}, \cdot \rangle_{\mathscr{H}}$ where $\langle \cdot, \cdot \rangle_{\mathscr{H}}$ denotes the inner product in $\mathscr{H}$. $\square$

3.2. Brownian Bridge

Let $D := (0, 1) =$ the open interval $0 < t < 1$, and set

$$(3.28) \qquad\qquad K_{bridge}(s, t) := s \wedge t - st;$$

then (3.28) is the covariance function for the Brownian bridge $B_{bri}(t)$, i.e.,

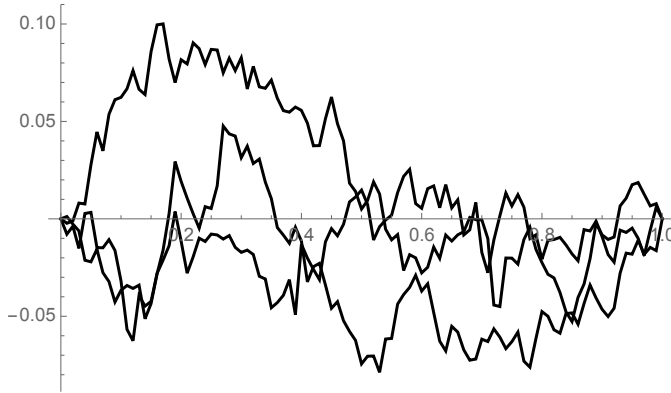$$(3.29) \qquad\qquad B_{bri}(0) = B_{bri}(1) = 0$$

Figure 3.1: Brownian bridge $B_{bri}(t)$, a simulation of three sample paths of the Brownian bridge.

$$(3.30) \qquad B_{bri}(t) = (1-t)\, B\left(\frac{t}{1-t}\right), \quad 0 < t < 1;$$

where $B(t)$ is Brownian motion; see Lemma 3.6.

The corresponding Cameron-Martin space is now

$$(3.31) \qquad \mathscr{H}_{bri} = \left\{ f \text{ on } [0,1]\,; f' \in L^2(0,1)\,, f(0) = f(1) = 0 \right\}$$

with

$$(3.32) \qquad \|f\|^2_{\mathscr{H}_{bri}} := \int_0^1 |f'(s)|^2\, ds < \infty.$$

If $V = \{x_i\}_{i=1}^\infty$, $x_1 < x_2 < \cdots < 1$, is the discrete subset of D, then we have for $F_n \in \mathscr{F}(V)$, $F_n = \{x_1, x_2, \ldots, x_n\}$,

$$(3.33) \qquad K_{F_n} = \left(K_{bridge}(x_i, x_j) \right)_{i,j=1}^n,$$

see (3.28), and

$$(3.34) \qquad \det K_{F_n} = x_1\,(x_2 - x_1) \cdots (x_n - x_{n-1})\,(1 - x_n).$$

As a result, we get $\delta_{x_i} \in \mathscr{H}_V^{(bri)}$ for all i, and

$$\|\delta_{x_i}\|^2_{\mathscr{H}_V^{(bri)}} = \frac{x_{i+1} - x_{i-1}}{(x_{i+1} - x_i)(x_i - x_{i-1})}.$$

Note $\lim_{x_i \to 1} \|\delta_{x_i}\|^2_{\mathscr{H}_V^{(bri)}} = \infty$.

3.3. Binomial RKHS

The purpose of the present subsection if to display a concrete RKHS $\mathscr{H}(K)$ in the discrete framework with the property that $\mathscr{H}(K)$ does not contain the Dirac masses δ_x. The RKHS in question is generated by the binomial coefficients, and it is relevant for a host of applications; see e.g., [JKS16, Dok14, Gal01].

Definition 3.14. Let $V = \mathbb{Z}_+ \cup \{0\}$; and

$$K_b(x, y) := \sum_{n=0}^{x \wedge y} \binom{x}{n} \binom{y}{n}, \quad (x, y) \in V \times V.$$

where $\binom{x}{n} = \frac{x(x-1)\cdots(x-n+1)}{n!}$ denotes the standard binomial coefficient from the binomial expansion.

Let $\mathscr{H} = \mathscr{H}(K_b)$ be the corresponding RKHS. Set

$$(3.35) \qquad\qquad e_n(x) = \begin{cases} \binom{x}{n} & \text{if } n \le x \\ 0 & \text{if } n > x. \end{cases}$$

Lemma 3.15 ([AJ14]).

(*i*) $e_n(\cdot) \in \mathscr{H}$, $n \in V$;

(*ii*) $\{e_n\}_{n \in V}$ *is an orthonormal basis (ONB) in the Hilbert space* $\mathscr{H}$.

(*iii*) *Given* $f \in \mathscr{F}unc(V)$; *then*

$$(3.36) \qquad\qquad f \in \mathscr{H} \iff \sum_{k=0}^{\infty} |\langle e_k, f \rangle_{\mathscr{H}}|^2 < \infty;$$

and, in this case,

$$\|f\|_{\mathscr{H}}^2 = \sum_{k=0}^{\infty} |\langle e_k, f \rangle_{\mathscr{H}}|^2.$$

(*iv*) *Set* $F_n = \{0, 1, 2, \ldots, n\}$, *and*

$$(3.37) \qquad\qquad P_{F_n} = \sum_{k=0}^{n} |e_k\rangle\langle e_k|$$

or equivalently

$$(3.38) \qquad P_{F_n} f = \sum_{k=0}^{n} \langle e_k, f \rangle_{\mathscr{H}} \, e_k \, .$$

Then formula (3.38) is well defined for all functions $f \in \mathscr{F}unc\,(V)$.

Fix $x_1 \in V$, then we shall apply Lemma 3.15 to the function $f_1 = \delta_{x_1}$ (in $\mathscr{F}unc\,(V)$).

Theorem 3.16. *We have*

$$\| P_{F_n} \left(\delta_{x_1} \right) \|_{\mathscr{H}}^2 = \sum_{k=x_1}^{n} \binom{k}{x_1}^2 \, .$$

The proof of the theorem will be subdivided in steps; see below.

Lemma 3.17 ([AJ14]).

(*i*) *For $\forall m, n \in V$, such that $m \leq n$, we have*

$$(3.39) \qquad \delta_{m,n} = \sum_{j=m}^{n} (-1)^{m+j} \binom{n}{j} \binom{j}{m} \, .$$

(*ii*) *For all $n \in \mathbb{Z}_+$, the inverse of the following lower triangle matrix is this: With (see Figure 3.2)*

$$(3.40) \qquad L_{xy}^{(n)} = \begin{cases} \binom{x}{y} & \text{if } y \leq x \leq n \\ 0 & \text{if } x < y \end{cases}$$

we have:

$$(3.41) \qquad \left(L^{(n)} \right)_{xy}^{-1} = \begin{cases} (-1)^{x-y} \binom{x}{y} & \text{if } y \leq x \leq n \\ 0 & \text{if } x < y. \end{cases}$$

Notation: The numbers in (3.41) are the entries of the matrix $\left(L^{(n)} \right)^{-1}$.

Proof. In rough outline, (ii) follows from (i). □

$$L^{(n)} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & & & & & \vdots & \vdots \\ 1 & 3 & 3 & 1 & \ddots & & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & 1 & 0 & & \vdots & \vdots \\ 1 & \cdots & \binom{x}{y} & \binom{x}{y+1} & \cdots & * & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & & 1 & 0 \\ 1 & \cdots & \binom{n}{y} & \binom{n}{y+1} & \cdots & \cdots & \cdots & \cdots & n & 1 \end{bmatrix}$$

Figure 3.2: The matrix L_n is simply a truncated Pascal triangle, arranged to fit into a lower triangular matrix.

Corollary 3.18. *Let K_b, $\mathscr{H}$, and $n \in \mathbb{Z}_+$ be as above with the lower triangle matrix L_n. Set*

$$(3.42) \qquad K_n(x,y) = K_b(x,y), \quad (x,y) \in F_n \times F_n,$$

i.e., an $(n+1) \times (n+1)$ matrix.

(i) *Then K_n is invertible with*

$$(3.43) \qquad K_n^{-1} = \left(L_n^{tr}\right)^{-1} (L_n)^{-1};$$

an (upper triangle) × (lower triangle) factorization.

(ii) *For the diagonal entries in the $(n+1) \times (n+1)$ matrix K_n^{-1}, we have:*

$$\langle x, K_n^{-1} x \rangle_{l^2} = \sum_{k=x}^{n} \binom{k}{x}^2$$

Conclusion: Since

$$(3.44) \qquad \|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2 = \langle x_1, K_n^{-1} x_1 \rangle_{\mathscr{H}}$$

for all $x_1 \in F_n$, we get

$$(3.45) \qquad \|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2 = \sum_{k=x_1}^{n} \binom{k}{x_1}^2$$

$$= 1 + \binom{x_1 + 1}{x_1}^2 + \binom{x_1 + 2}{x_1}^2 + \cdots + \binom{n}{x_1}^2;$$

and therefore,

$$\|\delta_{x_1}\|_{\mathscr{H}}^2 = \sum_{k=x_1}^{\infty} \binom{k}{x_1}^2 = \infty.$$

In other words, no δ_x is in $\mathscr{H}$.

References

[AACM11] Akram Aldroubi, Magalí Anastasio, Carlos Cabrelli, and Ursula Molter, *A dimension reduction scheme for the computation of optimal unions of subspaces*, Sampl. Theory Signal Image Process. **10** (2011), no. 1-2, 135–150.

[ABDdS93] Daniel Alpay, Vladimir Bolotnikov, Aad Dijksma, and Henk de Snoo, *On some operator colligations and associated reproducing kernel Hilbert spaces*, in: Operator Extensions, Interpolation of Functions and Related Topics, Oper. Theory Adv. Appl., Vol. 61, Birkhäuser, Basel, 1993, pp. 1–27.

[ACH+10] Akram Aldroubi, Carlos Cabrelli, Christopher Heil, Keri Kornelson, and Ursula Molter, *Invariance of a shift-invariant space*, J. Fourier Anal. Appl. **16** (2010), no. 1, 60–75.

[AD92] Daniel Alpay and Harry Dym, *On reproducing kernel spaces, the Schur algorithm, and interpolation in a general class of domains*, in: Operator Theory and Complex Analysis (Sapporo, 1991), Oper. Theory Adv. Appl., Vol. 59, Birkhäuser, Basel, 1992, pp. 30–77.

[AD93] Daniel Alpay and Harry Dym, *On a new class of structured reproducing kernel spaces*, J. Funct. Anal. **111** (1993), no. 1, 1–28.

[AJ14] Daniel Alpay and Palle Jorgensen, *Reproducing kernel Hilbert spaces generated by the binomial coefficients*, Illinois J. Math. **58** (2014), no. 2, 471–495.

[AJ15] Daniel Alpay and Palle Jorgensen, *Spectral theory for Gaussian processes: reproducing kernels, boundaries, and L^2-wavelet generators with fractional scales*, Numer. Funct. Anal. Optim. **36** (2015), no. 10, 1239–1285.

[AJSV13] Daniel Alpay, Palle Jorgensen, Ron Seager, and Dan Volok, *On discrete analytic functions: products, rational functions and reproducing kernels*, J. Appl. Math. Comput. **41** (2013), no. 1-2, 393–426.

[AJV14] Daniel Alpay, Palle Jorgensen, and Dan Volok, *Relative reproducing kernel Hilbert spaces*, Proc. Amer. Math. Soc. **142** (2014), no. 11, 3889–3895.

[Aro43] Nachman Aronszajn, *La théorie des noyaux reproduisants et ses applications. I*, Proc. Cambridge Philos. Soc. **39** (1943), 133–153.

[Aro48] Nachman Aronszajn, *Reproducing and pseudo-reproducing kernels and their application to the partial differential equations of physics*, Studies in Partial Differential Equations. Technical report 5, preliminary note, Harvard University, Graduate School of Engineering., 1948.

[BCL11] Radu Balan, Pete Casazza, and Zeph Landau, *Redundancy for localized frames*, Israel J. Math. **185** (2011), 445–476.

[BDP05] Stefan Bildea, Dorin Ervin Dutkay, and Gabriel Picioroaga, *MRA super-wavelets*, New York J. Math. **11** (2005), 1–19.

[CD93] Albert Cohen and Ingrid Daubechies, *Non-separable bidimensional wavelet bases*, Revista Matemática Iberoamericana (1993), 51–137.

[Chr14] Ole Christensen, *A short introduction to frames, Gabor systems, and wavelet systems*, Azerb. J. Math. **4** (2014), no. 1, 25–39.

[CS02] Felipe Cucker and Steve Smale, *On the mathematical foundations of learning*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 1, 1–49.

[Dok14] Nikolai Dokuchaev, *On strong causal binomial approximation for stochastic processes*, Discrete Contin. Dyn. Syst. Ser. B **19** (2014), no. 6, 1549–1562.

[Dut06] Dorin Ervin Dutkay, *Low-pass filters and representations of the baumslag solitar group*, Transactions of the American Mathematical Society **358** (2006), no. 12, 5271–5292.

[FJKO05] Matthew Fickus, Brody D. Johnson, Keri Kornelson, and Kasso A. Okoudjou, *Convolutional frames and the frame potential*, Appl. Comput. Harmon. Anal. **19** (2005), no. 1, 77–91.

[Gal01] Roza Galeeva, *Binomial trees as dynamical systems*, Phys. A **292** (2001), no. 1-4, 519–535.

[Hid80] Takeyuki Hida, Brownian Motion, Applications of Mathematics, Vol. 11, Springer-Verlag, New York, (1980), Translated from the Japanese by the author and T. P. Speed.

[HKL$^+$14] S. Haeseler, M. Keller, D. Lenz, J. Masamune, and M. Schmidt, *Global properties of Dirichlet forms in terms of Green's formula*, arXiv e-prints (2014).

[HKLW07] Deguang Han, Keri Kornelson, David Larson, and Eric Weber, Frames for Undergraduates, Student Mathematical Library, Vol. 40, American Mathematical Society, Providence, RI, (2007).

[HLS85] J. M. Harrison, H. J. Landau, and L. A. Shepp, *The stationary distribution of reflected Brownian motion in a planar region*, Ann. Probab. **13** (1985), no. 3, 744–757.

[HN14] Haakan Hedenmalm and Pekka J. Nieminen, *The Gaussian free field and Hadamard's variational formula*, Probab. Theory Related Fields **159** (2014), no. 1-2, 61–73.

[HQKL10] Minh Ha Quang, Sung Ha Kang, and Triet M. Le, *Image and video colorization using vector-valued reproducing kernel Hilbert spaces*, J. Math. Imaging Vision **37** (2010), no. 1, 49–65.

[JKS16] Norman Josephy, Lucia Kimball, and Victoria Steblovskaya, *Optimal hedging in an extended binomial market under transaction costs*, Quant. Finance **16** (2016), no. 5, 763–776.

[JT15a] P. Jorgensen and F. Tian, *Infinite weighted graphs with bounded resistance metric*, arXiv e-prints (2015).

[JT15b] Palle Jorgensen and Feng Tian, *Discrete reproducing kernel Hilbert spaces: sampling and distribution of Dirac-masses*, J. Mach. Learn. Res. **16** (2015), 3079–3114.

[JT16] Palle Jorgensen and Feng Tian, *Positive definite kernels and boundary spaces.*, Adv. Oper. Theory **1** (2016), no. 1, 123–133 (English).

[Kat04] Yitzhak Katznelson, An Introduction to Harmonic Analysis, third ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, (2004).

[KH11] Sanjeev Kulkarni and Gilbert Harman, An Elementary Introduction to Statistical Learning Theory, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, (2011).

[Lan60] H. J. Landau, *On the recovery of a band-limited signal, after instantaneous companding and subsequent band limiting*, Bell System Tech. J. **39** (1960), 351–364.

[Lan64] H. J. Landau, *A sparse regular sequence of exponentials closed on large sets*, Bull. Amer. Math. Soc. **70** (1964), 566–569.

[Lan67] H. J. Landau, *Necessary density conditions for sampling and interpolation of certain entire functions*, Acta Math. **117** (1967), 37–52.

[LB04] Yi Lin and Lawrence D. Brown, *Statistical properties of the method of regularization with periodic Gaussian reproducing kernel*, Ann. Statist. **32** (2004), no. 4, 1723–1743.

[LLSB84] H. J. Landau, B. F. Logan, L. A. Shepp, and N. Bauman, *Diffusion, cell mobility, and bandlimited functions*, SIAM Journal on Applied Mathematics **44** (1984), no. 6, 1232–1245.

[LP61] H. J. Landau and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty. II*, Bell System Tech. J. **40** (1961), 65–84.

[LP11] Sneh Lata and Vern Paulsen, *The Feichtinger conjecture and reproducing kernel Hilbert spaces*, Indiana Univ. Math. J. **60** (2011), no. 4, 1303–1317.

[Nel57] Edward Nelson, *Kernel functions and eigenfunction expansions*, Duke Math. J. **25** (1957), 15–27.

[NSW11] P. Niyogi, S. Smale, and S. Weinberger, *A topological view of unsupervised learning from noisy data*, SIAM J. Comput. **40** (2011), no. 3, 646–663.

[OST13] Kasso A. Okoudjou, Robert S. Strichartz, and Elizabeth K. Tuley, *Orthogonal polynomials on the Sierpinski gasket*, Constr. Approx. **37** (2013), no. 3, 311–340.

[Pau02] Vern Paulsen, Completely Bounded Maps and Operator Algebras, Cambridge Studies in Advanced Mathematics, Vol. 78, Cambridge University Press, Cambridge, (2002).

[RW06] Carl Edward Rasmussen and Christopher K. I. Williams, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, (2006).

[Sei04] Kristian Seip, Interpolation and Sampling in Spaces of Analytic Functions, University Lecture Series, Vol. 33, American Mathematical Society, Providence, RI, (2004).

[SRB+10] S. Smale, L. Rosasco, J. Bouvrie, A. Caponnetto, and T. Poggio, *Mathematics of the neural response*, Found. Comput. Math. **10** (2010), no. 1, 67–91.

[SS01] Bernhard Schlkopf and Alexander J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning), 1st ed., The MIT Press, (2001).

[SS13] Oded Schramm and Scott Sheffield, *A contour line of the continuum Gaussian free field*, Probab. Theory Related Fields **157** (2013), no. 1-2, 47–80.

[ST12] Robert S. Strichartz and Alexander Teplyaev, *Spectral analysis on infinite Sierpiński fractafolds*, J. Anal. Math. **116** (2012), 255–297.

[STC04] John Shawe-Taylor and Nello Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, (2004).

[Str10] Robert S. Strichartz, *Transformation of spectra of graph Laplacians*, Rocky Mountain J. Math. **40** (2010), no. 6, 2037–2062.

[SY06] Steve Smale and Yuan Yao, *Online learning algorithms*, Found. Comput. Math. **6** (2006), no. 2, 145–170.

[SZ05] Steve Smale and Ding-Xuan Zhou, *Shannon sampling. II. Connections to learning theory*, Appl. Comput. Harmon. Anal. **19** (2005), no. 3, 285–302.

[SZ09] Steve Smale and Ding-Xuan Zhou, *Online learning with Markov sampling*, Anal. Appl. (Singap.) **7** (2009), no. 1, 87–113.

[Vul13] Mirjana Vuletić, *The Gaussian free field and strict plane partitions*, in: 25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013), Discrete Math. Theor. Comput. Sci. Proc., AS, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, (2013), pp. 1041–1052.

[ZXZ12] Haizhang Zhang, Yuesheng Xu, and Qinghui Zhang, *Refinement of operator-valued reproducing kernels*, J. Mach. Learn. Res. **13** (2012), 91–136.

Department of Mathematics, The University of Iowa
Iowa City, IA 52242-1419, USA
E-mail address: palle-jorgensen@uiowa.edu

Mathematical Reviews
416 4th Street Ann Arbor, MI 48103-4816, USA
E-mail address: jft@ams.org