

NEW CHARACTERIZATIONS OF REPRODUCING KERNEL HILBERT SPACES AND APPLICATIONS TO METRIC GEOMETRY

Daniel Alpay and Palle E.T. Jorgensen

Communicated by Aurelian Gheondea

Abstract. We give two new global and algorithmic constructions of the reproducing kernel Hilbert space associated to a positive definite kernel. We further present a general positive definite kernel setting using bilinear forms, and we provide new examples. Our results cover the case of measurable positive definite kernels, and we give applications to both stochastic analysis and metric geometry and provide a number of examples.

Keywords: reproducing kernel, positive definite functions, approximation, algorithms, measures, stochastic processes.

Mathematics Subject Classification: 46E22, 43A35.

1. INTRODUCTION

Recently many researchers have made use of reproducing kernels in attacking diverse areas and problems from approximation and optimization; see [1, 6, 15, 17, 19, 20]. While there is a direct link from specific positive definite kernels, the link to the corresponding reproducing kernel Hilbert space (RKHS) is a rather abstract one, typically, one is faced only with an abstract completion, and the links to computation is often blurred. Our present aim is to offer two concrete approaches to RKHSs, explicit, and algorithmic. Our approaches are dictated directly by problems in constructive approximation and optimization.

Let S be some set and let $k(t, s)$ be positive definite on S . It is well known (see [4, 13, 16]) that one can associate to k a uniquely defined Hilbert space (which we will denote by $\mathfrak{H}(k)$) of functions defined on S with reproducing kernel $k(t, s)$, meaning that for every $s \in S$ and $f \in \mathfrak{H}(k)$, the function $t \mapsto k(t, s)$ belong to $\mathfrak{H}(k)$ and

$$\langle f(\cdot), k(\cdot, s) \rangle_{\mathfrak{H}(k)} = f(s).$$

The classical construction of the space $\mathfrak{H}(k)$ goes as follows. Consider the linear span $\mathfrak{L}(k)$ of the functions $k_s : u \mapsto k(u, s)$ endowed with the form

$$\langle k_s, k_t \rangle_k = k(t, s), \quad t, s \in S. \quad (1.1)$$

Then, one readily proves that the form (1.1) is well defined, and has the reproducing kernel property. Thus $(\mathfrak{L}(k), \langle \cdot, \cdot \rangle_k)$ is a pre-Hilbert space. Its metric space completion is unique up to a metric space isometry. One then shows that a uniquely defined completion exists, which is made of functions, and which still has the reproducing kernel property.

In general there is no direct characterization of the elements of $\mathfrak{H}(k)$ directly as functions on S (notable exceptions include for instance the Fock space and the Bergman space). In the present work we give two new global constructions of $\mathfrak{H}(k)$ and a global characterization of its elements. The first construction is in terms of a projective limit, and the second one uses a space of measures associated to the kernel. In the latter, we associate an explicit norm, for every positive kernel, and valid for every element in the reproducing kernel Hilbert space; see formula (5.4).

Another space plays an important role in our constructions besides the space $\mathfrak{L}(k)$, namely the linear span $\mathfrak{E}(k)$ of the Dirac measures $\delta_s, s \in S$, endowed with the Hermitian form defined first by

$$\langle \delta_s, \delta_t \rangle = k(t, s), \quad t, s \in S. \quad (1.2)$$

and extended by linearity to the finite linear combinations of Dirac measures. In this case, (1.2) generates a positive, but possibly degenerate, form, and one needs to mod-out via the linear space of finite linear combinations of Dirac measures which are self-orthogonal with respect to this Hermitian form.

We now describe the outline of the paper. In Section 2, starting from a fixed positive definite function (kernel) $k(t, s)$ defined on $S \times S$, we derive an associated metric d_k on S . Its properties are outlined; as well as its applications inside the paper. Section 3 deals with a duality lemma important in the main constructions inside the paper: The starting point here is a general pre-Hilbert space, so a normed vector space where the norm is defined from an inner product. Hence, we get a corresponding dual norm, i.e., a dual Banach space. With the use of a new transform, we show that this dual Banach space is a Hilbert space. In Section 4, starting from a fixed positive definite function (kernel) $k(t, s)$ defined on $S \times S$, we present a constructive approach to the reproducing kernel Hilbert space (RKHS) $\mathfrak{H}(k)$. This will make use of the following five steps: (i) the filter of all finite subsets of S , (ii) an explicit system of finite-rank operators defined on functions on S , and (iii) an algorithm which related the operators defined from two finite sets, with one contained in the other. With this, we establish (iv) a Kolmogorov consistency relation, and (v) we show that the constructive and algorithmic realization of the RKHS $\mathfrak{H}(k)$ follows directly from this (Theorem (2)). It is helpful to draw the following parallel between our particular choice of filter (i.e., the filter of all finite subsets) which deals with the most general setting, and other hand, related filters for special cases, in the literature. In more detail, on the one hand, we have (i) for a given positive definite kernel function $k(t, s)$ on a set S , our present

filter consists of all finite subsets of S . By contrast, (ii) for special RKHS constructions, one has a variety of alternative choices of filters. They are used in special cases of RKHS constructions from harmonic analysis; for example, for Hardy spaces $\mathbf{H}_2(D)$, with D some complex domain, it is popular to pick a filter consisting of closed curves inside D . There are interesting similarities and differences between (ii) special cases, special filters, and (i) our case: the general case, and with our choice of filter. Motivated by applications, in Section 5, we introduce a pre-Hilbert space of measures on S , with the measures defined relative to the Borel sigma-algebra from the metric d_k . Our duality lemma from Section 3 will be used in a characterization of the RKHS norm in $\mathfrak{H}(k)$, Theorem 5.4. The proof of this result is of independent interest, and is carried out in detail, in Section 6. Sections 8 and 9 in turn deal with a RKHS construction for a class of measurable positive definite kernels. Finally, Section 10 contains applications to stochastic processes and Section 11 contains applications to geometric measure theory.

2. THE d_k -METRIC

Our present starting point is a fixed set S and a positive definite function (kernel) $k(t, s)$ defined on $S \times S$. From this we derive an associated metric d_k on S . We shall use the name “metric” even for the cases when d_k might in fact not separate the points in S . It will be important for our subsequent construction to have d_k derived from first principles. For example, we shall specify our Borel sigma-algebra of subsets in S with use of the metric. This will be important in our constructive realization of the reproducing kernel Hilbert space (RKHS) $\mathfrak{H}(k)$ associated with the given kernel k . (Naturally, once $\mathfrak{H}(k)$ is available, then the metric d_k will be an easy byproduct. But our point in the present section is to derive the metric d_k directly from an easy matrix consideration.) For our present purpose, the metric d_k will be available at the outset, and we show in Proposition 2.4 that the kernel function k is automatically jointly continuous relative to the product metric.

The following is well known, and will hold even when $k(t, s)$ is only Hermitian. A proof is outlined for completeness.

Lemma 2.1. *Let S be a set and let $k(t, s)$ be positive definite on S . The formula (1.1) induces a well defined non-degenerate Hermitian form on the linear span $\mathfrak{L}(k)$ of the functions k_s , $s \in S$, and the space $\mathfrak{L}(k)$ endowed with this form is a reproducing kernel pre-Hilbert space.*

Proof. The form (1.1) extended to the linear span of the kernel functions, is well defined, and the reproducing kernel property stems directly from (1.1). We now prove that the form is non-degenerate. Let $f \in \mathfrak{L}(k)$ be orthogonal to all the elements in the space. We have in particular

$$\langle f, k_s \rangle_k = 0, \quad \forall s \in S$$

and so, using the reproducing kernel property, $f(s) = 0$ for all $s \in S$. □

The norm associated to this form induces a distance via

$$d(f, g) = \|f - g\|_k, \quad f, g \in \mathfrak{L}(k).$$

In particular

$$d(k_t, k_s)^2 = k(t, t) - 2\operatorname{Re} k(t, s) + k(s, s), \quad t, s \in S.$$

The right-hand side of the above equality is non negative since k is positive definite. We set

$$d_k(t, s) = d(k_t, k_s), \quad t, s \in S.$$

We note that d_k is symmetric and satisfies the triangle inequality since, for $t, s, u \in S$ we have

$$\begin{aligned} d_k(t, s) &= \|k_t - k_s\|_k \\ &\leq \|k_t - k_u\|_k + \|k_u - k_s\|_k \\ &= d_k(t, u) + d_k(u, s). \end{aligned}$$

It need not define a distance $d_k(t, s)$ on S , since it may happen that $d_k(t_0, s_0) = 0$ with $t_0 \neq s_0$. We then have

$$k(s, t_0) = k(s, s_0), \quad \forall s \in S.$$

We will say that two points t_0 and s_0 are equivalent if $d(k_{t_0}, k_{s_0}) = 0$. We have an equivalence relation \sim , and we consider the function $k(t, s)$ on the quotient space S/\sim . We set

$$\tilde{k}(\tilde{t}, \tilde{s}) = k(t, s), \quad t \in \tilde{t}, \quad s \in \tilde{s}$$

and

$$\tilde{d}_k(\tilde{t}, \tilde{s}) = d_k(t, s), \quad t \in \tilde{t}, \quad s \in \tilde{s}.$$

The proofs of the following two results are easy and omitted.

Lemma 2.2. \tilde{k} is well defined and positive definite on S/\sim .

Lemma 2.3. \tilde{d}_k is well defined and $(S/\sim, \tilde{d}_k)$ is a metric space.

In the sequel we assume, without loss of generality, that d_k is indeed a distance on S (i.e. we consider S/\sim rather than S).

Proposition 2.4. The function $k(t, s)$ on $S \times S$ is jointly continuous with respect to the product metric $D_k((t_1, s_1), (t_2, s_2)) = d_k(t_1, t_2) + d_k(s_1, s_2)$ and the diagonal function $t \mapsto k(t, t)$ is continuous with respect to d_k .

Proof. We first consider the second claim. The triangle inequality in $\mathfrak{L}(k)$ gives

$$\left| \|k_t\|_k - \|k_s\|_k \right| \leq \|k_t - k_s\|_k,$$

which can be rewritten as

$$\left| \sqrt{k(t, t)} - \sqrt{k(s, s)} \right| \leq d_k(t, s).$$

Thus the map $t \mapsto \sqrt{k(t, t)}$ is continuous with respect to d_k and so is its square.

To prove the joint continuity of $k(t, s)$ with respect to D_k we write:

$$\begin{aligned} |k(t_1, s_1) - k(t_2, s_2)| &\leq |k(t_1, s_1) - k(t_1, s_2)| + |k(t_1, s_2) - k(t_2, s_2)| \\ &= |\langle k_{s_1} - k_{s_2}, k_{t_1} \rangle| + |\langle k_{s_2}, k_{t_1} - k_{t_2} \rangle| \\ &\leq \sqrt{k(t_1, t_1)}d_k(s_1, s_2) + \sqrt{k(s_2, s_2)}d_k(t_1, t_2). \quad \square \end{aligned}$$

Example 2.5. For the covariance of the Brownian motion, $k(t, s) = t \wedge s$ with $t, s \geq 0$ we have

$$d_k(t, s) = \sqrt{|t - s|}.$$

We remark that:

Proposition 2.6. *The elements of $\mathfrak{L}(k)$ are Lipschitz continuous with respect to d_k*

Proof. The claim follows from the Cauchy-Schwarz inequality. Indeed, let $f \in \mathfrak{L}(k)$ and $t, s \in S$. Then

$$|f(t) - f(s)| = \langle f, k_t - k_s \rangle_k \leq \|f\|_k \cdot \|k_t - k_s\|_k = \|f\|_k \cdot d_k(t, s). \quad \square$$

The metric d_k induces an associated topology and sigma-algebra, denoted by \mathcal{T}_k and \mathcal{B}_k , respectively. We complete \mathcal{B}_k with the sets of outer measure 0, and still denote the completion \mathcal{B}_k .

3. A CANONICAL ISOMETRY OPERATOR

Let $(\mathfrak{H}_0, \langle \cdot, \cdot \rangle_{\mathfrak{H}_0})$ denote a pre-Hilbert space. We take the inner product linear in the first variable and anti-linear in the second variable. Furthermore, we denote by $(\mathfrak{H}, \langle \cdot, \cdot \rangle_{\mathfrak{H}})$ any of its completion. Recall that it is unique up to a Hilbert space isomorphism. Let \mathfrak{H}_0^* denote the Banach anti-dual of \mathfrak{H}_0 , that is the space of anti-linear continuous maps from \mathfrak{H}_0 into \mathbb{C} .

Lemma 3.1. *The dual norm defined by*

$$\|\varphi\|_{\mathfrak{H}_0^*} = \sup_{\substack{h \in \mathfrak{H}_0 \\ \|h\|_{\mathfrak{H}_0}=1}} |\varphi(h)|, \quad \varphi \in \mathfrak{H}_0^*, \tag{3.1}$$

is a Hilbert norm, and \mathfrak{H}_0^ is complete with this norm.*

Proof. It follows from (3.1) that φ is also continuous from the chosen closure \mathfrak{H} of \mathfrak{H}_0 into \mathbb{C} . By Riesz theorem applied to φ as a continuous map from \mathfrak{H} into \mathbb{C} , there exists an element $T\varphi \in \mathfrak{H}$ such that

$$\varphi(h) = \langle T\varphi, h \rangle_{\mathfrak{H}}. \tag{3.2}$$

It follows that

$$\|\varphi\|_{\mathfrak{H}_0^*} = \|T\varphi\|_{\mathfrak{H}},$$

and so the norm (3.1) is defined by the inner product

$$\langle \varphi, \psi \rangle_{\mathfrak{H}_0^*} \stackrel{\text{def.}}{=} \langle T\varphi, T\psi \rangle_{\mathfrak{H}}. \tag{3.3}$$

□

Lemma 3.2. *In the notation of the previous lemma and of its proof, the operator T is a Hilbert space isomorphism from \mathfrak{H}_0^* onto \mathfrak{H} .*

Proof. By (3.3), T is an isometry and hence is enough to show that T has dense range. This follows from (3.2). □

The function

$$K(\varphi, \psi) = \langle T\varphi, T\psi \rangle_{\mathfrak{H}}, \quad \varphi, \psi \in \mathfrak{H}_0^*, \tag{3.4}$$

is positive definite, and independent of the chosen completion of \mathfrak{H}_0 . We associate to it the Hilbert space of functions of the form

$$F(\varphi) = \langle T\varphi, f \rangle_{\mathfrak{H}}, \quad \|F\| = \|f\|_{\mathfrak{H}} \tag{3.5}$$

In the case of the point evaluations we have

$$\overline{f(t)} = \overline{\delta_t}(f) = \langle k(\cdot, t), f \rangle_{\mathfrak{H}}$$

and so

$$T(\delta_t) = k(\cdot, t)$$

and the reproducing kernel (3.4) restricted to the point measures is (with \mathfrak{H}_0 the linear span of the δ_t with the form (1.1))

$$K(\delta_t, \delta_s) = \langle k(\cdot, t), k(\cdot, s) \rangle_{\mathfrak{H}} = k(s, t).$$

Not always are point evaluations in \mathfrak{H}_0^* (in particular \mathfrak{H}_0 need not be a space of functions) but the kernel (3.4) is associated in a natural way to \mathfrak{H}_0 .

4. AN INJECTIVE LIMIT CONSTRUCTION

The idea is to look at the filter \mathcal{F} of finite subsets of S . For any such set F , one considers the finite dimensional reproducing kernel Hilbert space $\mathfrak{L}_F(k)$ with reproducing kernel $k(t, s)|_{F \times F}$. The argument below uses a limit over the filter of all finite subsets. We denote by K_F the corresponding Gram matrix. In order to get only invertible matrices we first extract from S a possibly smaller set S_0 for which the functions $k_t, t \in S_0$ are linearly independent and span the linear pan of the $k_t, t \in S$. We then define

the projection from the vector space of functions defined on $F = \{s_1, \dots, s_m\} \subset S_0$ onto $\mathfrak{L}_F(k)$ by

$$(Q_F f)(s) = \left(k(s, s_1) \ \cdots \ k(s, s_m) \right) K_F^{-1} \underbrace{\begin{pmatrix} f(s_1) \\ \vdots \\ f(s_m) \end{pmatrix}}_{\stackrel{\text{def.}}{=} f|_F}, \tag{4.1}$$

where the variable s runs through F . To ease the notation, we write S instead of S_0 in the sequel.

Lemma 4.1. *The map Q_F has the following properties:*

(1) *interpolation:*

$$(Q_F f)(s_j) = f(s_j), \quad j = 1, \dots, m, \tag{4.2}$$

(2) *projection:*

$$Q_F^2 = Q_F,$$

(3) *norm:*

$$\|Q_F f\|_{\mathfrak{L}_F(k)}^2 = (f|_F)^* K_F^{-1} (f|_F), \tag{4.3}$$

(4) *order preserving: if $F_1 \subset F_2$ are finite subsets of S ,*

$$Q_{F_2} Q_{F_1} = Q_{F_1} Q_{F_2} = Q_{F_1}.$$

Proof. Setting $s = s_i$ on the right-hand side of (4.1) we note that $(k(s_i, s_1) \dots k(s_i, s_m))$ is the i -th row of K_F , and so $(k(s_i, s_1) \dots k(s_i, s_m)) K_F^{-1}$ is the $1 \times m$ row vector with all entries equal to 0 besides the i -th entry equal to 1. Hence (4.2) holds. Applying the preceding argument to $Q_F f$ we get $Q_F^2 = Q_F$. (4.3) follows from

$$\begin{aligned} & \langle Q_F f, Q_F f \rangle_{\mathfrak{L}(k)} \\ &= \left\langle \sum_{k,j=1}^m k(\cdot, s_j) (K_F^{-1})_{jk} \overline{f(s_k)}, \sum_{i,\ell=1}^m k(\cdot, s_i) (K_F^{-1})_{i\ell} \overline{f(s_\ell)} (K_F^{-1})_{\ell i} (K_F)_{ij} \right\rangle_{\mathfrak{L}(k)} \\ &= \sum_{j,k,i,\ell=1}^m f(s_\ell) (K_F^{-1})_{\ell i} (K_F)_{ij} (K_F^{-1})_{jk} \overline{f(s_k)} \\ &= \sum_{j,k,\ell=1}^m f(s_\ell) \left(\sum_{i=1}^m (K_F^{-1})_{\ell i} (K_F)_{ij} \right) (K_F^{-1})_{jk} \overline{f(s_k)} \\ &= \sum_{j,k,\ell=1}^m f(s_\ell) \delta_{\ell j} (K_F^{-1})_{jk} \overline{f(s_k)} \\ &= \sum_{j,k,\ell=1}^m f(s_\ell) \delta_{\ell j} (K_F^{-1})_{jk} \overline{f(s_k)} \\ &= \sum_{j,k=1}^m f(s_j) (K_F^{-1})_{jk} \overline{f(s_k)}, \end{aligned}$$

which is (4.3).

Finally, let F_1 and F_2 be two finite subsets of S such that $F_1 \subset F_2$. Let f be a function defined on S . The functions f and $Q_{F_2}f$ coincide on F_2 and so on F_1 , and so

$$Q_{F_1}f = Q_{F_1}(Q_{F_2}f)$$

and so $Q_{F_1} = Q_{F_1}Q_{F_2}$. Next, $Q_{F_2}(Q_{F_1}f) = Q_{F_1}$ since the span of the kernels for $s \in F_1$ is isometrically included in the span of the kernels for $s \in F_2$. \square

As a definition we introduce the subspace of all the functions on S for which the sup

$$\sup_{F \in \mathcal{F}} \|Q_F f\|_{\mathfrak{L}_F(k)}^2$$

is finite. In the next theorem we show that this is the reproducing kernel Hilbert space associated to k .

Theorem 4.2. *Let $k(t, s)$ be positive definite on S .*

(1) *The following injective limit of Hilbert spaces*

$$\lim_{\substack{F \in \mathcal{F} \\ K_F > 0}} \mathfrak{L}_F(k) \tag{4.4}$$

as $F \rightarrow S$, with norm

$$\|f\|_{\mathfrak{H}}^2 = \sup_{F \in \mathcal{F}} \|Q_F f\|_{\mathfrak{L}_F(k)}^2 \tag{4.5}$$

exists.

(2) *The supremum in (4.5) defines a Hilbert space norm, and the above injective limit coincides with the reproducing kernel Hilbert space with reproducing kernel $k(t, s)$.*

Proof. We first note that when $F_1 \subset F_2$, the space $\mathfrak{L}_{F_1}(k)$ is isometrically included in the space $\mathfrak{L}_{F_2}(k)$. We denote by $J = J_{F_1 F_2}$ the isometric inclusion from $\mathfrak{L}_{F_1}(k)$ into $\mathfrak{L}_{F_2}(k)$. The operator $JJ^* = Q_{F_1 F_2}$ is the orthogonal projection from $\mathfrak{L}_{F_2}(k)$ onto $\mathfrak{L}_{F_1}(k)$. By Kolmogorov’s consistency theorem there is a uniquely defined Hilbert space \mathfrak{H} such that

$$\begin{array}{ccc} \mathfrak{H} & \xrightarrow{i} & \mathfrak{L}_{F_2}(k) \\ \swarrow i & & \searrow J \\ & \mathfrak{L}_{F_1}(k) & \end{array} .$$

To conclude the proof we now proceed in a number of steps.

Step 1. (4.5) defines a norm. The arguments for a norm are easily checked, but it is important to note that the supremum is in fact an increasing limit since

$$F_1 \subset F_2 \implies \|Q_{F_1} f\|_{\mathfrak{L}_{F_1}(k)}^2 \leq \|Q_{F_2} f\|_{\mathfrak{L}_{F_2}(k)}^2 .$$

Step 2. Any Cauchy sequence in the norm (4.5) converges pointwise. To see this we take F to be a singleton.

Step 3. Every Cauchy sequence converges to a function in \mathfrak{H} . Indeed, from (4.5), convergence in the norm implies convergence on finite sets, and the limiting function is obtained in the previous step. It remains to show that the function so obtained belongs to the space (4.4). This follows from the definition of (4.5) since the operators Q_F are defined on all functions, and we use the fact that a Cauchy sequence is uniformly bounded in the (4.5) norm and so gives a uniform bound on the $\|Q_F f\|_{\mathfrak{L}_F(k)}$.

Step 4. The norm is defined by an inner product. Since the norm is defined by an increasing limit, one can take the polarization formula for every F and takes limit to obtain the corresponding inner product.

Step 5. The reproducing kernel property holds. Let $s_0 \in S$ and let $F = \{s_0, s_1, \dots, s_n\} \subset S$, and let $f \in \mathfrak{H}$. We have

$$\begin{pmatrix} k(s_0, s_1) & k(s_0, s_1) & \cdots & k(s_0, s_n) \end{pmatrix} Q_F^{-1} \begin{pmatrix} f(s_0) \\ f(s_1) \\ \vdots \\ f(s_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} f(s_0) \\ f(s_1) \\ \vdots \\ f(s_n) \end{pmatrix} = f(s_0).$$

To conclude one takes the limit for those $F \subset S$ which contain s_0 . □

As a transition to the second construction, we note the following. Denote by $\mathfrak{M}(k)$ the linear span of the delta measures δ_s , and let $\mathfrak{M}_F(k)$ denote the linear span of the δ_s for $s \in F$. The map T_k defined by

$$T_k \left(\sum_{t \in F} z_t \delta_t \right) = \sum_{t \in F} z_t k_t \tag{4.6}$$

is unitary from $\mathfrak{M}_F(k)$ onto $\mathfrak{L}_F(k)$, with matrix representation $k(t, s)|_{F \times F}$.

The map T_k extends to a unitary map between the projective limits $\mathfrak{D}_1(k) = \lim_{F \xrightarrow{\mathcal{F}} S} \mathfrak{M}_F(k)$ onto $\mathfrak{H}(k) = \lim_{F \xrightarrow{\mathcal{F}} S} \mathfrak{L}_F(k)$.

Remark 4.3. Rather than restricting to a subset S_0 one can use the Moore–Penrose inverse $K_F^{[-1]}$ of the matrix K_F ; see [5] for the latter. Then, one needs to use Schur complement formulas. See [7, 11] for the latter.

5. DUAL NORMS AND HILBERT SPACE CONSTRUCTION

Recall that \mathcal{B}_k is the Borel sigma-algebra generated by the metric d_k defined in Section 2. Let ξ be a signed measure on \mathcal{B}_k . It follows from the definition of the latter and the positivity of $k(t, s)$ that the integral

$$\langle \xi, \xi \rangle_k = \iint_{S \times S} \overline{d\xi(t)} k(t, s) d\xi(s) \tag{5.1}$$

is non-negative, but possibly infinite. The order of integration in (5.1) is not important since $k(t, s)$ is continuous with respect to the underlying topology (see Proposition 2.4). For two measures ξ and η for which (5.1) is finite we set

$$\langle \xi, \eta \rangle_k = \iint_{S \times S} \overline{d\eta(t)} k(t, s) d\xi(s).$$

Since (5.1) is a positive (possibly degenerate) Hermitian form, the Cauchy–Schwarz inequality holds

$$|\langle \xi, \eta \rangle_k|^2 \leq \langle \xi, \xi \rangle_k \langle \eta, \eta \rangle_k \tag{5.2}$$

for any pair of signed matrices for which (5.1) is finite.

We will say that two signed measures for which (5.1) is finite are equivalent if

$$\langle \xi_1 - \xi_2, \xi_1 - \xi_2 \rangle_k = 0. \tag{5.3}$$

Thanks to (5.2) we have an equivalence relation, which we denote by \sim_k . We use the same symbol for the equivalence class and for an element in the class, and still denote by $\langle \cdot, \cdot \rangle_k$ the corresponding Hermitian form.

Definition 5.1. We denote by $\mathfrak{M}_1(k)$ the set of equivalence classes of signed measures on \mathcal{B}_k for which (5.1) is finite, and set

$$\mathfrak{B}_1(k) = \{ \xi \in \mathfrak{M}_1(k) : \langle \xi, \xi \rangle_k \leq 1 \}.$$

The space $\mathfrak{M}_1(k)$ endowed with the form (5.1) is a pre-Hilbert space, and its dual (space of anti-linear continuous functionals) $(\mathfrak{M}_1(k))^*$ is complete. We will consider the elements of the previously constructed space $\mathfrak{H}(k)$ as linear functionals on $\mathfrak{M}_1(k)$. By the results of Section 3 $(\mathfrak{M}_1(k))^*$ is a Hilbert space.

Proposition 5.2. *The linear span of the delta measures $\delta_s, s \in S$ is dense in $\mathfrak{M}_1(k)$.*

Proof. Let μ be orthogonal to the measures $\delta_t, t \in S$. Then,

$$\iint_{S \times S} \delta_{t_0}(t) k(t, s) d\mu(s) = 0, \quad \forall t_0 \in S,$$

i.e.

$$\int_S k(t, s) d\mu(s) = 0, \quad \forall t \in S.$$

Hence

$$\langle \mu, \mu \rangle_k = \iint_{S \times S} \overline{d\mu(t)} k(t, s) d\mu(s) = 0$$

and so $\mu = 0$. □

The operator T_k defined in (4.6) extends to $\mathfrak{M}_1(k)$:

Proposition 5.3. *The map T_k defined by*

$$(T_k\mu)(t) = \int_S k(t, s)d\mu(s)$$

is an isometry from $\mathfrak{M}_1(k)$ into $(\mathfrak{M}_1(k))^$, which extends to a unitary map between $\mathfrak{D}_1(k)$ (the Hilbert space closure $\mathfrak{M}_1(k)$) and $(\mathfrak{M}_1(k))^*$.*

Proof. Let $\mu \in \mathfrak{M}_1(k)$. It defines an element in $(\mathfrak{M}_1(k))^*$ via the formula

$$(\widetilde{T_k\mu})(\xi) = \int_S \overline{d\xi(t)}T_k(\mu)(t) = \langle \mu, \xi \rangle$$

which is continuous in the $\mathfrak{M}_1(k)$ topology thanks to (5.2). □

Theorem 5.4. *Let $k(t, s)$ be positive definite on the set S , with associated sigma-algebra of Borel sets \mathcal{B}_k associated to d_k . Let $\mathfrak{M}_1(k)$ be the space of signed measures on \mathcal{B}_k such that*

$$\iint_{S \times S} \overline{d\xi(t)}k(t, s)d\xi(s) < \infty.$$

The space of functions of the form $\widetilde{f}(\xi) = \langle f, \xi \rangle$ with $f \in (\mathfrak{M}_1(k))^$ endowed with the norm*

$$\|\widetilde{f}\| = \|f\|_{\mathfrak{M}_1(k)^*} \tag{5.4}$$

(where $\|f\|_{\mathfrak{M}_1(k)^}$ denotes the dual norm) is a Hilbert space. The restrictions of its elements to the jump measures δ_s is the reproducing kernel Hilbert space with reproducing kernel k .*

6. PROOF OF THEOREM 5.4

Step 1. The dual $(\mathfrak{M}_1(k))^$ is a Hilbert space.*

This follows from Section 3 but we give a specific argument here. By a general result in functional analysis (see e.g. [8]), the dual is a Banach space. We show that its norm is defined by an inner product. To that purpose, let $f \in (\mathfrak{M}_1(k))^*$. It extends to the completion $\mathfrak{D}_1(k)$ of $\mathfrak{M}_1(k)$. Since $\mathfrak{M}_1(k)$ is a pre-Hilbert space, then $\mathfrak{D}_1(k)$ is a Hilbert space and so, by Riesz theorem applied to $\mathfrak{D}_1(k)$ there is an operator U_k from $(\mathfrak{M}_1(k))^*$ into $\mathfrak{D}_1(k)$ such that

$$\langle f, \xi \rangle = \langle U_k f, \xi \rangle_{\mathfrak{D}_1(k)},$$

where the brackets on the left denote the duality between $\mathfrak{M}_1(k)$ and its anti-dual. Setting $f = T_k(\mu)$ in the above equality we obtain

$$\langle U_k T_k(\mu), \xi \rangle = \int_S \overline{(T_k\mu)(s)}d\xi(s) = \langle \mu, \xi \rangle_{\mathfrak{D}_1(k)}$$

and so $U_k T_k = I$.

We define an inner product on $(\mathfrak{M}_1(k))^*$ as follows:

$$\langle f, g \rangle = \langle U_k f, U_k g \rangle_{\mathfrak{D}_1(k)}.$$

For $f = T_k \xi$ and $g = T_k \eta$ two elements of $(\mathfrak{M}_1(k))^*$ we have

$$\langle T_k \xi, T_k \eta \rangle_{(\mathfrak{M}_1(k))^*} = \langle U_k T_k \xi, U_k T_k \eta \rangle_{\mathfrak{D}_1(k)} = \langle \xi, \eta \rangle_{\mathfrak{D}_1(k)}.$$

For $\xi = \eta$ we have

$$\langle T_k \xi, T_k \xi \rangle_{(\mathfrak{M}_1(k))^*} = \|\xi\|_{\mathfrak{D}_1(k)}^2.$$

Step 2. The Banach norm in $(\mathfrak{M}_1(k))^$ coincide with the Hilbert space norm.*

It follows from the fact that the Hilbert norm is the supremum of the inner products on the unit ball of $\mathfrak{M}_1(k)$. But, by the definition of the norm as a supremum we have

$$\begin{aligned} \|T_k \xi\|_{(\mathfrak{M}_1(k))^*}^2 &= \sup_{\eta \in \mathfrak{B}_1(k)} |(T_k \xi)(\eta)|^2 \\ &= \sup_{\eta \in \mathfrak{B}_1(k)} |\langle \eta, \xi \rangle_k|^2 \leq \langle \eta, \eta \rangle_k \langle \xi, \xi \rangle_k \end{aligned}$$

so that $\|T_k \xi\|^2 \leq \langle \xi, \xi \rangle_k$. Taking $\eta = \xi / \sqrt{\langle \xi, \xi \rangle_k}$ leads to equality.

Step 3. $(\mathfrak{M}_1(k))^$ is the reproducing kernel Hilbert space with reproducing kernel*

$$\langle T_k \delta_t, T_k \delta_s \rangle = k(t, s).$$

Let $f \in (\mathfrak{M}_1(k))^*$ and $s \in S$. We have

$$\langle f, T_k \delta_s \rangle = f(s) = \tilde{f}(\delta_s).$$

7. AN EXAMPLE

Every positive definite function is the covariance function of a centered Gaussian process; see e.g. [14, pp. 38–39]. Applying this result to the inner product of a real Hilbert space \mathfrak{H} with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\|\cdot\|$, we obtain a process called *the associated Gaussian space*. The elements of this space are the centered Gaussian variables W_h , $h \in \mathfrak{H}$, with law $N(0, \|h\|^2)$ and such that

$$\mathbb{E}(W_{h_1} W_{h_2}) = \langle h_1, h_2 \rangle, \quad h_1, h_2 \in \mathfrak{H}.$$

Let

$$k(t, s) = \sum_{n=0}^{\infty} \frac{t^n s^n}{(n!)^2},$$

This gives an isometry sending a measure into the corresponding sequence of moments in the weighted l_2 space. Then

$$d\mu \in \mathfrak{M}_1(k) \iff \sum_{n=0}^{\infty} \frac{|\int_{\mathbb{R}} t^n d\mu(t)|^2}{(n!)^2} < \infty$$

(condition on the moments of the measure).

The Gaussian measures

$$d\mu_h(t) = \frac{1}{\sqrt{2\pi}\|h\|} e^{-\frac{t^2}{2\|h\|^2}} dt$$

belong to $\mathfrak{M}_1(k)$. There are explicit formulas for the moments. We recall in particular that

$$\mathbb{E}(W_h^{2n}) = \frac{(2n)!}{2^n n!} \|h\|^{2n}, \quad n = 0, 1, 2, \dots \tag{7.1}$$

Proposition 7.1. *In the above notation,*

$$\langle \mu_{h_1}, \mu_{h_2} \rangle_k = k\left(\frac{\|h_1\|_2}{2}, \frac{\|h_2\|_2}{2}\right), \quad h_1, h_2 \in \mathfrak{H}.$$

Proof. Using (7.1) we can write for $h_1, h_2 \in \mathfrak{H}$:

$$\begin{aligned} \langle \mu_{h_1}, \mu_{h_2} \rangle_k &= \sum_{n=0}^{\infty} \frac{\mathbb{E}(W_{h_1}^{2n})\mathbb{E}(W_{h_2}^{2n})}{((2n)!)^2} \\ &= \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \left(\frac{\|h_1\|_2^2}{2}\right)^n \left(\frac{\|h_2\|_2^2}{2}\right)^n \\ &= k\left(\frac{\|h_1\|_2}{2}, \frac{\|h_2\|_2}{2}\right). \quad \square \end{aligned}$$

We note that this last kernel was studied in [2, 3].

8. MEASURABLE KERNELS: DUAL NORM CONSTRUCTION

In the previous we considered a positive definite function $k(t, s)$ defined for $(t, s) \in S \times S$ and associated to it in a natural way a metric and a corresponding sigma-algebra. The arguments still work when S is endowed ahead of time with a sigma-algebra, say \mathcal{B} , and $k(t, s)$ is assumed only jointly measurable in (t, s) with respect to \mathcal{B} , and has the property that the double integral (5.1)

$$\langle \xi, \xi \rangle_k = \iint_{S \times S} \overline{d\xi(t)} k(t, s) d\xi(s)$$

is positive or possibly infinite for every signed measure on \mathcal{B} . As in Section 5 we define two signed measures ξ_1 and ξ_2 for which the above integral is finite, to be equivalent if (5.3) holds, and the space $\mathfrak{M}_1(k)$ is defined as in Definition 5.1.

Definition 8.1. With the above notation, the function $k(t, s)$ is called positive definite.

For $\xi \in \mathfrak{M}_1(k)$ the operator T_k defined by

$$(T_k \xi)(t) = \int_S k(t, s) d\xi(s)$$

defines an element of $(\mathfrak{M}_1(k))^*$ via

$$(\widetilde{T_k \xi})(\mu) = \langle \xi, \mu \rangle_k.$$

Theorem 5.4 takes now the following form:

Theorem 8.2. *Let (S, \mathcal{B}) be a measurable space, and let $k(t, s)$ jointly measurable for $t, s \in S$ and positive definite on S . The space of functions of the form $\tilde{f}(\xi) = \langle f, \xi \rangle$ with $f \in (\mathfrak{M}_1(k))^*$ endowed with the norm*

$$\|\tilde{f}\| = \|f\|_{(\mathfrak{M}_1(k))^*}.$$

(where $\|f\|_{(\mathfrak{M}_1(k))^*}$ is the dual norm) is a Hilbert space of functions defined on $\mathfrak{M}_1(k)$, with the reproducing kernel $\langle T_k \mu, T_k \nu \rangle$ and reproducing kernel property

$$\tilde{f}(\xi) = \langle f, T_k \xi \rangle_{(\mathfrak{M}_1(k))^*}.$$

Remark 8.3. The framework considered here includes for instance the singular integral kernel appearing in the study of the fractional Brownian motion, namely

$$k(t, s) = |t - s|^{2H-2}$$

with Hurst constant $H \in [1/2, 1)$; see [9]. For Lebesgue measurable functions f and associated signed measure $f(t)dt$ the condition $\langle f(t)dt, f(t)dt \rangle_k < \infty$ can be rewritten as

$$\int_{\mathbb{R}} |\widehat{f}(u)|^2 |u|^{1-2H} du < \infty,$$

as follows from Parseval’s equality and since the distributional Fourier transform of $|t|^{2H-2}$ is $c_H |u|^{1-2H}$ with $c_H = 2\Gamma(2H - 1) \sin((1 - H)\pi)$. See e.g. [10, p. 170]. For more general information on such integral operators, see [18].

9. QUADRATIC FORMS, OPERATOR RANGES AND REPRODUCING KERNELS

Connections between operator ranges and reproducing kernel spaces are well known, but usually considered in the setting of bounded operators. We here consider a more general case, using the theory of quadratic forms. In the notation of the previous section we still assume that S is a measure space, with σ -algebra \mathcal{B} and we fix a positive σ -finite measure μ on S . We define

$$\mathcal{F}_\mu = \{A \in \mathcal{B}; \mu(A) < \infty\}, \tag{9.1}$$

assume that $k(t, s)$ is jointly measurable and positive in the sense that

$$Q(\varphi) = \iint_{S \times S} \overline{\varphi(t)} k(t, s) \varphi(s) (d\mu \times d\mu)(t, s) \geq 0$$

for every φ in the linear span \mathfrak{D} of the characteristic functions 1_A , with $A \in \mathcal{F}_\mu$.

Assuming Q densely defined and closed, Kato's theorem on quadratic forms, see [12], ensures that there is a positive self-adjoint operator T such that

$$Q(\varphi) = \langle \varphi, T\varphi \rangle_{\mathbf{L}_2(S, d\mu)} = \|T^{1/2}\varphi\|_{\mathbf{L}_2(S, d\mu)}^2, \quad \varphi \in \mathfrak{D}.$$

We then defined the associated reproducing kernel Hilbert space to be the range of $T^{1/2}$ with the operator range norm, meaning

$$\langle T^{1/2}\varphi, T^{1/2}\psi \rangle_T = \langle \varphi, (I - \pi)\psi \rangle_{\mathbf{L}_2(S, d\mu)},$$

where π is the orthogonal projection onto the kernel (null space) of T . We then have

$$\begin{aligned} \langle T\varphi, T\psi \rangle_T &= \langle T\varphi, \psi \rangle_{\mathbf{L}_2(S, d\mu)}, \\ \langle T^{1/2}\varphi, T1_A \rangle_T &= \langle T\varphi, 1_A \rangle_{\mathbf{L}_2(S, d\mu)} = \int_A (T\varphi)(t) d\mu(t) \end{aligned}$$

and the reproducing kernel is now

$$k(A, B) = Q(1_{A \cap B}) = \iint_{A \times B} d\mu(t) d\mu(s).$$

Within the present setting one can also construct the above reproducing kernel space as a projective limit. We define

$$\mathfrak{H}_A(k) = \left\{ F(t) = \int_A k(t, s) \varphi(s) d\mu(s), \varphi \in \mathfrak{D} \right\}$$

with inner product (with $G(t) = \int_A k(t, s) \psi(s) d\mu(s)$, where $\psi \in \mathfrak{D}$)

$$\langle F, G \rangle = \iint_{A \times A} \overline{\psi(t)} k(t, s) \varphi(s) (d\mu \times d\mu)(t, s)$$

We define the injective limit

$$\mathfrak{H} = \lim_{\substack{A \rightarrow S \\ A \in \mathcal{F}_\mu}} \mathfrak{H}_A(k)$$

and the following arguments are similar to the case of finite space considered earlier.

Proposition 9.1. *Let $A \in \mathcal{F}_\mu$. The map T_A from $\mathbf{L}_2(A, \mu)$ onto \mathfrak{H}_A defined by*

$$T_A(\varphi d\mu) = \int_A k(t, s)\varphi(s)d\mu(s)$$

is an isometry, and $T_A T_A^ = Q_A$ is the projection from $\mathbf{L}_2(S, \mu)$ onto $\mathbf{L}_2(A, \mu)$:*

$$(Q_A g)(t) = \int_A k(t, s)T_A^*(g|_A)(s)d\mu(s)$$

The system of projections $\{Q_A, A \in \mathcal{F}_\mu\}$ is the counterpart of the system $\{Q_F, F \in \mathcal{F}\}$ in Section 4, and when $A \subset B$ we have $Q_A = Q_A Q_B$.

10. AN APPLICATION TO STOCHASTIC PROCESSES

We now give an application to random variables of the construction presented in the previous section. More precisely, we start as before with $k(t, s)$ positive definite on S , and we associate to $k(t, s)$ a measurable positive kernel defined in an underlying probability space. Let thus (Ω, \mathcal{B}, P) be a probability space. For every random variable X with values in S , we set

$$k_X(u, v) = k(X(u), X(v)), \quad u, v \in \Omega$$

to be the induced kernel on $\Omega \times \Omega$ defined from X . For a complex-valued random variable X on Ω we defined a signed measure on S via

$$W_X(P) = P \circ X^{-1}. \tag{10.1}$$

We then have

$$\iint_{\Omega \times \Omega} k(X(u), X(v))dP(u)dP(v) = \iint_{S \times S} (d\overline{(P \circ X^{-1})(t)})k(t, s)d(P \circ X^{-1})(s) \tag{10.2}$$

when the second double integral is finite. We denote by $\mathfrak{M}_1(k_X) \subset \mathfrak{M}_1(k)$ the corresponding set of signed measures. We define a norm on the corresponding random variables $\text{RV}_k(\Omega)$ via

$$\langle P, P \rangle = \langle W_P X, W_P X \rangle_{\mathfrak{M}_1(k)}$$

From (10.3) we obtain the following result.

Theorem 10.1. *The map W_X is an isometry from $\text{RV}_k(P)$ into $\mathfrak{M}_1(k)$.*

We have here the generalization of the notion of distribution of a random variable.

We give a Hilbert structure on a family of random variables on Ω using the previous analysis.

$$k_{X,Y}(u, v) = k(X(u), Y(v)), \quad u, v \in \Omega$$

$$\iint_{\Omega \times \Omega} k(X(u), Y(v))dP(u)dP(v) = \iint_{S \times S} (d\overline{(P \circ X^{-1})(t)})k(t, s)d(P \circ Y^{-1})(s) \tag{10.3}$$

11. HAUSDORFF DISTANCE

Finally we make some connections between our analysis and the Hausdorff distance of two measures in $\mathfrak{M}_1(k)$. Recall that the Hausdorff distance is defined to be

$$d_{\text{Haus}}(\mu, \nu) = \sup_{f \in \text{Lip}_1(S)} \left| \int_S f(t) d\mu(t) - \int_S f(t) d\nu(t) \right|,$$

where Lip_1 denotes the set of functions on S such that

$$\|f\|_{\text{Lip}} = \sup_{\substack{s, t \in S \\ s \neq t}} \frac{|f(t) - f(s)|}{d_k(t, s)} \leq 1.$$

Proposition 11.1. *Let $\mu, \nu \in \mathfrak{M}_1(k)$. It holds that*

$$d_{\text{Haus}}(\mu, \nu) \leq d_{\mathfrak{H}(k)}(\mu, \nu) = \|T_k(\mu) - T_k(\nu)\|_{\mathfrak{H}(k)}.$$

Proof. By Cauchy–Schwarz inequality,

$$\frac{|f(t) - f(s)|}{d_k(t, s)} = \frac{|\langle f, k_t - k_s \rangle_{\mathfrak{H}(k)}}{\|k_t - k_s\|_{\mathfrak{H}(k)}} \leq \|f\|_{\mathfrak{H}(k)}. \quad \square$$

REFERENCES

- [1] R.A. Aliev, C.A. Gadjieva, *Approximation of hypersingular integral operators with Cauchy kernel*, Numer. Funct. Anal. Optim. **37** (2016), no. 9, 1055–1065.
- [2] D. Alpay, M. Porat, *Generalized Fock spaces and the Stirling numbers*, J. Math. Phys. **59** (2018), 063509.
- [3] D. Alpay, P. Jorgensen, R. Seager, D. Volok, *On discrete analytic functions: Products, rational functions and reproducing kernels*, J. Appl. Math. Comput. **41** (2013), 393–426.
- [4] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404.
- [5] A. Ben-Israel, T.N. Greville, *Generalized Inverses*, 2nd ed., CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 15, Springer-Verlag, New York, 2003.
- [6] J. Bouvrie, B. Hamzi, *Kernel methods for the approximation of nonlinear systems*, SIAM J. Control Optim. **55** (2017), no. 4, 2460–2492.
- [7] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
- [8] H. Brezis, *Analyse fonctionnelle*, Masson, Paris, 1987.
- [9] T.E. Duncan, *Some applications of fractional Brownian motion to linear systems*, [in:] *System theory: modeling, analysis and control (Cambridge, MA, 1999)*, volume 518 of *Kluwer Internat. Ser. Engrg. Comput. Sci.*, pages 97–105. Kluwer Acad. Publ., Boston, MA, 2000.

- [10] I.M. Guel'fand, G.E. Shilov, *Les distributions. Tome 1*, Collection Universitaire de Mathématiques, no. 8, Dunod, Paris, 1972, Nouveau tirage.
- [11] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1994, Corrected reprint of the 1991 original.
- [12] T. Kato, *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition.
- [13] H. Meschkovski, *Hilbertsche Räume mit Kernfunktion*, Springer-Verlag, 1962.
- [14] J. Neveu, *Processus aléatoires gaussiens*, Number 34 in Séminaires de mathématiques supérieures. Les presses de l'université de Montréal, 1968.
- [15] P. Saint-Pierre, *Approximation of the viability kernel*, Appl. Math. Optim. **29** (1994), no. 2, 187–209.
- [16] S. Saitoh, *Theory of Reproducing Kernels and its Applications*, vol. 189, Longman Scientific and Technical, 1988.
- [17] A.J. Smola, B. Schölkopf, *On a kernel-based method for pattern recognition, regression, approximation and operator inversion*, Algorithmica **22** (1998), no. 1–2, 211–231.
- [18] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Mathematical Series, vol. 30, Princeton University Press, Princeton, N.J., 1970.
- [19] S.N. Vasil'eva, Yu.S. Kan, *Approximation of probabilistic constraints in stochastic programming problems using a probability measure kernel*, Avtomat. i Telemekh. **80** (2019), no. 11, 93–107.
- [20] M. Yousefi, K. van Heusden, I.M. Mitchell, G.A. Dumont, *Model-invariant viability kernel approximation*, Systems Control Lett. **127** (2019), 13–18.

Daniel Alpay (corresponding author)

alpay@chapman.edu

 <https://orcid.org/0000-0002-7612-3598>

Schmid College of Science and Technology

Chapman University

One University Drive Orange, California 92866, USA

Palle E.T. Jorgensen

palle-jorgensen@uiowa.edu

 <https://orcid.org/0000-0003-2681-5753>

The University of Iowa

Department of Mathematics

14C McLean Hall, Iowa City, IA 52246, USA

Received: November 20, 2020.

Accepted: January 23, 2021.