Journal of Machine Learning Research 16 (2015) 3079-3114 Submitted 1/15; Revised 3/15; Published 12/15

Discrete Reproducing Kernel Hilbert Spaces: Sampling and Distribution of Dirac-masses

Palle Jorgensen PALLE-JORGENSEN@UIOWA.EDU
Department of Mathematics
The University of Iowa
Iowa City, IA 52242-1419, U.S.A.

Feng Tian TIANF@TRINE.EDU
Department of Mathematics, Informatics, and Cybersecurity
Trine University
Angola, IN 46703, U.S.A.

Editor: John Shawe-Taylor

Abstract

We study reproducing kernels, and associated reproducing kernel Hilbert spaces (RKHSs) $\mathscr{H}$ over infinite, discrete and countable sets V. In this setting we analyze in detail the distributions of the corresponding Dirac point-masses of V. Illustrations include certain models from neural networks: An Extreme Learning Machine (ELM) is a neural network-configuration in which a hidden layer of weights are randomly sampled, and where the object is then to compute resulting output. For RKHSs $\mathscr{H}$ of functions defined on a prescribed countable infinite discrete set V, we characterize those which contain the Dirac masses δ_x for all points x in V. Further examples and applications where this question plays an important role are: (i) discrete Brownian motion-Hilbert spaces, i.e., discrete versions of the Cameron-Martin Hilbert space; (ii) energy-Hilbert spaces corresponding to graph-Laplacians where the set V of vertices is then equipped with a resistance metric; and finally (iii) the study of Gaussian free fields.

Keywords: Gaussian reproducing kernel Hilbert spaces, sampling in discrete systems, resistance metric, graph Laplacians, discrete Green's functions

1. Introduction

A reproducing kernel Hilbert space (RKHS) is a Hilbert space $\mathscr{H}$ of functions on a prescribed set, say V, 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 $x \in V$, the point-evaluation for functions $f \in \mathscr{H}$, $f(x)$ must then be given as a $\mathscr{H}$-inner product of f and a vector k_x, in $\mathscr{H}$; called the kernel.

The RKHSs have been studied extensively since the pioneering papers by Aronszajn (1943; 1948). 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 (Nelson, 1957; Haeseler et al., 2014). Other applications include engineering, physics, machine-learning theory (Kulkarni and Harman, 2011; Smale and Zhou, 2009; Cucker and Smale, 2002), stochastic processes (Alpay and Dym, 1993; Alpay et al., 1993; Alpay and Dym, 1992; Alpay et al., 2013, 2014), numerical analysis, and more (Lin and Brown, 2004; Ha Quang et al., 2010; Zhang et al., 2012; Lata and Paulsen, 2011; Vuletić, 2013; Schramm and Sheffield, 2013; Hedenmalm and Nieminen, 2014; Shawe-Taylor and Cristianini, 2004; Schlkopf and Smola, 2001). But the literature so far has focused on the theory of kernel functions defined on continuous domains, either domains in Euclidean space, or complex domains in one or more variables. For these cases, the Dirac δ_x distributions do not have finite $\mathscr{H}$-norm. But for RKHSs over discrete point distributions, it is reasonable to expect that the Dirac δ_x functions will in fact have finite $\mathscr{H}$-norm.

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 (Rasmussen and Williams, 2006), 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.

Here we consider the discrete case, i.e., RKHSs of functions defined on a prescribed countable infinite discrete set V. We are concerned with a characterization of those RKHSs $\mathscr{H}$ which contain the Dirac masses δ_x for all points $x \in V$. 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).

Our setting is a given positive definite function k on $V \times V$, where V is discrete. We study the corresponding RKHS $\mathscr{H}\,(= \mathscr{H}\,(k))$ in detail. Our main results are Theorems 1, 2, and 3 which give explicit answers to the question of which point-masses from V are in $\mathscr{H}$. Applications include Corollaries 29, 41, 46, 48, 52, and 53.

The paper is organized as follows: Section 2 leads up to our characterization (Theorem 1) of point-masses which have finite $\mathscr{H}$-norm. It is applied in Sections 3 and 4 to a variety of classes of discrete RKHSs. Section 3 deals with samples from Brownian motion, and from the Brownian bridge process, and binomial kernels, and with kernels on sets $V \times V$ which arise as restrictions to sample-points. Section 4 covers the case of infinite network of resistors. By this we mean an infinite graph with assigned resistors on its edges. In this family of examples, the associated RKHSs vary with the assignment of resistors on the edges in G, and are computed explicitly from a resulting energy form. Our result Corollary 46 states that, for the network models, all point-masses have finite energy. Furthermore, we compute the value, and we study V as a metric space w.r.t. the corresponding resistance metric. These results, in turn, have direct implications (Corollaries 48, 52 and 55) for the family of Gaussian free fields associated with our infinite network models.

A positive definite kernel k is said to be *universal* (Steinwart, 2002; Caponnetto et al., 2008) if, every continuous function, on a compact subset of the input space, can be uniformly approximated by sections of the kernel, i.e., by continuous functions in the RKHS. In Theorem 3 we show that for the RKHSs from kernels k_c in electrical network G of resistors, this universality holds. The metric in this case is the resistance metric on the vertices of G, determined by the assignment of a conductance function c on the edges in G.

Infinite vs finite graphs. We study “large weighted graphs” (vertices V, edges E, and weights as functions assigned on the edges E), and our motivation derives from learning where “learning” is understood broadly to include (machine) learning of suitable probability distribution, i.e., meaning learning from samples of training data. Other applications of an analysis of weighted graphs include statistical mechanics, such as infinite spin models, and large digital networks. It is natural to ask then how one best approaches analysis on "large" systems. We propose an analysis via infinite weighted graphs. This is so even if some of the questions in learning theory may in fact refer to only "large" finite graphs.

One reason for this (among others) is that statistical features in such an analysis are best predicted by consideration of probability spaces corresponding to measures on infinite sample spaces. Moreover the latter are best designed from consideration of infinite weighted graphs, as opposed to their finite counterparts. Examples of statistical features which are relevant even for finite samples is long-range order; i.e., the study of correlations between distant sites (vertices), and related phase-transitions, e.g., sign-flips at distant sites. In designing efficient learning models, it is important to understand the possible occurrence of unexpected long-range correlations; e.g., correlations between distant sites in a finite sample.

A second reason for the use of infinite sample-spaces is their use in designing efficient sampling procedures. The interesting solutions will often occur first as vectors in an infinite-dimensional reproducing-kernel Hilbert space RKHS. Indeed, such RKHSs serve as powerful tools in the solution of a kernel-optimization problems with penalty terms. Once an optimal solution is obtained in infinite dimensions, one may then proceed to study its restrictions to suitably chosen finite subgraphs.

In general when reproducing kernels and their Hilbert spaces are used, one ends up with functions on a suitable set, and so far we feel that the dichotomy discrete vs continuous has not yet received sufficient attention. After all, a choice of sampling points in relevant optimization models based on kernel theory suggests the need for a better understanding of point masses as they are accounted for in the RKHS at hand. In broad outline, this is a leading theme in our paper.

2. Discrete RKHSs

Definition 1 *Let V be a countable and infinite set, and $\mathscr{F}(V)$ the set of all* finite *subsets of V. A function $k : V \times V \to \mathbb{C}$ is said to be* positive definite*, if*

$$\sum\sum_{(x,y)\in F\times F} k(x,y)\,\overline{c_x}c_y \geq 0 \tag{1}$$

holds for all coefficients $\{c_x\}_{x\in F} \subset \mathbb{C}$, and all $F \in \mathscr{F}(V)$.

Definition 2 *Fix a set V, countable infinite.*

1. *For all $x \in V$, set*
$$k_x := k(\cdot, x) : V \to \mathbb{C} \tag{2}$$
as a function on V.

2. *Let $\mathscr{H} := \mathscr{H}(k)$ be the Hilbert-completion of the span $\{k_x : x \in V\}$, with respect to the inner product*
$$\left\langle \sum c_x k_x, \sum d_y k_y \right\rangle_{\mathscr{H}} := \sum\sum \overline{c_x} d_y k(x,y) \tag{3}$$

modulo the subspace of functions of zero $\mathscr{H}$-norm. $\mathscr{H}$ is then a reproducing kernel Hilbert space (HKRS), with the reproducing property:

$$\langle k_x, \varphi \rangle_{\mathscr{H}} = \varphi(x), \ \forall x \in V, \ \forall \varphi \in \mathscr{H}. \tag{4}$$

Note. *The summations in (3) are all finite. Starting with finitely supported summations in (3), the RKHS $\mathscr{H} = \mathscr{H}(k)$ is then obtained by Hilbert space completion. We use physicists' convention, so that the inner product is conjugate linear in the first variable, and linear in the second variable.*

3. *If $F \in \mathscr{F}(V)$, set $\mathscr{H}_F$ = closed span$\{k_x\}_{x \in F} \subset \mathscr{H}$, (closed is automatic if F is finite.) And set*

$$P_F := \text{the orthogonal projection onto } \mathscr{H}_F. \tag{5}$$

4. *For $F \in \mathscr{F}(V)$, set*

$$K_F := (k(x,y))_{(x,y) \in F \times F} \tag{6}$$

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

Remark 3 *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?" Now, before completion, the functions are as specified in Definition 2, but 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 (Okoudjou et al., 2013; Strichartz and Teplyaev, 2012; Strichartz, 2010) (for real domains, or for fractals), band-limited L^2 functions (from signal analysis), and Cameron-Martin Hilbert spaces from Gaussian processes (in continuous time domain).*

Our focus here is on discrete analogues of the classical RKHSs from real or complex analysis. These discrete RKHSs in turn are dictated by applications, and their features are quite different from those of their continuous counterparts.

Definition 4 *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, $\delta_x(y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$, i.e., the Dirac mass at $x \in V$.*

Lemma 5 *Let $F \in \mathscr{F}(V)$, $x_1 \in F$. Assume $\delta_{x_1} \in \mathscr{H}$. Then*

$$P_F(\delta_{x_1})(\cdot) = \sum_{y \in F} \left(K_F^{-1} \delta_{x_1}\right)(y)\, k_y(\cdot). \tag{7}$$

Proof Show that

$$\delta_{x_1} - \sum_{y \in F} \left(K_F^{-1} \delta_{x_1}\right)(y)\, k_y(\cdot) \in \mathscr{H}_F^{\perp}. \tag{8}$$

The remaining part follows easily from this.

(The notation $(\mathscr{H}_F)^{\perp}$ stands for orthogonal complement, also denoted $\mathscr{H} \ominus \mathscr{H}_F = \left\{\varphi \in \mathscr{H} \;\middle|\; \langle f, \varphi\rangle_{\mathscr{H}} = 0,\ \forall f \in \mathscr{H}_F\right\}$.) ■

Lemma 6 *Using Dirac's bra-ket, and ket-bra notation (for rank-one operators), the orthogonal projection onto* $\mathscr{H}_F$ *is*

$$P_F = \sum_{y\in F} \left|k_y\right\rangle\!\left\langle k_y^*\right|; \tag{9}$$

where

$$k_x^* := \sum_{y\in F} \left(K_F^{-1}\right)_{yx} k_y \tag{10}$$

is the dual vector to k_x*, for all* $x \in V$*.*

Proof Let k_x^* be specified as in (10), then

$$\begin{aligned}
\langle k_x^*, k_z\rangle_{\mathscr{H}} &= \sum_{y\in F} \left\langle \left(K_F^{-1}\right)_{yx} k_y, k_z \right\rangle_{\mathscr{H}} \\
&= \sum_{y\in F} \left(K_F^{-1}\right)_{xy} \langle k_y, k_z\rangle_{\mathscr{H}} \\
&= \sum_{y\in F} \left(K_F^{-1}\right)_{xy} \left(K_F\right)_{yz} = \delta_{x,z},
\end{aligned}$$

i.e., k_x^* is the dual vector to k_x, for all $x \in V$.

For $f \in \mathscr{H}$, and $F \in \mathscr{F}(V)$, we have

$$\begin{aligned}
\sum_{y\in F} \left|k_y\right\rangle\!\left\langle k_y^*\right| f &= \sum_{y\in F} \left\langle k_y^*, f\right\rangle_{\mathscr{H}} k_y \\
&= \sum\sum_{(y,z)\in F\times F} \left(K_F^{-1}\right)_{z,y} \langle k_z, f\rangle_{\mathscr{H}} \\
&= P_F f.
\end{aligned}$$

This yields the orthogonal projection realized as stated in (9).

Now, applying (9) to δ_{x_1}, we get

$$\begin{aligned}
P_F\left(\delta_{x_1}\right) &= \sum_{y\in F} \left\langle k_y^*, \delta_{x_1}\right\rangle_{\mathscr{H}} k_y \\
&= \sum_{y\in F} \left(\sum_{z\in F} \left(K_F^{-1}\right)_{yz} \langle k_z, \delta_{x_1}\rangle_{\mathscr{H}}\right) k_y \\
&= \sum_{y\in F} \left(\sum_{z\in F} \left(K_F^{-1}\right)_{yz} \delta_{x_1}(z)\right) k_y \\
&= \sum_{y\in F} \left(K_F^{-1}\delta_{x_1}\right)(y)\, k_y,
\end{aligned}$$

where

$$\left(K_F^{-1}\delta_{x_1}\right)(y) := \sum_{z\in F}\left(K_F^{-1}\right)_{yz}\delta_{x_1}(z).$$

This verifies (7). ∎

Remark 7 *Note a slight abuse of notations: We make formally sense of the expressions for $P_F(\delta_x)$ in (7) even in the case when δ_x might not be in $\mathscr{H}$. For all finite F, we showed that $P_F(\delta_x) \in \mathscr{H}$. But for δ_x be in $\mathscr{H}$, we must have the additional boundedness assumption (18) satisfied; see Theorem 1.*

Lemma 8 *Let $F \in \mathscr{F}(V)$, $x_1 \in F$, then*

$$\left(K_F^{-1}\delta_{x_1}\right)(x_1) = \left\|P_F\delta_{x_1}\right\|_{\mathscr{H}}^2. \tag{11}$$

Proof Setting $\zeta^{(F)} := K_F^{-1}(\delta_{x_1})$, we have

$$P_F(\delta_{x_1}) = \sum_{y\in F}\zeta^{(F)}(y)\,k_F(\cdot, y)$$

and for all $z \in F$,

$$\begin{aligned}\underbrace{\sum_{z\in F}\zeta^{(F)}(z)\,P_F(\delta_{x_1})(z)}_{\zeta^{(F)}(x_1)} &= \sum_F\sum_F\zeta^{(F)}(z)\,\zeta^{(F)}(y)\,k_F(z,y) \qquad (12)\\ &= \left\|P_F\delta_{x_1}\right\|_{\mathscr{H}}^2.\end{aligned}$$

By Lemma 6, the LHS of (12) is given by

$$\begin{aligned}\left\|P_F\delta_{x_1}\right\|_{\mathscr{H}}^2 &= \left\langle P_F\delta_{x_1}, \delta_{x_1}\right\rangle_{\mathscr{H}}\\ &= \sum_{y\in F}\left(K_F^{-1}\delta_{x_1}\right)(y)\left\langle k_y, \delta_{x_1}\right\rangle_{\mathscr{H}}\\ &= \left(K_F^{-1}\delta_{x_1}\right)(x_1) = K_F^{-1}(x_1, x_1).\end{aligned}$$

∎

Corollary 9 *If $\delta_{x_1} \in \mathscr{H}$ (see Theorem 1), then*

$$\sup_{F\in\mathscr{F}(V)}\left(K_F^{-1}\delta_{x_1}\right)(x_1) = \left\|\delta_{x_1}\right\|_{\mathscr{H}}^2. \tag{13}$$

The following condition is satisfied in some examples, but not all:

Corollary 10 $\exists F \in \mathscr{F}(V)$ *s.t.* $\delta_{x_1} \in \mathscr{H}_F \Longleftrightarrow$

$$K_{F'}^{-1}(\delta_{x_1})(x_1) = K_F^{-1}(\delta_{x_1})(x_1)$$

for all $F' \supset F$.

Corollary 11 (Monotonicity) *If F and F' are in $\mathscr{F}(V)$ and $F \subset F'$, then*

$$\left(K_{F}^{-1}\delta_{x_1}\right)(x_1) \leq \left(K_{F'}^{-1}\delta_{x_1}\right)(x_1) \tag{14}$$

and

$$\lim_{F \nearrow V}\left(K_{F}^{-1}\delta_{x_1}\right)(x_1) = \left\|\delta_{x_1}\right\|_{\mathscr{H}}^{2}. \tag{15}$$

Proof By (11),

$$\left(K_{F}^{-1}\delta_{x_1}\right)(x_1) = \left\|P_F\delta_{x_1}\right\|_{\mathscr{H}}^{2}.$$

Since $\mathscr{H}_F \subset \mathscr{H}_{F'}$, we have $P_F P_{F'} = P_F$, so

$$\left\|P_F\delta_{x_1}\right\|_{\mathscr{H}}^{2} = \left\|P_F P_{F'}\delta_{x_1}\right\|_{\mathscr{H}}^{2} \leq \left\|P_{F'}\delta_{x_1}\right\|_{\mathscr{H}}^{2}$$

i.e.,

$$\left(K_{F}^{-1}\delta_{x_1}\right)(x_1) \leq \left(K_{F'}^{-1}\delta_{x_1}\right)(x_1).$$

So (14) follows; and the limit in (15) is monotone. ■

Theorem 1 *Given V, $k : V \times V \to \mathbb{R}$ positive definite (p.d.). Let $\mathscr{H} = \mathscr{H}(k)$ be the corresponding RKHS. Assume V is countable and infinite. Then the following three conditions (i)-(iii) are equivalent; $x_1 \in V$ is fixed:*

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

(ii) $\exists C_{x_1} < \infty$ *such that for all $F \in \mathscr{F}(V)$, the following estimate holds:*

$$\left|\xi(x_1)\right|^2 \leq C_{x_1} \sum\sum_{F\times F} \overline{\xi(x)}\xi(y)\,k(x,y) \tag{16}$$

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

$$K_F = \left(k(x,y)\right)_{(x,y)\in F\times F} \tag{17}$$

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

$$\sup_{F\in\mathscr{F}(V)}\left(K_{F}^{-1}\delta_{x_1}\right)(x_1) < \infty. \tag{18}$$

Proof (i)⇒(ii) For $\xi \in l^2(F)$, set

$$h_\xi = \sum_{y\in F}\xi(y)\,k_y(\cdot) \in \mathscr{H}_F.$$

Then $\left\langle \delta_{x_1}, h_\xi\right\rangle_{\mathscr{H}} = \xi(x_1)$ for all ξ.

Since $\delta_{x_1} \in \mathscr{H}$, then by Schwarz:

$$\left|\left\langle \delta_{x_1}, h_\xi\right\rangle_{\mathscr{H}}\right|^2 \leq \left\|\delta_{x_1}\right\|_{\mathscr{H}}^{2} \sum\sum_{F\times F}\overline{\xi(x)}\xi(y)\,k(x,y). \tag{19}$$

But $\langle \delta_{x_1}, k_y \rangle_{\mathscr{H}} = \delta_{x_1,y} = \begin{cases} 1 & y = x_1 \\ 0 & y \neq x_1 \end{cases}$; hence $\langle \delta_{x_1}, h_\xi \rangle_{\mathscr{H}} = \xi(x_1)$, and so (19) implies (16).

(ii)⇒(iii) Recall the matrix

$$K_F := (\langle k_x, k_y \rangle)_{(x,y)\in F\times F}$$

as a linear operator $l^2(F) \to l^2(F)$, where

$$(K_F\varphi)(x) = \sum_{y\in F} K_F(x,y)\,\varphi(y),\ \varphi \in l^2(F). \tag{20}$$

By (16), we have

$$\ker(K_F) \subset \left\{\varphi \in l^2(F) : \varphi(x_1) = 0\right\}. \tag{21}$$

Equivalently,

$$\ker(K_F) \subset \{\delta_{x_1}\}^{\perp} \tag{22}$$

and so $\delta_{x_1}\Big|_F \in \ker(K_F)^{\perp} = \operatorname{ran}(K_F)$, and $\exists\ \zeta^{(F)} \in l^2(F)$ s.t.

$$\delta_{x_1}\Big|_F = \underbrace{\sum_{y\in F} \zeta^{(F)}(y)\,k(\cdot,y)}_{=:h_F}. \tag{23}$$

Claim. $P_F(\delta_{x_1}) = h_F$, where P_F = projection onto $\mathscr{H}_F$; see (5) and Lemma 5. (See Figure 1.) Indeed, we only need to prove that $\delta_{x_1} - h_F \in \mathscr{H} \ominus \mathscr{H}_F$, i.e.,

$$\langle \delta_{x_1} - h_F, k_z \rangle_{\mathscr{H}} = 0,\ \forall z \in F. \tag{24}$$

But, by (23),

$$\text{LHS}_{(24)} = \delta_{x_1,z} - \sum_{y\in F} k(z,y)\,\zeta^{(F)}(y) = 0.$$

This proves the claim.

If $F \subset F'$, $F, F' \in \mathscr{F}(V)$, then $\mathscr{H}_F \subset \mathscr{H}_{F'}$, and $P_F P_{F'} = P_F$ by easy facts for projections. Hence

$$\|P_F\delta_{x_1}\|^2_{\mathscr{H}} \leq \|P_{F'}\delta_{x_1}\|^2_{\mathscr{H}}, \quad h_F := P_F(\delta_{x_1})$$

and

$$\lim_{F\nearrow V} \|\delta_{x_1} - h_F\|_{\mathscr{H}} = 0.$$

(iii)⇒(i) Follows from Lemma 8 and Corollary 9. ■

Corollary 12 *The numbers* $\left(\zeta^{(F)}(y)\right)_{y\in F}$ *in (23) satisfies*

$$\zeta^{(F)}(x_1) = \sum\sum_{(y,z)\in F\times F} \zeta^{(F)}(y)\,\zeta^{(F)}(z)\,k(y,z). \tag{25}$$

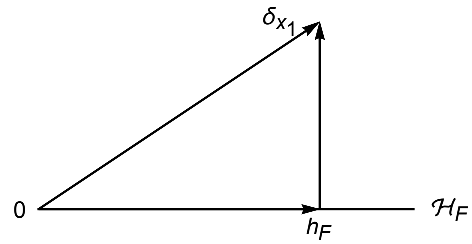


Figure 1: $h_F := P_F(\delta_{x_1})$

Proof Multiply (23) by $\zeta^{(F)}(z)$ and carry out the summation. ■

Remark 13 *To see that (23) is a solution to a linear algebra problem, with $F = \{x_i\}_{i=1}^n$, note that (23) $\Longleftrightarrow$*

$$\begin{bmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_n) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_n) \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ k(x_n,x_1) & k(x_n,x_2) & \cdots & k(x_n,x_n) \end{bmatrix} \begin{bmatrix} \zeta^{(F)}(x_1) \\ \zeta^{(F)}(x_2) \\ \vdots \\ \zeta^{(F)}(x_{n-1}) \\ \zeta^{(F)}(x_n) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{26}$$

We now resume the general case of k given and positive definite on $V \times V$.

Corollary 14 *We have*

$$\zeta^{(F)}(x_1) = \|P_F(\delta_{x_1})\|^2_{\mathscr{H}} \tag{27}$$

where

$$P_F(\delta_{x_1}) = \sum_{y \in F} \zeta^{(F)}(y)\, k_y(\cdot) \tag{28}$$

and

$$\zeta^{(F)} = K_N^{-1}(\delta_{x_1}), \quad N := \#F. \tag{29}$$

Proof It follows from (26) that

$$\sum_j k(x_i,x_j)\,\zeta^{(F)}(x_j) = \delta_{1,i}$$

and so multiplying by $\zeta^{(F)}(i)$, and summing over i, gives

$$\underbrace{\sum_i \sum_j k(x_i,x_j)\,\zeta^{(F)}(x_i)\,\zeta^{(F)}(x_j)}_{=\|P_F(\delta_{x_1})\|^2_{\mathscr{H}}} = \zeta^{(F)}(x_1).$$

■

Corollary 15 *We have*

(i)

$$P_F\left(\delta_{x_1}\right) = \zeta^{(F)}\left(x_1\right) k_{x_1} + \sum_{y \in F \setminus \{x_1\}} \zeta^{(F)}\left(y\right) k_y \tag{30}$$

where ζ_F solves (26), for all $F \in \mathscr{F}(V)$;

(ii)

$$\left\| P_F\left(\delta_{x_1}\right)\right\|_{\mathscr{H}}^2 = \zeta^{(F)}\left(x_1\right) \tag{31}$$

and so in particular:

(iii)

$$0 < \zeta^{(F)}\left(x_1\right) \leq \left\|\delta_{x_1}\right\|_{\mathscr{H}}^2 \tag{32}$$

Proof Formula (31) follows from the definition of $\zeta^{(F)}$ as a solution to the matrix problem $K_N \zeta^{(F)} = \delta_{x_1}$, but we may also prove (31) directly from

$$P_F\left(\delta_{x_1}\right) = \sum_y \zeta^{(F)}\left(y\right) k_y \,. \tag{33}$$

Apply $\langle \cdot, \delta_{x_1} \rangle_{\mathscr{H}}$ to both sides in (33), we get

$$\underbrace{\left\langle \delta_{x_1}, P_F\left(\delta_{x_1}\right)\right\rangle_{\mathscr{H}}}_{\left\|P_F\left(\delta_{x_1}\right)\right\|_{\mathscr{H}}^2} = \zeta^{(F)}\left(x_1\right)$$

since $P_F = P_F^* = P_F^2$; i.e., a projection in the RKHS $\mathscr{H} = \mathscr{H}_V$ of k. ∎

Example 1 ($\#F = 2$) *Let $F = \{x_1, x_2\}$, $K_F = \left(k_{ij}\right)_{i,j=1}^2$, where $k_{ij} := k\left(x_i, x_j\right)$. Then (26) reads*

$$\begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} \zeta_F\left(x_1\right) \\ \zeta_F\left(x_2\right) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{34}$$

Set $D := \det\left(K_F\right) = k_{11} k_{22} - k_{12} k_{21}$, then:

$$\zeta_F\left(x_1\right) = \frac{k_{22}}{D}, \quad \zeta_F\left(x_2\right) = -\frac{k_{21}}{D}.$$

Example 2 *Let $V = \{x_1, x_2, \ldots\}$ be an ordered set. Set $F_n := \{x_1, \ldots, x_n\}$. Note that with*

$$D_n = \det\left(K_{F_n}\right) = \det\left(\left(k\left(x_i, x_j\right)\right)_{i,j=1}^n\right), \text{ and} \tag{35}$$

$$D'_{n-1} = (1,1) \text{ minor of } K_{F_n} = \det\left(\left(k\left(x_i, x_j\right)\right)_{i,j=2}^n\right); \tag{36}$$

then

$$\zeta^{(F_n)}\left(x_1\right) = \frac{D'_{n-1}}{D_n} = \left(K_{F_n}^{-1} \delta_{x_1}\right)\left(x_1\right). \tag{37}$$

Corollary 16 *We have*

$$\frac{1}{k(x_1,x_1)} \leq \frac{k(x_2,x_2)}{D_2} \leq \cdots \leq \frac{D'_{n-1}}{D_n} \leq \cdots \leq \|\delta_{x_1}\|^2_{\mathscr{H}}.$$

Proof Follows from (37), and if $F \subset F'$ are two finite subsets, then

$$\|P_F(\delta_{x_1})\|^2_{\mathscr{H}} \leq \|P_{F'}(\delta_{x_1})\|^2_{\mathscr{H}} \leq \|\delta_{x_1}\|^2_{\mathscr{H}}.$$

■

Let $k : V \times V \to \mathbb{R}$ be as specified above. Let $\mathscr{H} = \mathscr{H}(k)$ be the RKHS. We set $\mathscr{F}(V)$:= all finite subsets of V; and if $x \in V$ is fixed, $\mathscr{F}_x(V) := \{F \in \mathscr{F}(V) \mid x \in F\}$.

For $F \in \mathscr{F}(V)$, let K_F be the $\#F \times \#F$ matrix given by $(k(x,y))_{(x,y)\in F\times F}$. Following Karlin and Ziegler (1996), we say that k is *strictly positive* iff $\det K_F > 0$ for all $F \in \mathscr{F}(V)$.

Set $D_F := \det K_F$. If $x \in V$, and $F \in \mathscr{F}_x(V)$, set $K'_F :=$ the minor in K_F obtained by omitting row x and column x, see Figure 2.

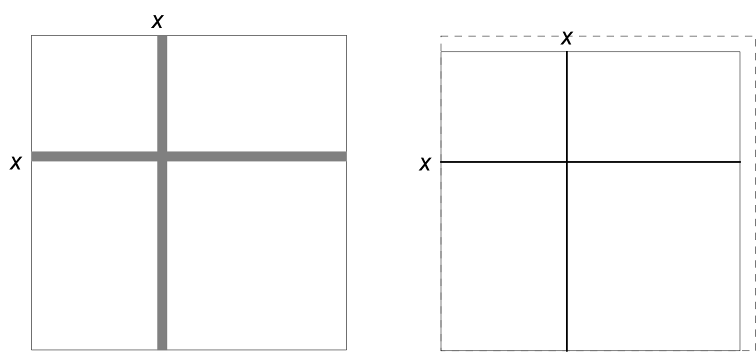


Figure 2: The (x,x) minors, $K_F \to K'_F$.

Corollary 17 *Suppose $k : V \times V \to \mathbb{R}$ is strictly positive. Let $x \in V$. Then*

$$\delta_x \in \mathscr{H} \Longleftrightarrow \sup_{F\in\mathscr{F}_x(V)} \frac{D'_F}{D_F} < \infty. \tag{38}$$

2.1 Unbounded Containment in RKHSs

Definition 18 *Let $\mathscr{K}$ and $\mathscr{H}$ be two Hilbert spaces. We say that $\mathscr{K}$ is* unboundedly contained *in $\mathscr{H}$ if there is a dense subspace $\mathscr{K}_0 \subset \mathscr{K}$ such that $\mathscr{K}_0 \subset \mathscr{H}$; and the inclusion operator, with $\mathscr{K}_0$ as its dense domain, is* closed, *i.e.,*

$$\mathscr{K} \overset{incl}{\hookrightarrow} \mathscr{H}, \quad dom(incl) = \mathscr{K}_0.$$

Let $k : V \times V \to \mathbb{R}$ be a p.d. kernel, and let $\mathscr{H}$ be the corresponding RKHS. Set $\mathscr{K} = l^2(V)$, and

$$\mathscr{K}_0 = span\{\delta_x \mid x \in V\}. \tag{39}$$

Proposition 19 *If $\delta_x \in \mathscr{H}$ for $\forall x \in V$, then $l^2(V)$ is unboundedly contained in $\mathscr{H}$.*

Proof Recall that $\mathscr{H}$ is the RKHS defined for a fixed p.d. kernel $k : V \times V \to \mathbb{R}$. Let k_x be the vector in $\mathscr{H}$, given by $k_x(y) = k(x, y)$, s.t.

$$f(x) = \langle k_x, f \rangle_{\mathscr{H}}, \quad \forall f \in \mathscr{H}. \tag{40}$$

To finish the proof we will need: ■

Lemma 20 *The following equation*

$$\langle \delta_x, k_y \rangle_{\mathscr{H}} = \delta_{x,y} \tag{41}$$

holds if $\delta_x \in \mathscr{H}$ for $\forall x \in V$.

Proof (41) is immediate from (40). ■

Lemma 21 *On*

$$span\{k_x \mid x \in V\} \subset \mathscr{H} \tag{42}$$

define $Mk_x := \delta_x$, then by Lemma 20, M extends to be a well defined operator $M : \mathscr{H} \to l^2(V)$ with dense domain (42). We have

$$\langle k, Mf \rangle_{l^2(V)} = \langle k, f \rangle_{\mathscr{H}}, \quad \forall k \in span\{\delta_x\}, \forall f \in dom(M). \tag{43}$$

Proof By linearity, it is enough to prove that

$$\langle \delta_x, \delta_y \rangle_{l^2} = \langle \delta_x, k_y \rangle_{\mathscr{H}} \tag{44}$$

holds for $\forall x, y \in V$. But (44) follows immediate from Lemma 20. ■

Corollary 22 *If $L : l^2(V) \to \mathscr{H}$ denotes the inclusion mapping with*

$$dom(L) = span\{\delta_x : x \in V\},$$

then we conclude that

$$L \subset M^*, \text{ and } M \subset L^*. \tag{45}$$

Since $dom(M)$ is dense in $\mathscr{H}$, it follows that L^* has dense domain; and that therefore L is closable.

Remark 23 *This also completes the proof of Proposition 19.*

Corollary 24 *Suppose $k : V \times V \to \mathbb{R}$ is as given, and that $\mathscr{H} = RKHS(k)$. Let L be the densely defined inclusion mapping $l^2(V) \to \mathscr{H}$. Then L^*L is selfadjoint with dense domain in $l^2(V)$; and LL^* is selfadjoint with dense domain in $\mathscr{H}$. Moreover, the following polar decomposition holds:*

$$L = U(L^*L)^{1/2} = (LL^*)^{1/2} U \tag{46}$$

where U is a partial isometry $l^2(V) \to \mathscr{H}$.

3. Point-masses in Concrete Models

Suppose $V \subset D \subset \mathbb{R}^d$ where V is countable and discrete, but D is open. In this case, we get two kernels: k on $D \times D$, and $k_V := k\big|_{V \times V}$ on $V \times V$ by restriction. If $x \in V$, then $k_x^{(V)}(\cdot) = k(\cdot, x)$ is a function on V, while $k_x(\cdot) = k(\cdot, x)$ is a function on D.

This means that the corresponding RKHSs are different, $\mathscr{H}_V$ vs $\mathscr{H}$, where $\mathscr{H}_V$ = a RKHS of functions on V, and $\mathscr{H}$ = a RKHS of functions on D.

Lemma 25 *$\mathscr{H}_V$ is isometrically contained in $\mathscr{H}$ via $k_x^{(V)} \longmapsto k_x$, $x \in V$.*

Proof If $F \subset V$ is a finite subset, and $\xi = \xi_F$ is a function on F, then

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

The desired result follows from this. ■

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 (restriction) $\delta_x\big|_V \in \mathscr{H}_V$, but $\delta_x \notin \mathscr{H}$; indeed this happens for the kernel k of standard Brownian motion:

$D = \mathbb{R}_+$;

V = an ordered subset $0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots$, $V = \{x_i\}_{i=1}^{\infty}$.

In this case, we compute $\mathscr{H}_V$, and we show that $\delta_{x_i}\big|_V \in \mathscr{H}_V$; while for $\mathscr{H}_m$ = the Cameron-Martin Hilbert space, we have $\delta_{x_i} \notin \mathscr{H}_m$.

Also note that δ_{x_1} has a different meaning with reference to $\mathscr{H}_V$ vs $\mathscr{H}_m$. In the first case, it is simply $\delta_{x_1}(y) = \begin{cases} 1 & y = x_1 \\ 0 & y \in V \setminus \{x_1\} \end{cases}$. In the second case, δ_{x_1} is a Schwartz distribution. We shall abuse notation, writing δ_x in both cases.

In the following, we will consider restriction to $V \times V$ of a special continuous p.d. kernel k on $\mathbb{R}_+ \times \mathbb{R}_+$. It is $k(s,t) = s \wedge t = \min(s,t)$. Before we restrict, note that the RKHS of this k is the Cameron-Martin Hilbert space of function f on $\mathbb{R}_+$ with distribution derivative $f' \in L^2(\mathbb{R}_+)$, and

$$\|f\|_{\mathscr{H}}^2 := \int_0^\infty \left| f'(t) \right|^2 dt < \infty. \tag{47}$$

For details, see below.

Remark 26 (Application) *The Hilbert space given by $\|\cdot\|_{\mathscr{H}}^2$ in (47) is called the Cameron-Martin Hilbert space, and, as noted, it is the RKHS of $k : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R} : k(s,t) := s \wedge t$. Now pick a discrete subset $V \subset \mathbb{R}_+$; then Lemma 25 states that the RKHS of the $V \times V$ restricted kernel, $k^{(V)}$ is isometrically embedded into $\mathscr{H}$, i.e., setting*

$$J^{(V)}\left(k_x^{(V)}\right) = k_x, \quad \forall x \in V; \tag{48}$$

$J^{(V)}$ extends by "closed span" to an isometry $\mathscr{H}_V \xrightarrow{J^{(V)}} \mathscr{H}$. It further follows from the lemma, that the range of $J^{(V)}$ may have infinite co-dimension.

Note that $P_V := J^{(V)} \left(J^{(V)}\right)^*$ is the projection onto the range of $J^{(V)}$. The ortho-complement is as follow:

$$\mathscr{H} \ominus \mathscr{H}_V = \left\{ \psi \in \mathscr{H} \;\middle|\; \psi(x) = 0,\ \forall x \in V \right\}. \tag{49}$$

Example 3 *Let k and $k^{(V)}$ be as in (48), and set $V := \pi\mathbb{Z}_+$, i.e., integer multiples of π. Then easy generators of wavelet functions (Bratteli and Jorgensen, 2002) yield non-zero functions ψ on $\mathbb{R}_+$ such that*

$$\psi \in \mathscr{H} \ominus \mathscr{H}_V. \tag{50}$$

More precisely,

$$0 < \int_0^\infty \left|\psi'(t)\right|^2 dt < \infty, \tag{51}$$

where ψ' is the distribution (weak) derivative; and

$$\psi(n\pi) = 0, \quad \forall n \in \mathbb{Z}_+. \tag{52}$$

An explicit solution to (50)-(52) is

$$\psi(t) = \prod_{n=1}^{\infty} \cos\left(\frac{t}{2^n}\right) = \frac{\sin t}{t}, \quad \forall t \in \mathbb{R}. \tag{53}$$

From this, one easily generates an infinite-dimensional set of solutions.

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

$$\mathbb{E}(B_s B_t) = s \wedge t = \min(s,t). \tag{54}$$

We now show that the restriction of (54) to $V \times V$ for an ordered subset (we fix such a set V):

$$V:\ 0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots \tag{55}$$

has the discrete mass property (Definition 4).

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

$$k_V(x_i, x_j) = x_i \wedge x_j. \tag{56}$$

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

$$K_n = k^{(F_n)} = \begin{bmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & x_2 & \cdots & x_2 \\ x_1 & x_2 & x_3 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} = (x_i \wedge x_j)_{i,j=1}^n. \tag{57}$$

We will show that condition (iii) in Theorem 1 holds for k_V. For this, we must compute all the determinants, $D_n = \det(K_F)$ etc. ($n = \#F$), see Corollary 17.

Lemma 27

$$D_n = \det\left((x_i \wedge x_j)_{i,j=1}^n\right) = x_1(x_2 - x_1)(x_3 - x_2)\cdots(x_n - x_{n-1}). \tag{58}$$

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. ■

Lemma 28 *Let*

$$\zeta_{(n)} := K_n^{-1} \left(\delta_{x_1}\right)(\cdot) \tag{59}$$

be as in (11), so that

$$\left\| P_{F_n}\left(\delta_{x_1}\right)\right\|_{\mathscr{H}_V}^2 = \zeta_{(n)}\left(x_1\right). \tag{60}$$

Then,

$$\begin{aligned} \zeta_{(1)}\left(x_1\right) &= \frac{1}{x_1} \\ \zeta_{(n)}\left(x_1\right) &= \frac{x_2}{x_1\left(x_2 - x_1\right)}, \quad \textit{for } n = 2, 3, \ldots, \end{aligned}$$

and

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

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

$$D'_{n-1} = \det\left(\left(x_i \wedge x_j\right)_{i,j=2}^n\right) = x_2\left(x_3 - x_2\right)\left(x_4 - x_3\right)\cdots\left(x_n - x_{n-1}\right) \tag{61}$$

and so

$$\begin{aligned} \zeta_{(1)}\left(x_1\right) &= \frac{1}{x_1}, \quad \text{and} \\ \zeta_{(2)}\left(x_1\right) &= \frac{x_2}{x_1\left(x_2 - x_1\right)} \\ \zeta_{(3)}\left(x_1\right) &= \frac{x_2\left(x_3 - x_2\right)}{x_1\left(x_2 - x_1\right)\left(x_3 - x_2\right)} = \frac{x_2}{x_1\left(x_2 - x_1\right)} \\ \zeta_{(4)}\left(x_1\right) &= \frac{x_2\left(x_3 - x_2\right)\left(x_4 - x_3\right)}{x_1\left(x_2 - x_1\right)\left(x_3 - x_2\right)\left(x_4 - x_3\right)} = \frac{x_2}{x_1\left(x_2 - x_1\right)} \\ &\vdots \end{aligned}$$

The result follows from this, and from Corollary 9. ■

Corollary 29 $P_{F_n}(\delta_{x_1}) = P_{F_2}(\delta_{x_1})$, $\forall n \geq 2$. *Therefore,*

$$\delta_{x_1} \in \mathscr{H}_V^{(F_2)} := span\{k_{x_1}^{(V)}, k_{x_2}^{(V)}\} \tag{62}$$

and

$$\delta_{x_1} = \zeta_{(2)}(x_1)\, k_{x_1}^{(V)} + \zeta_{(2)}(x_2)\, k_{x_2}^{(V)} \tag{63}$$

where

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

Specifically,

$$\zeta_{(2)}(x_1) = \frac{x_2}{x_1(x_2 - x_1)} \tag{64}$$

$$\zeta_{(2)}(x_2) = \frac{-1}{x_2 - x_1}; \tag{65}$$

and

$$\|\delta_{x_1}\|_{\mathscr{H}_V}^2 = \frac{x_2}{x_1(x_2 - x_1)}. \tag{66}$$

Proof Follows from the lemma. Note that

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

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 (66). ■

Remark 30 *We showed that* $\delta_{x_1} \in \mathscr{H}_V$, $V = \{x_1 < x_2 < \cdots\} \subset \mathbb{R}_+$, *with the restriction of* $s \wedge t$ *= the covariance kernel of Brownian motion.*

The same argument also shows that $\delta_{x_i} \in \mathscr{H}_V$ *when* $i > 1$. *We only need to modify the index notation from the case of the proof for* $\delta_{x_1} \in \mathscr{H}_V$. *The details are sketched below.*

Fix $V = \{x_i\}_{i=1}^{\infty}$, $x_1 < x_2 < \cdots$, *then*

$$P_{F_n}(\delta_{x_i}) = \begin{cases} 0 & \text{if } n < i - 1 \\ \sum_{s=1}^{n} \left(K_{F_n}^{-1}\delta_{x_i}\right)(x_s)\, k_{x_s} & \text{if } n \geq i \end{cases}$$

and

$$\|P_{F_n}(\delta_{x_i})\|_{\mathscr{H}}^2 = \begin{cases} 0 & \text{if } n < i - 1 \\ \frac{1}{x_i - x_{i-1}} & \text{if } n = i \\ \frac{x_{i+1} - x_{i-1}}{(x_i - x_{i-1})(x_{i+1} - x_i)} & \text{if } n > i \end{cases}$$

Conclusion.

$$\delta_{x_i} \in span\left\{k_{x_{i-1}}^{(V)}, k_{x_i}^{(V)}, k_{x_{i+1}}^{(V)}\right\}, \quad \text{and} \tag{67}$$

$$\|\delta_{x_i}\|_{\mathscr{H}}^2 = \frac{x_{i+1} - x_{i-1}}{(x_i - x_{i-1})(x_{i+1} - x_i)}. \tag{68}$$

Corollary 31 *Let $V \subset \mathbb{R}_+$ be countable. If $x_a \in V$ is an accumulation point (from V), then $\|\delta_a\|_{\mathscr{H}_V} = \infty$.*

Remark 32 *This computation will be revisited in Section 4, in a much wider context.*

Example 4 *An illustration for $0 < x_1 < x_2 < x_3 < x_4$:*

$$\begin{aligned} P_F(\delta_{x_3}) &= \sum_{y \in F} \zeta^{(F)}(y)\, k_y(\cdot) \\ \zeta^{(F)} &= K_F^{-1} \delta_{x_3}\,. \end{aligned}$$

That is,

$$\underbrace{\begin{bmatrix} x_1 & x_1 & x_1 & x_1 \\ x_1 & x_2 & x_2 & x_2 \\ x_1 & x_2 & x_3 & x_3 \\ x_1 & x_2 & x_3 & x_4 \end{bmatrix}}_{(K_F(x_i,x_j))_{i,j=1}^4} \begin{bmatrix} \zeta^{(F)}(x_1) \\ \zeta^{(F)}(x_2) \\ \zeta^{(F)}(x_3) \\ \zeta^{(F)}(x_4) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

and

$$\begin{aligned} \zeta^{(F)}(x_3) &= \frac{x_1(x_2 - x_1)(x_4 - x_2)}{x_1(x_2 - x_1)(x_3 - x_2)(x_4 - x_3)} \\ &= \frac{x_4 - x_2}{(x_3 - x_2)(x_4 - x_3)} = \|\delta_{x_3}\|_{\mathscr{H}}^2\,. \end{aligned}$$

Example 5 (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 33 *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:*

$$\delta_{x_1} = \sum_{y \in V} \zeta^{(x_1)}(y)\, k_y\,, \tag{69}$$

then

$$\|\delta_{x_1}\|_{\mathscr{H}}^2 = \zeta^{(x_1)}(x_1)\,. \tag{70}$$

Proof Substitute both sides of (69) into $\langle\delta_{x_1},\cdot\rangle_{\mathscr{H}}$ where $\langle\cdot,\cdot\rangle_{\mathscr{H}}$ denotes the inner product in $\mathscr{H}$. ■

Example 6 (Application) *Suppose $V=\cup_n F_n$, $F_n\subset F_{n+1}$, where each $F_n\in\mathscr{F}(V)$, then if $x_1\in F_n$, we have*

$$P_{F_n}(\delta_{x_1})=\sum_{y\in F_n}\left\langle x_1,K_{F_n}^{-1}y\right\rangle_{l^2}k_y \tag{71}$$

and

$$\|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2=\left\langle x_1,K_{F_n}^{-1}x_1\right\rangle_{l^2}=\left(K_{F_n}^{-1}\delta_{x_1}\right)(x_1) \tag{72}$$

and the expression $\|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2$ is monotone in n, i.e.,

$$\|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2\leq\left\|P_{F_{n+1}}(\delta_{x_1})\right\|_{\mathscr{H}}^2\leq\cdots\leq\|\delta_{x_1}\|_{\mathscr{H}}^2$$

with

$$\sup_{n\in\mathbb{N}}\|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2=\lim_{n\to\infty}\|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2=\|\delta_{x_1}\|_{\mathscr{H}}^2.$$

Question 34 *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, e.g., $V=\mathbb{Z}^d$. When does $k\big|_{V\times V}$ have the discrete mass property?*

Examples of the affirmative, or not, will be discussed below.

3.2 Discrete RKHS from Restrictions

Let $D:=[0,\infty)$, and $k:D\times D\to\mathbb{R}$, with

$$k(x,y)=x\wedge y=\min(x,y).$$

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

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

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

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

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

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

Lemma 35 *We have $\delta_n=2k_n-k_{n+1}-k_{n-1}$.*

Proof Introduce the discrete Laplacian Δ, where

$$(\Delta f)(n) = 2f(n) - f(n-1) - f(n+1),$$

then $\Delta k_n = \delta_n$, and

$$\langle 2k_n - k_{n+1} - k_{n-1}, k_m \rangle_{\mathscr{H}_V} = \langle \delta_n, k_m \rangle_{\mathscr{H}_V} = \delta_{n,m}.$$

■

Remark 36 *The same argument as in the proof of the lemma shows (*mutatis mutandis*) that any ordered discrete countable infinite subset $V \subset [0, \infty)$ yields*

$$\mathscr{H}_V := \mathscr{H}\left(k\big|_{V \times V}\right)$$

as a RKHS which is discrete in that (Definition 4) if $V = \{x_i\}_{i=1}^{\infty}$, $x_i \in \mathbb{R}_+$, then $\delta_{x_i} \in \mathscr{H}_V$, $\forall i \in \mathbb{N}$.

Proof Fix vertices $V = \{x_i\}_{i=1}^{\infty}$,

$$0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \infty, \quad x_i \to \infty. \tag{73}$$

Assign conductance

$$c_{i,i+1} = c_{i+1,i} = \frac{1}{x_{i+1} - x_i} \left(= \frac{1}{\text{dist}} \right) \tag{74}$$

Let

$$\begin{aligned} (\Delta f)(x_i) &= \left(\frac{1}{x_{i+1} - x_i} + \frac{1}{x_i - x_{i-1}} \right) f(x_i) \\ &\quad - \frac{1}{x_i - x_{i-1}} f(x_{i-1}) - \frac{1}{x_{i+1} - x_i} f(x_{i+1}) \end{aligned} \tag{75}$$

Equivalently,

$$(\Delta f)(x_i) = (c_{i,i+1} + c_{i,i-1}) f(x_i) - c_{i,i-1} f(x_{i-1}) - c_{i,i+1} f(x_{i+1}). \tag{76}$$

Remark 37 *The most general graph-Laplacians will be discussed in detail in Section 4 below.*

Then, with (76) we have:

$$\Delta k_{x_i} = \delta_{x_i}$$

where $k(\cdot, \cdot)$ = restriction of $s \wedge t$ from $[0, \infty) \times [0, \infty)$ to $V \times V$; and therefore

$$\delta_{x_i} = (c_{i,i+1} + c_{i,i-1}) k_{x_i} - c_{i,i+1} k_{x_{i+1}} - c_{i,i-1} k_{x_{i-1}} \in \mathscr{H}_V \tag{77}$$

as the right-side in the last equation is a finite sum. Note that now the RKHS is

$$\mathscr{H}_V = \left\{ f : V \to \mathbb{C} \;\Big|\; \sum_{i=1}^{\infty} c_{i,i+1} \left| f(x_{i+1}) - f(x_i) \right|^2 < \infty \right\}.$$

■

3.3 Brownian Bridge

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

$$k_{bridge}(s,t) := s \wedge t - st; \tag{78}$$

then (78) is the covariance function for the Brownian bridge $B_{bri}(t)$, i.e.,

$$B_{bri}(0) = B_{bri}(1) = 0 \tag{79}$$

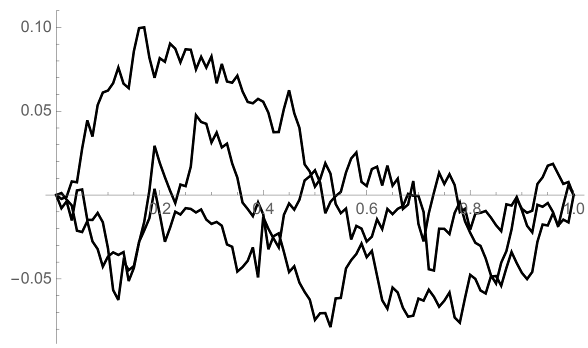


Figure 3: Brownian bridge $B_{bri}(t)$, a simulation of three sample paths of the Brownian bridge.

$$B_{bri}(t) = (1-t)\,B\left(\frac{t}{1-t}\right), \quad 0 < t < 1; \tag{80}$$

where $B(t)$ is Brownian motion; see Lemma 25.

The corresponding Cameron-Martin space is now

$$\mathscr{H}_{bri} = \left\{ f \text{ on } [0,1]\,;\, f' \in L^2(0,1)\,,\, f(0) = f(1) = 0 \right\} \tag{81}$$

with

$$\|f\|^2_{\mathscr{H}_{bri}} := \int_0^1 \left| f'(s) \right|^2 ds < \infty. \tag{82}$$

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

$$K_{F_n} = \left(k_{bridge}(x_i, x_j)\right)_{i,j=1}^{n}, \tag{83}$$

see (78), and

$$\det K_{F_n} = x_1 (x_2 - x_1) \cdots (x_n - x_{n-1})(1 - x_n). \tag{84}$$

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

3.4 Binomial RKHS

Definition 38 *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*

$$e_n(x) = \begin{cases} \binom{x}{n} & \text{if } n \leq x \\ 0 & \text{if } n > x. \end{cases} \tag{85}$$

Lemma 39 (Alpay and Jorgensen, 2015)

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

$$P_{F_n} = \sum_{k=0}^{n} |e_k\rangle\langle e_k| \tag{86}$$

or equivalently

$$P_{F_n} f = \sum_{k=0}^{n} \langle e_k, f\rangle_{\mathscr{H}} e_k\,. \tag{87}$$

then,

(iv) Formula (87) is well defined for all functions $f : V \to \mathbb{C}$, $f \in \mathscr{F}unc(V)$.

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

$$f \in \mathscr{H} \Longleftrightarrow \sum_{k=0}^{\infty} \left|\langle e_k, f\rangle_{\mathscr{H}}\right|^2 < \infty; \tag{88}$$

and, in this case,

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

Fix $x_1 \in V$, then we shall apply Lemma 39 to the function $f_1 = \delta_{x_1}$ (in $\mathscr{F}unc(V)$), $f_1(y) = \begin{cases} 1 & \text{if } y = x_1 \\ 0 & \text{if } y \neq x_1. \end{cases}$

Theorem 2 *We have*

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

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

Lemma 40 (Alpay and Jorgensen, 2015)

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

$$\delta_{m,n}=\sum_{j=m}^{n}(-1)^{m+j}\binom{n}{j}\binom{j}{m}. \tag{89}$$

(ii) For all $n\in\mathbb{Z}_+$, the inverse of the following lower triangle matrix is this: With (see Figure 4)

$$L_{xy}^{(n)}=\begin{cases}\binom{x}{y} & \text{if } y\le x\le n\\ 0 & \text{if } x<y\end{cases} \tag{90}$$

we have:

$$\left(L^{(n)}\right)_{xy}^{-1}=\begin{cases}(-1)^{x-y}\binom{x}{y} & \text{if } y\le x\le n\\ 0 & \text{if } x<y.\end{cases} \tag{91}$$

Notation: The numbers in (91) are the entries of the matrix $\left(L^{(n)}\right)^{-1}$.

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

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

Figure 4: The matrix L_n is simply a truncated Pascal triangle, arranged to fit into a lower triangular matrix.

Corollary 41 *Let k_b, $\mathscr{H}$, and $n \in \mathbb{Z}_+$ be as above with the lower triangle matrix L_n. Set*

$$K_n(x, y) = k_b(x, y), \quad (x, y) \in F_n \times F_n, \tag{92}$$

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

(i) Then K_n is invertible with

$$K_n^{-1} = \left(L_n^{tr}\right)^{-1} (L_n)^{-1}; \tag{93}$$

an (upper triangle) $\times$ (lower triangle) factorization.

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

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

Conclusion: Since

$$\|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2 = \left\langle x_1, K_n^{-1} x_1 \right\rangle_{\mathscr{H}} \tag{94}$$

for all $x_1 \in F_n$, we get

$$\begin{aligned} \|P_{F_n}(\delta_{x_1})\|_{\mathscr{H}}^2 &= \sum_{k=x_1}^{n} \binom{k}{x_1}^2 \\ &= 1 + \binom{x_1+1}{x_1}^2 + \binom{x_1+2}{x_1}^2 + \cdots + \binom{n}{x_1}^2; \end{aligned} \tag{95}$$

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 $\mathscr{H}$.

4. Infinite Network of Resistors

Here we introduce a family of positive definite kernels $k : V \times V \to \mathbb{R}$, defined on infinite sets V of vertices for a given graph $G = (V, E)$ with edges $E \subset V \times V \backslash$(diagonal).

There is a large literature dealing with analysis on infinite graphs (Jorgensen and Pearse, 2010, 2011, 2013; Okoudjou and Strichartz, 2005; Boyle et al., 2007; Cho and Jorgensen, 2011).

Our main purpose here is to point out that every assignment of resistors on the edges E in G yields a p.d. kernel k, and an associated RKHS $\mathscr{H} = \mathscr{H}(k)$ such that

$$\delta_x \in \mathscr{H}, \quad \text{for all } x \in V. \tag{96}$$

Definition 42 *Let $G = (V, E)$ be as above. Assume*

1. *$(x, y) \in E \Longleftrightarrow (y, x) \in E$;*

2. $\exists c : E \to \mathbb{R}_+$ *(a conductance function = 1 / resistance) such that*

(i) $c_{(xy)} = c_{(yx)}$, $\forall (xy) \in E$;

(ii) *for all* $x \in V$, $\#\left\{y \in V \mid c_{(xy)} > 0\right\} < \infty$; *and*

(iii) $\exists o \in V$ *s.t. for* $\forall x \in V \setminus \{o\}$, $\exists$ *edges* $(x_i, x_{i+1})_0^{n-1} \in E$ *s.t.* $x_o = 0$, *and* $x_n = x$; *called connectedness.*

Given $G = (V, E)$, and a fixed conductance function $c : E \to \mathbb{R}_+$ as specified above, we now define a corresponding Laplace operator $\Delta = \Delta^{(c)}$ acting on functions on V, i.e., on $\mathscr{F}unc(V)$ by

$$(\Delta f)(x) = \sum_{y \sim x} c_{xy} \left(f(x) - f(y)\right). \tag{97}$$

Let $\mathscr{H}$ be the Hilbert space defined as follows: A function f on V is in $\mathscr{H}$ iff $f(o) = 0$, and

$$\|f\|_{\mathscr{H}}^2 := \frac{1}{2} \sum\sum_{\substack{(x,y) \in E \\ \subset V \times V}} c_{xy} \left|f(x) - f(y)\right|^2 < \infty. \tag{98}$$

Lemma 43 (Jorgensen and Pearse, 2010) *For all* $x \in V \setminus \{o\}$, $\exists v_x \in \mathscr{H}$ *s.t.*

$$f(x) - f(o) = \langle v_x, f \rangle_{\mathscr{H}}, \quad \forall f \in \mathscr{H} \tag{99}$$

where

$$\langle h, f \rangle_{\mathscr{H}} = \frac{1}{2} \sum\sum_{(x,y) \in E} c_{xy} \left(\overline{h(x)} - \overline{h(y)}\right) \left(f(x) - f(y)\right), \quad \forall h, f \in \mathscr{H}. \tag{100}$$

(The system $\{v_x\}$ *is called a system of dipoles.)*

Proof Let $x \in V \setminus \{o\}$, and use (97) together with the Schwarz-inequality to show that

$$\left|f(x) - f(o)\right|^2 \leq \sum_i \frac{1}{c_{x_i x_{i+1}}} \sum_i c_{x_i x_{i+1}} \left|f(x_i) - f(x_{i+1})\right|^2.$$

An application of Riesz' lemma then yields the desired conclusion.

Note that $v_x = v_x^{(c)}$ depends on the choice of base point $o \in V$, and on conductance function c; see (i)-(ii) and (98). ■

Now set

$$k^{(c)}(x, y) = \langle v_x, v_y \rangle_{\mathscr{H}}, \quad \forall (xy) \in (V \setminus \{o\}) \times (V \setminus \{o\}). \tag{101}$$

It follows from a theorem that $k^{(c)}$ is a Green's function for the Laplacian $\Delta^{(c)}$ in the sense that

$$\Delta^{(c)} k^{(c)}(x, \cdot) = \delta_x \tag{102}$$

where the dot in (102) is the dummy-variable in the action. Note that the solution to (102) is not unique.

Lemma 44 (Jorgensen and Pearse, 2011) *Let* $G = (V, E)$*, and conductance function* $c : E \to \mathbb{R}_+$ *be a s specified above; then* $k^{(c)}$ *in (101) is positive definite, and the corresponding RKHS* $\mathscr{H}\left(k^{(c)}\right)$ *is the Hilbert space introduced in (98) and (100), called the energy-Hilbert space.*

Proof See Jorgensen et al. (2010; 2011; 2013). ■

Proposition 45 *Let* $x \in V \backslash \{o\}$*, and let* $c : E \to \mathbb{R}_+$ *be specified as above. Let* $\mathscr{H} = \mathscr{H}(k^c)$ *be the corresponding RKHS. Then* $\delta_x \in \mathscr{H}$*, and*

$$\|\delta_x\|^2_{\mathscr{H}} = \sum_{y \sim x} c_{(xy)} =: c(x). \tag{103}$$

Proof We study the finite matrices, defined for $\forall F \in \mathscr{F}(V)$, by

$$K_F(x, y) = k^c(x, y), \quad (x, y) \in F \times F. \tag{104}$$

Fix $x \in V \backslash \{o\}$, and pick $F \in \mathscr{F}(V)$ such that

$$\{x\} \cup \{y \in V \mid y \sim x\} \subset F, \tag{105}$$

see Figure 5; an interior point:

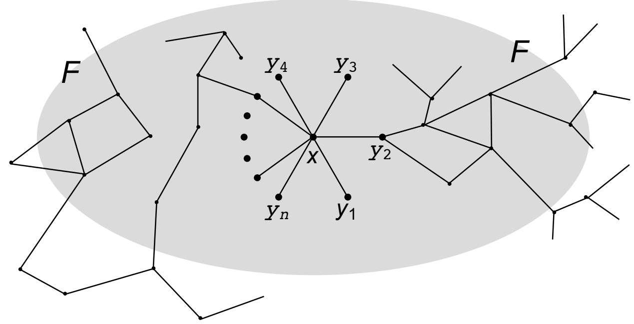


Figure 5: Neighborhood of x, see Definition 42 (ii). An interior point x.

Let $F \in \mathscr{F}(V)$ be as in (104) and in Figure 5, and let $\Delta = \Delta^{(c)}$ be the Laplace operator (97), then for all $(x, y) \in F \times F$, we have:

$$\begin{aligned} \left\langle x, K_F^{-1} y \right\rangle_{l^2} &= \langle \delta_x, \Delta \delta_y \rangle_{l^2} \\ &= (\Delta \delta_y)(x) \\ &= \begin{cases} c(x) & \text{if } y = x; \text{ see } (103) \\ -c_{(xy)} & \text{if } y \sim x \\ 0 & \text{for all other values of } y \end{cases} \end{aligned} \tag{106}$$

In particular,

$$\sup_{F\in\mathscr{F}(V)}\left(K_F\delta_x\right)(x)<\infty;$$

and in fact,

$$\|\delta_x\|^2_{\mathscr{H}}=c\left(x\right),\text{ for all }x\in V\backslash\left\{o\right\},$$

as claimed in the Proposition.

The last step in the present proof uses the equivalence (i)⇔(ii)⇔(iii) from Theorem 1 above.

Finally, we note that the assertion in (106) follows from

$$\Delta v_x=\delta_x-\delta_o,\quad \forall x\in V\backslash\left\{o\right\}. \tag{107}$$

And (107) in turn follows from (99), (97) and a straightforward computation. ■

Corollary 46 *Let $G=(V,E)$ and conductance $c:E\rightarrow\mathbb{R}_+$ be as specified above. Let $\Delta=\Delta^{(c)}$ be the corresponding Laplace operator. Let $\mathscr{H}=\mathscr{H}\left(k^c\right)$ be the RKHS. Then*

$$\left\langle\delta_x,f\right\rangle_{\mathscr{H}}=\left(\Delta f\right)(x) \tag{108}$$

and

$$\delta_x=c\left(x\right)v_x-\sum_{y\sim x}c_{xy}v_y \tag{109}$$

holds for all $x\in V$.

Proof Since the system $\{v_x\}$ of dipoles in (99) span a dense subspace in $\mathscr{H}$, it is enough to verify (108) when $f=v_y$ for $y\in V\backslash\left\{o\right\}$. But in this case, (108) follows from (102) and (106). ■

Corollary 47 *Let $G=(V,E)$, and conductance $c:E\rightarrow\mathbb{R}_+$ be as before; let $\Delta^{(c)}$ be the Laplace operator, and $\mathscr{H}_E^{(c)}$ the energy-Hilbert space in Definition 42 (Equation (98)). Let $k^{(c)}\left(x,y\right)=\left\langle v_x,v_y\right\rangle_{\mathscr{H}_E}$ be the kernel from (101), i.e., the Green's function of $\Delta^{(c)}$. Then the two Hilbert spaces $\mathscr{H}_E$, and $\mathscr{H}\left(k^{(c)}\right)=RKHS\left(k^{(c)}\right)$, are naturally isometrically isomorphic via $v_x\longmapsto k_x^{(c)}$ where $k_x^{(c)}=k^{(c)}\left(x,\cdot\right)$ for all $x\in V$.*

Proof Let $F\in\mathscr{F}\left(V\right)$, and let ξ be a function on F; then

$$\begin{aligned}\left\|\sum\nolimits_{x\in F}\xi\left(x\right)k_x^{(c)}\right\|^2_{\mathscr{H}\left(k^{(c)}\right)} &= \sum\sum_{F\times F}\overline{\xi\left(x\right)}\xi\left(y\right)k^{(c)}\left(x,y\right)\\ &\underset{(101)}{=} \sum\sum_{F\times F}\overline{\xi\left(x\right)}\xi\left(y\right)\left\langle v_x,v_y\right\rangle_{\mathscr{H}_E}\\ &= \left\|\sum\nolimits_{x\in F}\xi\left(x\right)v_x\right\|^2_{\mathscr{H}_E}.\end{aligned}$$

The remaining steps in the proof of the Corollary now follows from the standard completion from dense subspaces in the respective two Hilbert spaces $\mathscr{H}_E$ and $\mathscr{H}\left(k^{(c)}\right)$. ■

In the following we show how the kernels $k^{(c)} : V \times V \to \mathbb{R}$ from (101) in Lemma 43 are related to metrics on V; so called *resistance metrics* (Jorgensen and Pearse, 2010; Alpay et al., 2013).

Corollary 48 *Let $G = (V, E)$, and conductance $c : E \to \mathbb{R}_+$ be as above; and let $k^{(c)}(x, y) := \langle v_x, v_y \rangle_{\mathscr{H}_E}$ be the corresponding Green's function for the graph Laplacian $\Delta^{(c)}$.*

Then there is a metric $R\left(= R^{(c)} = \text{the resistance metric}\right)$, such that

$$k^{(c)}(x, y) = \frac{R^{(c)}(o, x) + R^{(c)}(o, y) - R^{(c)}(x, y)}{2} \tag{110}$$

holds on $V \times V$. Here the base-point $o \in V$ is chosen and fixed s.t.

$$\langle V_x, f \rangle_{\mathscr{H}_E} = f(x) - f(o), \quad \forall f \in \mathscr{H}_E,\ \forall x \in V. \tag{111}$$

Proof Set

$$R^{(c)}(x, y) = \|v_x - v_y\|^2_{\mathscr{H}_E}. \tag{112}$$

We proved (Jorgensen and Pearse, 2010) that $R^{(c)}(x, y)$ in (112) indeed defines a metric on V; the so called *resistance metric*. It represents the voltage-drop from x to y when 1 Amp is fed into (G, c) at the point x, and then extracted at y.

The verification of (110) is now an easy computation, as follows:

$$\begin{aligned} & \frac{R^{(c)}(o, x) + R^{(c)}(o, y) - R^{(c)}(x, y)}{2} \\ = & \frac{\|v_x\|^2_{\mathscr{H}_E} + \|v_y\|^2_{\mathscr{H}_E} - \|v_x - v_y\|^2_{\mathscr{H}_E}}{2} \\ = & \langle v_x, v_y \rangle_{\mathscr{H}_E} \\ = & k^{(c)}(x, y) \quad \text{by (101).} \end{aligned}$$

■

Proposition 49 *In the two cases: (i) $B(t)$, Brownian motion on $0 < t < \infty$; and (ii) the Brownian bridge $B_{bri}(t)$, $0 < t < 1$, from Section 3 (Figure 3), the corresponding resistance metric R is as follows:*

(i) If $V = \{x_i\}_{i=1}^{\infty} \subset (0, \infty)$, $x_1 < x_2 < \cdots$, then

$$R_B^{(V)}(x_i, x_j) = |x_i - x_j|. \tag{113}$$

(ii) If $W = \{x_i\}_{i=1}^{\infty} \subset (0, 1)$, $0 < x_1 < x_2 < \cdots < 1$, then

$$R_{bridge}^{(W)}(x_i, x_j) = |x_i - x_j| \cdot (1 - |x_i - x_j|). \tag{114}$$

In the completion w.r.t. the resistance metric $R_{bridge}^{(W)}$, the two endpoints $x = 0$ and $x = 1$ are identified.

4.1 Gaussian Processes

Definition 50 *A* Gaussian realization *of an infinite graph-network $G = (V, E)$, with prescribed conductance function $c : E \to \mathbb{R}_+$, and* dipoles $(v_x^c)_{x \in V \setminus \{o\}}$, *is a Gaussian process $(X_x)_{x \in V}$ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, where Ω is a sample space; $\mathscr{F}$ a sigma-algebra of events, and $\mathbb{P}$ a probability measure s.t., for $\forall F \in \mathscr{F}(V)$, the random variables $(X_x)_{x \in F}$, are jointly Gaussian with*

$$\mathbb{E}(X_x) = \int_\Omega X_x d\mathbb{P} = 0 \tag{115}$$

and covariance

$$\mathbb{E}(X_x X_y) = k^{(c)}(x, y) = \left\langle v_x^{(c)}, v_y^{(c)} \right\rangle_{\mathscr{H}_E}; \tag{116}$$

i.e., the covariance matrix $(\mathbb{E}(X_x X_y))_{(x,y) \in F \times F}$ is

$$K_F(x, y) := k^{(c)}(x, y) \text{ on } F \times F. \tag{117}$$

Lemma 51 (Jorgensen and Pearse, 2010) *For all $G = (V, E)$, and $c : E \to \mathbb{R}_+$, as specified, Gaussian realizations exist; they are called Gaussian free fields.*

Corollary 52 *Let $G = (V, E)$, $c : E \to \mathbb{R}_+$ be as above; and let $(X_x)_{x \in V}$ be an associated Gaussian free field. Then the point Dirac-masses $(\delta_x)_{x \in V}$ have Gaussian realizations*

$$\widetilde{\delta_x} = c(x) X_x - \sum_{y \sim x} c_{xy} X_y, \quad \forall x \in V. \tag{118}$$

Corollary 53 *Let $G = (V, E)$, and $c : E \to \mathbb{R}_+$ be as above. Let $\{X_x\}_{x \in V}$ be the corresponding Gaussian free field, i.e., with correlation*

$$\mathbb{E}(X_x X_y) = k^{(c)}(x, y) = \left\langle v_x^{(c)}, v_y^{(c)} \right\rangle_{\mathscr{H}_E} \tag{119}$$

where the dipoles $\{v_x^{(c)}\} \subset \mathscr{H}_E$ are computed w.r.t. a chosen (and fixed) based-point $o \in V$, i.e.,

$$\left\langle v_x^{(c)}, f \right\rangle_{\mathscr{H}_E} = f(x) - f(o), \quad \forall f \in \mathscr{H}_E, \ x \in V. \tag{120}$$

Finally, let $R^{(c)}(x, y)$ be the corresponding resistance metric on V. Then

$$\mathbb{E}(X_x X_z) + \mathbb{E}(X_z X_y) \leq \mathbb{E}(X_x X_y) + R^{(c)}(o, z) \tag{121}$$

holds for all vertices $x, y, z \in V$; see Figure 6.

Proof Use Corollary 48, and (112). We have

$$\|v_x - v_y\|_{\mathscr{H}}^2 \leq \|v_x - v_z\|_{\mathscr{H}}^2 + \|v_z - v_y\|_{\mathscr{H}}^2,$$

and (121) now follows from (116). ∎

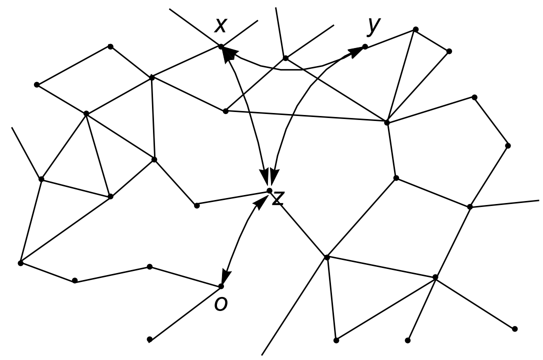


Figure 6: Covariance vs resistance distance $R^{(c)}(o,z)$ for three vertices $x,y,z\in V$.

4.2 Metric Completion

The next theorem illustrates a connection between the universal property of a kernel in a RKHS $\mathscr{H}$, on the one hand, and the distribution of the Dirac point-masses δ_x, on the other. We make "distribution" precise by the quantity $E(x) := \|\delta_x\|^2_{\mathscr{H}}$, the energy of the point-mass at the vertex point x. We introduce a metric completion M, and the universal property of the RKHS $\mathscr{H}$ asserts that the functions from $\mathscr{H}$ are continuous and 1/2-Lipschitz on M, and that they approximate every continuous function on M in the uniform norm. Recall, the vertex set V is equipped with its resistance metric. The universal property here refers to the corresponding metric completion M of the discrete vertex set. In the interesting cases (see e.g., Example 7), M is a continuum; in the case of the example below, the boundary of V is a Cantor set. One expects the value of $E(x)$ to go to infinity as x approaches the boundary M, and this is illustrated in the example; with an explicit formula for $E(x)$.

Of special interest is the class of networks (V,E) where the resistance metric R (on the given vertex vertex-set V) is bounded; see (ii) in Theorem 3 below. This class of networks, for which the diameter of V measured in the resistance metric R is bounded, includes networks having lots of edges with resistors occurring in parallel (Jorgensen and Pearse, 2011).

Theorem 3 *Let $G=(V,E)$, $c:E\rightarrow\mathbb{R}_+$ be as above, and let $R^{(c)}:V\times V\rightarrow\mathbb{R}_+$ be the resistance-metric in (112). Let M be the metric completion of $\left(V,R^{(c)}\right)$. Then:*

(i) For every $f\in\mathscr{H}$, the function

$$V\ni x\longmapsto f(x)\in\mathbb{C} \tag{122}$$

extends by closure to a uniformly continuous function $\widetilde{f}:M\mapsto\mathbb{C}$.

(ii) If $R^{(c)}$ is assumed bounded, then the RKHS $\mathscr{H}$ is an algebra under point-wise product:

$$(f_1f_2)(x)=f_1(x)\,f_2(x),\quad f_i\in\mathscr{H},\ i=1,2,\ x\in V. \tag{123}$$

(iii) If M is compact, then $\{\widetilde{f}\mid f\in\mathscr{H}\}$ is dense in $C(M)$ in the uniform norm.

Proof The assertions in (i) follow from the following two estimates:

Let $f \in \mathscr{H}$, then

$$|f(x) - f(y)|^2 \leq \|f\|^2_{\mathscr{H}} R^{(c)}(x,y), \quad \forall x, y \in V; \tag{124}$$

and

$$|f(x)| \leq |f(o)| + R^{(c)}(o,x)^{\frac{1}{2}}. \tag{125}$$

The estimates in (124)-(125), in turn, follow from Corollaries 47 and 48.

To prove (ii), we compute the energy-norm of the product $f_1 \cdot f_2$ where $f_i \in \mathscr{H}$, $i = 1, 2$; and we use Corollary 47:

$$\begin{aligned}
& \sum_x \sum_y c_{xy} |f_1(x) f_2(x) - f_1(y) f_2(y)|^2 \\
= & \sum_x \sum_y c_{xy} |(f_1(x) - f_1(y)) f_2(x) + f_1(y)(f_2(x) - f_2(y))|^2 \\
\leq & \sum_x \sum_y c_{xy} \left(|f_1(x) - f_1(y)|^2 + |f_2(x) - f_2(y)|^2\right) \cdot \left(|f_2(x)|^2 + |f_1(y)|^2\right) \\
& \text{(by Schwarz inside)} \\
\leq & \left(\|f_1\|^2_\infty + \|f_2\|^2_\infty\right) \cdot \left(\|f_1\|^2_{\mathscr{H}} + \|f_2\|^2_{\mathscr{H}}\right);
\end{aligned}$$

and we note that the right-side is finite subject to the assumption in (ii).

Proof of (iii): We are assuming here that M is *compact*, and we shall apply the Stone-Weierstrass theorem to the subalgebra

$$\left\{\widetilde{f} \mid f \in \mathscr{H}\right\} \subset C(M). \tag{126}$$

Indeed, the conditions for Stone-Weierstrass are satisfied: The functions on LHS in (126) form an algebra, by (ii), closed under complex conjugation; and it separates points in M by Corollary 48. ∎

Example 7 (The binary tree) *Let $A = \{0,1\}$, and $M := \prod_{\mathbb{N}} A$ the infinite Cartesian product, as a Cantor space. Set $V :=$ all finite words:*

$$V = \bigcup_{n\in\mathbb{N}} \left\{(\alpha_1, \alpha_2, \cdots, \alpha_n) \;\middle|\; \alpha_i \in \{0,1\}\right\}; \tag{127}$$

and set $l((\alpha_1, \alpha_2, \cdots, \alpha_n)) =: n$.

For $\omega = (\omega_k)_1^\infty \in M$, set

$$\omega\big|_n := (\omega_1, \omega_2, \cdots, \omega_n) \in V. \tag{128}$$

For two points $\omega, \omega' \in M$, we shall need the number

$$l(\omega \cap \omega') = \sup\left\{n \,:\, \omega\big|_n = \omega'\big|_n\right\}. \tag{129}$$

Let $r:\mathbb{N}\rightarrow\mathbb{R}_{+}$ *be given such that*

$$r\left(\emptyset\right)=0,\quad\sum_{n\in\mathbb{N}}r\left(n\right)<\infty. \tag{130}$$

For conductance function $c:E\rightarrow\mathbb{R}_{+}$*, set*

$$c_{\alpha,\left(\alpha t\right)}=\frac{1}{r\left(l\left(\alpha\right)\right)},\quad\forall\alpha\in V,\;t\in\left\{0,1\right\}. \tag{131}$$

One checks that, when (130) holds, then

$$\lim_{n,m\rightarrow\infty}R^{(c)}\left(\omega\big|_{n},\omega\big|_{m}\right)=0.$$

Consider the graph $G_{2}=(V,E)$ *where the edges are "lines" between* α *and* (αt)*, where* $t\in\{0,1\}$*. See Figure 7.*

Lemma 54 *With the settings above, the metric completion* $\widetilde{R^{(c)}}$ *w.r.t. the resistance metric on* V *is as follows: For* $\omega,\omega'\in M$ *(see Figure 9),*

$$\widetilde{R^{(c)}}\left(\omega,\omega'\right)=2\sum_{n=l(\omega\cap\omega')}^{\infty}r\left(n\right). \tag{132}$$

Let $\mathscr{H}$ *be the corresponding energy-Hilbert space* $\simeq$ *the RKHS of* k_{c}*. For* $\alpha\in V$*, let* δ_{α} *be the Dirac-mass at the vertex point* α*. Then*

$$\left\Vert \delta_{\alpha}\right\Vert _{\mathscr{H}}^{2}=\frac{2}{r\left(l\left(\alpha\right)\right)}+\frac{1}{r\left(l\left(\alpha\right)-1\right)}. \tag{133}$$

(See Figure 8.)

Proof To see this, note that α has the three neighbors sketched in Figure 7, i.e., α^{*}, $(\alpha0)$, and $(\alpha1)$, where α^{*} is the one-truncated word,

$$\widetilde{R^{(c)}}\left(\omega,\omega'\right)=2\sum_{n=l(\omega\cap\omega')}^{\infty}r\left(n\right). \tag{134}$$

One checks that when (130) is assumed, then the conditions in point (iii) of the theorem are satisfied. ■

Corollary 55 *Now return to the discrete restriction of Brownian motion in Section 3.1. Set* $V=\{x_{1},x_{2},x_{3},\cdots\}$ *where the points* $\{x_{i}\}_{i=1}^{\infty}$ *are prescribed such that* $x_{1}<x_{2}<\cdots<x_{i}<x_{i+1}<\cdots$*. We turn* V *into a weighted graph* G *as follows: The edges* E *in* G *are nearest neighbors; and we define a conductance function* $c:E\rightarrow\mathbb{R}_{+}$ *by setting*

$$c_{x_{i}x_{i+1}}:=\frac{1}{x_{i+1}-x_{i}}, \tag{135}$$

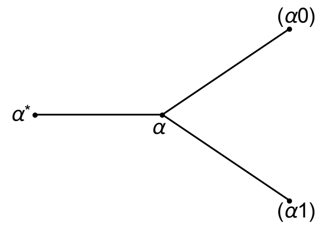
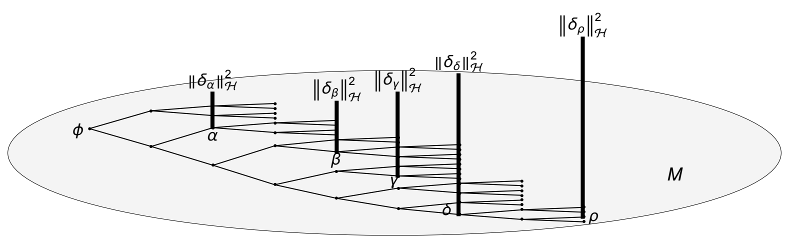
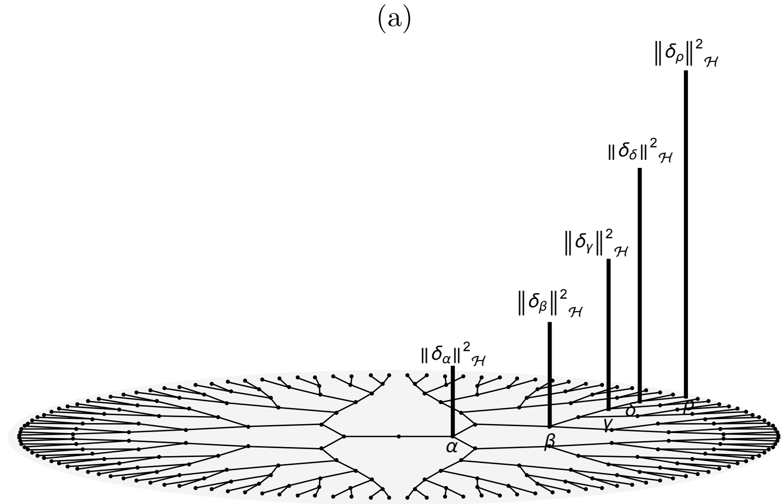


Figure 7: Edges in G_2.



(a)



(b)

Figure 8: Histogram for $\|\delta_\alpha\|^2_{\mathscr{H}}$ as vertices $\alpha \in V$ approach the boundary. See (133), and note $\|\delta_\alpha\|^2_{\mathscr{H}} \to \infty$ as $\alpha \to M$.

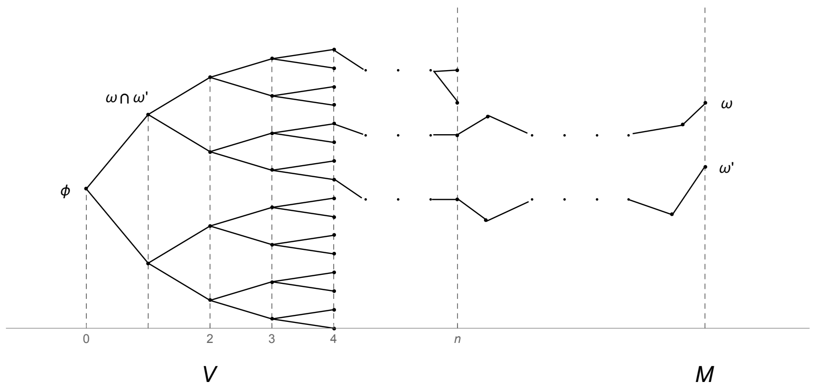


Figure 9: The binary tree and its boundary, the Cantor-set.

and Laplace operator,

$$(\Delta f)(x_i) = \frac{1}{x_{i+1} - x_i}(f(x_i) - f(x_{i+1})) + \frac{1}{x_i - x_{i-1}}(f(x_i) - f(x_{i-1})). \tag{136}$$

Then the RKHS associated with the Green's function of Δ in (136) agrees with that from the kernel construction in Section 3.1, i.e., the discrete Cameron-Martin Hilbert space.

Proof Immediate from the previous Proposition and its corollaries. ■

Acknowledgments

The co-authors thank the following colleagues for helpful and enlightening discussions: Professors Daniel Alpay, Sergii Bezuglyi, Ilwoo Cho, Ka Sing Lau, Paul Muhly, Myung-Sin Song, Wayne Polyzou, Gestur Olafsson, Keri Kornelson, and members in the Math Physics seminar at the University of Iowa. We are grateful to the referees for their care to details and for their kind and very helpful suggestions. We have revised following them all.

References

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)*, volume 59 of *Oper. Theory Adv. Appl.*, pages 30–77. Birkhäuser, Basel, 1992.

Daniel Alpay and Harry Dym. On a new class of structured reproducing kernel spaces. *J. Funct. Anal.*, 111(1):1–28, 1993.

Daniel Alpay and Palle Jorgensen. Reproducing kernel Hilbert spaces generated by the binomial coefficients. To appear. *ArXiv e-prints*, 2015.

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*, volume 61 of *Oper. Theory Adv. Appl.*, pages 1–27. Birkhäuser, Basel, 1993.

Daniel Alpay, Palle Jorgensen, Ron Seager, and Dan Volok. On discrete analytic functions: products, rational functions and reproducing kernels. *J. Appl. Math. Comput.*, 41(1-2): 393–426, 2013.

Daniel Alpay, Palle Jorgensen, and Dan Volok. Relative reproducing kernel Hilbert spaces. *Proc. Amer. Math. Soc.*, 142(11):3889–3895, 2014.

Nachman Aronszajn. La théorie des noyaux reproduisants et ses applications. I. *Proc. Cambridge Philos. Soc.*, 39:133–153, 1943.

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.

Brighid Boyle, Kristin Cekala, David Ferrone, Neil Rifkin, and Alexander Teplyaev. Electrical resistance of N-gasket fractal networks. *Pacific J. Math.*, 233(1):15–40, 2007.

Ola Bratteli and Palle Jorgensen. *Wavelets Through a Looking Glass*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2002.

Andrea Caponnetto, Charles A. Micchelli, Massimiliano Pontil, and Yiming Ying. Universal multi-task kernels. *J. Mach. Learn. Res.*, 9:1615–1646, 2008.

Ilwoo Cho and Palle Jorgensen. Free probability induced by electric resistance networks on energy Hilbert spaces. *Opuscula Math.*, 31(4):549–598, 2011.

Felipe Cucker and Steve Smale. On the mathematical foundations of learning. *Bull. Amer. Math. Soc. (N.S.)*, 39(1):1–49, 2002.

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(1):49–65, 2010.

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.

Haakan Hedenmalm and Pekka J. Nieminen. The Gaussian free field and Hadamard's variational formula. *Probab. Theory Related Fields*, 159(1-2):61–73, 2014.

Palle Jorgensen and Erin P.J. Pearse. A Hilbert space approach to effective resistance metric. *Complex Anal. Oper. Theory*, 4(4):975–1013, 2010.

Palle Jorgensen and Erin P.J. Pearse. Resistance boundaries of infinite networks. In *Random walks, boundaries and spectra*, volume 64 of *Progr. Probab.*, pages 111–142. Birkhäuser/Springer Basel AG, Basel, 2011.

Palle Jorgensen and Erin P.J. Pearse. A discrete Gauss-Green identity for unbounded Laplace operators, and the transience of random walks. *Israel J. Math.*, 196(1):113–160, 2013.

Samuel Karlin and Zvi Ziegler. Some inequalities of total positivity in pure and applied mathematics. In *Total positivity and its applications (Jaca, 1994)*, volume 359 of *Math. Appl.*, pages 247–261. Kluwer Acad. Publ., Dordrecht, 1996.

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.

Sneh Lata and Vern Paulsen. The Feichtinger conjecture and reproducing kernel Hilbert spaces. *Indiana Univ. Math. J.*, 60(4):1303–1317, 2011.

Yi Lin and Lawrence D. Brown. Statistical properties of the method of regularization with periodic Gaussian reproducing kernel. *Ann. Statist.*, 32(4):1723–1743, 2004.

Edward Nelson. Kernel functions and eigenfunction expansions. *Duke Math. J.*, 25:15–27, 1957.

Kasso A. Okoudjou and Robert S. Strichartz. Weak uncertainty principles on fractals. *J. Fourier Anal. Appl.*, 11(3):315–331, 2005.

Kasso A. Okoudjou, Robert S. Strichartz, and Elizabeth K. Tuley. Orthogonal polynomials on the Sierpinski gasket. *Constr. Approx.*, 37(3):311–340, 2013.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2006.

Bernhard Schlkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. The MIT Press, 1st edition, 12 2001.

Oded Schramm and Scott Sheffield. A contour line of the continuum Gaussian free field. *Probab. Theory Related Fields*, 157(1-2):47–80, 2013.

John Shawe-Taylor and Nello Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

Steve Smale and Ding-Xuan Zhou. Online learning with Markov sampling. *Anal. Appl. (Singap.)*, 7(1):87–113, 2009.

Ingo Steinwart. On the influence of the kernel on the consistency of support vector machines. *J. Mach. Learn. Res.*, 2:67–93, March 2002.

Robert S. Strichartz. Transformation of spectra of graph Laplacians. *Rocky Mountain J. Math.*, 40(6):2037–2062, 2010.

Robert S. Strichartz and Alexander Teplyaev. Spectral analysis on infinite Sierpiński fractafolds. *J. Anal. Math.*, 116:255–297, 2012.

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, pages 1041–1052. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013.

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