2002

Hilbert Spaces Induced by Toeplitz Covariance Kernels

Mihaela Teodora Matache
University of Nebraska at Omaha, dvelcsov@unomaha.edu

Valentin Matache
University of Nebraska at Omaha, vmatache@unomaha.edu

Recommended Citation

Matache, M. T., & Matache, V. (2002). Hilbert spaces induced by Toeplitz covariance kernels In B. Pasik-Duncan (Ed.), Stochastic Theory and Control (pp. 319-333). New York, NY: Springer.

HILBERT SPACES INDUCED BY TOEPLITZ COVARIANCE KERNELS

MIHAELA T. MATACHE
AND
VALENTIN MATACHE

Abstract. We consider the reproducing kernel Hilbert space $\mathcal{H}_\mu$ induced by a kernel which is obtained using the Fourier-Stieltjes transform of a regular, positive, finite Borel measure μ on a locally compact abelian topological group Γ. Denote by G the dual of Γ. We determine $\mathcal{H}_\mu$ as a certain subspace of the space $\mathcal{C}_0(G)$ of all continuous function on G vanishing at infinity. Our main application is calculating the reproducing kernel Hilbert spaces induced by the Toeplitz covariance kernels of some well-known stochastic processes.

AMS Subj. Class. Primary: 60B15, Secondary: 60G10, 46C15.
Keywords: Covariance Kernel, Fourier Transform, Reproducing Kernel.

1. Introduction

Let K denote a reproducing kernel on a nonempty set X. Such a kernel is called a Toeplitz kernel if X is an abelian group and there exists a function $\Phi : X \to \mathbb{C}$ such that

$$K(x,y) = \Phi(x-y), \qquad \forall x,y \in X.$$

The covariance kernels associated to wide-sense stationary stochastic processes (see Definition 2 in Section 3 of this paper) are Toeplitz reproducing kernels. Let G be a locally compact, abelian, topological group and K a continuous Toeplitz reproducing kernel on G. A well known theorem of Bochner, [23, 1.4.3], states that K is necessarily induced by a positive, finite, regular Borel measure μ on Γ, the dual of G, in the sense that

$$K(x,y) = \hat{\mu}(y-x) \qquad \forall x,y \in G$$

where $\hat{\mu}$ is the Fourier-Stieltjes transform of μ. For that reason the reproducing kernel Hilbert space (RKHS) induced by such a kernel K is denoted by $\mathcal{H}_\mu$. The main result of this paper is Theorem 2 in which we describe $\mathcal{H}_\mu$ as follows.

If $\hat{\mu} \in L^1_G(dx)$, then μ is absolutely continuous with respect to the Haar measure $d\gamma$ of Γ, there is a continuous function φ such that $\varphi = d\mu/d\gamma$, $d\gamma$-a.e., and $f \in \mathcal{H}_\mu$ if and only if $f \in \mathcal{C}_0(G) \cap L^1_G(dx)$, $\hat{f}(\gamma) = 0$ on $\{\gamma \in \Gamma : \varphi(\gamma) = 0\}$ and

$$\int_{\{\gamma\in\Gamma:\varphi(\gamma)\neq 0\}} \frac{|\hat{f}(\gamma)|^2}{\varphi(\gamma)} d\gamma < \infty.$$

In the text above $\hat{f}$ denotes the Fourier transform of f and $\mathcal{C}_0(G)$ is the algebra of all continuous functions vanishing at infinity on G. The space $L^1_G(dx)$ is calculated with respect to the Haar measure dx of G.

Let $\mathbb{T}$ be the unit circle in the complex plane and $\mathbb{Z}$ the set of all integers. The representation of $\mathcal{H}_\mu$ contained by Theorem 2 generalizes similar results which describe the RKHS induced by a continuous, Toeplitz reproducing kernel in the particular cases $G = \mathbb{T}$, $\Gamma = \mathbb{Z}$, [19, page 84], respectively $G = \Gamma = \mathbb{R}$, [13], [21].

Section 2 of this paper is dedicated to introducing in detail the notions mentioned without many details or explanations in this section. We also prove some preliminary results and technical lemmas necessary to the proof of the main result. Section 3 contains that result (Theorem 2), its particular form for compact, abelian, topological groups (Theorem 3), and several examples of RKHS spaces induced by Toeplitz covariance kernels associated to well known stochastic processes, (such as the increment-process associated to a Poisson process, the Ornstein-Uhlenbeck process, some discrete first order autoregressive processes, and some moving average processes). Theorems 2 and 3 are used to determine those spaces.

There are many applications of the theory of RKHS spaces in various fields. To give an example, T. Kailath, E. Parzen, and some of their co-workers developed techniques of RKHS spaces to solve detection and estimation problems, [4], [5], [6], [13], [14], [21]. RKHS techniques play a central role in these papers. They are used to solve problems of extraction, detection, and prediction of signals in the presence of noise. The main message in [21] is that it is important to know if the signal belongs or not to the RKHS induced by the covariance kernel of the noise. For more on the importance of RKHS theory and its applications in this area of mathematics we refer to [2], [3], [11], [12], [16], [17], to quote only few of many eligible references.

2. Preliminary Results

Let X denote a nonempty set. A reproducing kernel on X is any function $K : X \times X \to \mathbb{C}$ with the property that

$$\sum_{i,j=1}^{n} K(x_i, x_j) c_i \bar{c}_j \geq 0 \qquad \forall x_1, \ldots, x_n \in X, \quad \forall c_1, \ldots, c_n \in \mathbb{C}.$$

Each reproducing kernel K on X induces in a unique way a Hilbert space $\mathcal{H}_K$ consisting of complex valued functions on X called the Reproducing Kernel Hilbert Space (RKHS) induced by K.

For each $y \in X$ denote by k_y the kernel function associated to y, i.e. the function

$$k_y(x) = K(x, y), \qquad x \in X.$$

Denote by $\mathcal{K}$ the set of all kernel functions, i.e.

$$\mathcal{K} = \{k_y : y \in X\}.$$

$\mathcal{H}_K$ is the completion of the linear space Span$\mathcal{K}$ spanned by the functions in $\mathcal{K}$ endowed with the inner product $< \cdot, \cdot >$ determined by the following relation

$$< k_y, k_x >= K(x,y) \qquad \forall x, y \in X.$$

The reason why $\mathcal{H}_K$ is called an RKHS with kernel K is the so called reproducing property

$$f(x) =< f, k_x > \qquad \forall x \in X, \quad \forall f \in \mathcal{H}_K.$$

All these facts are well known. We refer the reader to [1] and [8] for the basics on RKHS.

For each $f \in \mathcal{H}_K$ we denote by $\|f\|$ the $\mathcal{H}_K$-norm of f, while $\|f\|_\infty$ denotes the following quantity

$$\|f\|_\infty := \sup\{f(x) : x \in X\}.$$

Clearly $\|f\|_\infty \in [0, \infty]$, $\forall f \in \mathcal{H}_K$. The following lemma contains an elementary remark.

Lemma 1. *Under the assumptions above, if $\mathcal{K}$ is a norm bounded set then there is $M > 0$ such that*

$$\|f\|_\infty \leq M\|f\| \qquad \forall f \in \mathcal{H}_K. \tag{1}$$

Therefore a norm-convergent sequence in $\mathcal{H}_K$ must also be uniformly convergent on X toward the same limit.

Proof. Since $\mathcal{K}$ is norm bounded there is $M > 0$ such that

$$\|k_y\| \leq M \qquad \forall y \in X.$$

Therefore for each $x \in X$ one can write

$$|f(x)| = | < f, k_x > | \leq \|f\|\|k_x\| \leq M\|f\|.$$

□

Let G be an abelian, locally compact, topological group having the dual group denoted by Γ. It is well known that the reproducing kernels $K(x,y)$ on G of the form $K(x,y) = \Phi(x-y)$ for some $\Phi : G \to \mathbf{C}$ are necessarily induced by a finite, positive, regular Borel measure μ on Γ in the sense that such a μ with the following property always exists

$$\Phi(x) = \int_\Gamma (x,\gamma) d\mu(\gamma) \qquad x \in G. \tag{2}$$

In (2) (x, γ) denotes $\gamma(x)$. We will use this notation all over this paper. Let $\hat{\mu}$ denote the Fourier-Stieltjes transform of μ. Identifying as usual G to the dual group of Γ, observe that (2) can be rewritten in the following form

$$\Phi(x) = \hat{\mu}(-x) \qquad x \in G. \tag{3}$$

In (3) and all over this paper we use additive notation for the group law of G. Since $\hat{\mu}$ is the Fourier-Stieltjes transform of a complex regular Borel measure on a locally compact abelian group, Φ must be uniformly continuous and bounded on G and hence

$$k_y \in \mathcal{U}_b(G) \qquad \forall y \in G$$

where for each y, k_y is the kernel-function associated to the reproducing kernel $K(x,y) = \Phi(x-y)$, $x, y \in G$, Φ and μ are related as in (2), and $\mathcal{U}_b(G)$ is the space of all bounded, uniformly continuous, complex functions on G. Recall also the notation $\mathcal{H}_K = \mathcal{H}_\mu$ for the RKHS with the previously described reproducing kernel. Our first observation on $\mathcal{H}_\mu$ is the following.

Theorem 1. *Under the assumptions above one has that*

$$\mathcal{H}_\mu \subseteq \mathcal{U}_b(G).$$

Proof. Denote as before by $\mathcal{K}$ the set of all kernel-functions of $\mathcal{H}_\mu$. Clearly the linear space spanned by $\mathcal{K}$, Span$\mathcal{K}$ is a dense subset of $\mathcal{H}_\mu$. On the other hand, Span$\mathcal{K} \subseteq \mathcal{U}_b(G)$ and $\mathcal{K}$ is a norm bounded set since one can write

$$\|k_y\|^2 =< k_y, k_y >= \Phi(y-y) = \Phi(0) \qquad \forall y \in G.$$

By the density of Span$\mathcal{K}$ in $\mathcal{H}_\mu$ and Lemma 1, we deduce that

$$\mathcal{H}_\mu \subseteq \mathcal{U}_b(G).$$

□

Corollary 1. *If G is separable, then $\mathcal{H}_\mu$ is also separable.*

Proof. Let $S = \{x_n : n \in I\}$ be a countable dense subset of G. Observe that if $f \in \mathcal{H}_\mu$ is perpendicular to k_{x_n}, $\forall n \in I$, then $f(x_n) = 0$, $\forall n \in I$. Since f is a continuous function and S a dense subset of G, it follows the f is the null function. The immediate consequence of this fact is that Span$\{k_y : y \in S\}$ is dense in $\mathcal{H}_\mu$. Therefore, since S is countable, if one considers the linear span of the vectors in $\{k_y : y \in S\}$ with coefficients chosen in the set $\mathbb{Q}[i]$ of all complex numbers with rational real and imaginary parts, one obtains a countable, dense subset of $\mathcal{H}_\mu$. □

The Haar measures on G and Γ will be denoted by dx and $d\gamma$ respectively. It will always be assumed that they are normalized in such a way that the inversion theorem ([23, 1.5.1]) holds i.e. the following formula holds

$$f(x) = \int_\Gamma \hat{f}(\gamma)(x,\gamma)\, d\gamma \qquad x \in G. \tag{4}$$

In order that (4) holds f must be an $L^1_G(dx)$-function which belongs to the class $B(G)$ of all functions on G which are the Fourier-Stieltjes transforms of complex, Borel measures on Γ. Assume now that μ is absolutely continuous with respect to the Haar measure of Γ, $\varphi \in B(\Gamma)$ and

$$d\mu = \varphi(\gamma)\, d\gamma.$$

Since μ is a finite, positive measure, it follows that the Radon-Nykodim derivative φ is a nonnegative $L^1_\Gamma(d\gamma)$-function. The following lemma shows that the kernel-functions of $\mathcal{H}_\mu$ are complex conjugates of shifted Fourier transforms of φ, more precisely we can prove the following.

Lemma 2. *Let $y \in G$ be arbitrary and fixed. The kernel-function k_y can be calculated by the following formula.*

$$k_y(x) = \overline{(\widehat{y,\gamma)\varphi(}\gamma)(x)} \qquad x \in G. \tag{5}$$

Proof. Indeed, one can write the following.

$$k_y(x) = \Phi(x-y) = \int_\Gamma (x-y,\gamma)\, d\mu(\gamma) =$$
$$\int_\Gamma (x-y,\gamma)\varphi(\gamma)\, d\gamma = \int_\Gamma \overline{\gamma(y-x)}\varphi(\gamma)\, d\gamma =$$
$$\overline{\int_\Gamma (-x,\gamma)(y,\gamma)\varphi(\gamma)\, d\gamma} = \overline{(\widehat{y,\gamma)\varphi(}\gamma)(x)}.$$

□

Corollary 2. *All the kernel-functions of $\mathcal{H}_\mu$ are $L^1_G(dx)$-functions if and only if $\hat{\varphi} \in L^1_G(dx)$.*

Proof. Observe that $k_0 = \overline{\hat{\varphi}}$, hence $k_0 \in L^1_G(dx)$ if and only if $\hat{\varphi} \in L^1_G(dx)$. Given that $k_y(x) = \Phi(x-y) = k_0(x-y)$ and dx is translation-invariant, it follows that $k_y \in L^1_G(dx)$, $\forall y \in G$ if and only if $k_0 \in L^1_G(dx)$. □

Under the assumptions $d\mu/d\gamma = \varphi\ d\gamma$-a.e and $\varphi \in B(\Gamma) \cap L^1_\Gamma(d\gamma)$ one has the following useful formula for the Fourier transform of a kernel-function.

Lemma 3. *If $\varphi \in B(\Gamma) \cap L^1_\Gamma(d\gamma)$ the Fourier transform $\hat{k}_y$ of an arbitrary, fixed kernel-function k_y is given by the following formula.*

$$\hat{k}_y(\gamma) = \overline{(y,\gamma)\varphi(\gamma)} = (-y,\gamma)\varphi(y) \qquad \gamma \in \Gamma. \tag{6}$$

Proof. The equality (6) is a direct consequence of Lemma 2 and the inversion formula (4). Indeed

$$\hat{k}_y(\gamma) = \int_G (-x,\gamma)k_y(x)\, dx = \overline{\int_G (x,\gamma)\overline{k_y(x)}\, dx} =$$

$$\overline{\int_G (x,\gamma)\widehat{(y,\gamma)\varphi(}\gamma)\,dx} = \overline{(y,\gamma)\varphi(\gamma)} = (-y,\gamma)\varphi(y).$$

Above we were able to apply formula (4) to the function $(y,\gamma)\varphi(\gamma)$ because $B(\Gamma)$ is invariant under multiplication by (y,γ), [23, 1.3.3]. □

From now on, we will work under the assumptions in Lemma 3, namely assume that μ is a finite, positive Borel measure on Γ such that $d\mu << d\gamma$, there is a finite, positive Borel measure λ on G such that $d\mu/d\gamma$ is equal $d\gamma$-a.e to φ, the Fourier-Stieltjes transform of λ i.e. such that

$$\varphi(\gamma) = \int_G (-x,\gamma)\,d\lambda(x) \qquad \gamma \in \Gamma, \tag{7}$$

and $\varphi \in L^1_\Gamma(d\gamma)$.

Remark 1. *The assumptions above hold if and only if* $\hat{\mu} \in L^1_G(dx)$.

Proof. By [23, 1.7.3], if $\hat{\mu} \in L^1_G(dx)$, then $d\mu << d\gamma$, $d\mu/d\gamma \in L^1_\Gamma(d\gamma)$, and $d\mu/d\gamma$ is equal $d\gamma$-a.e. to the Fourier transform of $g(x) = \hat{\mu}(-x)$, an $L^1_G(dx)$-function. The converse implication is a direct consequence of [23, 1.5.1]. □

Under these assumptions we introduce a positive, not necessarily finite measure $\tilde{\mu}$ associated to μ as follows.

Definition 1. *Let S denote the following open subset of Γ, $S := \{\gamma \in \Gamma : \varphi(\gamma) \neq 0\}$. Let E denote an arbitrary, fixed, Borel subset of Γ. The measure $\tilde{\mu}$ is the Borel measure on Γ given by the following equality*

$$\tilde{\mu}(E) := \int_{S\cap E} \frac{1}{\varphi(\gamma)}\,d\gamma.$$

From now on we will use the notation $S = \mathrm{supp}\tilde{\mu}$. The following is the last technical lemma we need prior to proving the theorem containing the description of $\mathcal{H}_\mu$.

Lemma 4. *Let μ, $\tilde{\mu}$, λ, and φ be as described above. The following inequality holds.*

$$\frac{\lambda(G)}{\varphi(\gamma)} \geq 1 \qquad \forall \gamma \in \mathrm{supp}\tilde{\mu}. \tag{8}$$

Let f be any function in $B(G) \cap L^1_G(dx)$ such that $\hat{f}(\gamma) = 0$ $d\gamma$-a.e. on $\Gamma \setminus \mathrm{supp}\tilde{\mu}$. For such f the following equality holds.

$$f(x) = \int_\Gamma \hat{f}(\gamma)\bar{\hat{k}}_y(\gamma)\,d\tilde{\mu}(\gamma). \tag{9}$$

Proof. Relation (8) is an immediate consequence of (7), as for equality (9), observe that one can write

$$\int_\Gamma \hat{f}(\gamma)\bar{\hat{k}}_x(\gamma)\, d\tilde{\mu}(\gamma) = \int_{\text{supp}\tilde{\mu}} \hat{f}(\gamma)(x,\gamma)\varphi(\gamma)\frac{1}{\varphi(\gamma)}\, d\gamma =$$

$$\int_\Gamma \hat{f}(\gamma)(x,\gamma)\, d\gamma = f(x).$$

Above we made use of both Lemma 3 and the inversion formula (4). □

3. The Main Results

Let $\mathcal{C}_0(G)$ denote the space of all continuous, complex functions on G which vanish at infinity. We are ready to characterize the space $\mathcal{H}_\mu$ now.

Theorem 2. *If $\hat{\mu} \in L^1_G(dx)$ then the space $\mathcal{H}_\mu$ consists of those functions $f \in L^1_G(dx) \cap \mathcal{C}_0(G)$ which satisfy the following two conditions*

$$\hat{f}(\gamma) = 0 \qquad \forall\gamma \in \Gamma \setminus \text{supp}\tilde{\mu} \tag{10}$$

$$\int_\Gamma |\hat{f}(\gamma)|^2\, d\tilde{\mu}(\gamma) < \infty. \tag{11}$$

Any function $f \in \mathcal{H}_\mu$ has the property $\|f\|_2 < \infty$ where $\|\cdot\|_2$ is the norm of $L^2_G(dx)$.

Proof. Let $\mathcal{H}_0$ denote the space of all functions $f \in L^1_G(dx) \cap \mathcal{C}_0(G)$ satisfying conditions (10) and (11). First we will show that $\mathcal{H}_0$ is complete under the norm induced by the inner product

$$< f, g >= \int_\Gamma \hat{f}(\gamma)\bar{\hat{g}}(\gamma)d\tilde{\mu}(\gamma)$$

$\|f\|_2 < \infty$, $\forall f \in \mathcal{H}_0$, and $\mathcal{H}_0 \subseteq B(G)$. First, observe that Span$\mathcal{K} \subseteq \mathcal{H}_0$. This is a direct consequence of Lemma 3 and the following computation

$$\int_\Gamma |\hat{k_y}(\gamma)|^2 d\tilde{\mu}(\gamma) = \int_{\text{supp}\tilde{\mu}} \varphi(\gamma)\, d\gamma = \mu(\Gamma) < \infty.$$

Denote by $\|\cdot\|$ the norm of $\mathcal{H}_0$. If $f \in \mathcal{H}_0$ then

$$\int_\Gamma |\hat{f}(\gamma)| d\gamma < \infty \quad \text{and} \quad \int_\Gamma |\hat{f}(\gamma)|^2 d\gamma < \infty.$$

Indeed, by Lemma 4

$$\frac{1}{\varphi(\gamma)} \geq \frac{1}{\lambda(G)} \qquad \forall\gamma \in \text{supp}\tilde{\mu}.$$

Therefore

$$\int_\Gamma |\hat{f}(\gamma)|^2 d\gamma = \int_{\text{supp}\tilde{\mu}} |\hat{f}(\gamma)|^2 d\gamma \leq \lambda(G) \int_{\text{supp}\tilde{\mu}} \frac{|\hat{f}(\gamma)|^2}{\varphi(\gamma)} d\gamma = \lambda(G)\|f\|^2 < \infty.$$

Also

$$\int_\Gamma |\hat{f}(\gamma)| d\gamma = \int_{\mathrm{supp}\tilde{\mu}} |\hat{f}(\gamma)| d\gamma \le \sqrt{\int_{\mathrm{supp}\tilde{\mu}} \frac{|\hat{f}(\gamma)|^2}{\varphi(\gamma)} d\gamma} \sqrt{\int_{\mathrm{supp}\tilde{\mu}} \varphi(\gamma) d\gamma} =$$

$$= \|f\| \sqrt{\mu(\Gamma)} < \infty.$$

Observe that we established the inequalities

$$\|\hat{f}\|_1 \le \sqrt{\mu(\Gamma)} \|f\| \qquad \forall f \in \mathcal{H}_0$$

where $\| \cdot \|_1$ is the norm of $L^1_\Gamma(d\gamma)$, and

$$\|\hat{f}\|_2 \le \sqrt{\lambda(G)} \|f\| \qquad \forall f \in \mathcal{H}_0.$$

We will also denote by $\|\cdot\|_1$ and $\|\cdot\|_2$ the norms of $L^1_G(dx)$ and $L^2_G(dx)$ respectively.

Now assume that $(f_n)_n$ is a Cauchy sequence in $\mathcal{H}_0$ endowed with the norm $\| \cdot \|$. Then $(\hat{f}_n)_n$ will be Cauchy in both $L^1_\Gamma(d\gamma)$ and $L^2_\Gamma(d\gamma)$. Since both these spaces are complete, there is a $g \in L^1_\Gamma(d\gamma) \cap L^2_\Gamma(d\gamma)$ such that $\|\hat{f}_n - g\|_1 \to 0$ and $\|\hat{f}_n - g\|_2 \to 0$. The reason why the limit is the same modulo equality $d\gamma$-a.e. is the fact that convergent sequences in L^p-spaces have subsequences converging a.e. toward the limit-function, [24, 3.12]. Also, the sequence $(\hat{f}_n)_n$ is Cauchy in the norm $\|\cdot\|_\infty$ because $\|\hat{f}_n\|_\infty \le \|f_n\|_1$, $\forall n$, [23, 1.2.4]. Denote by h the uniform limit of $(\hat{f}_n)_n$. Clearly $h = g$ $d\gamma$-a.e. and $h \in B(\Gamma)$ because, by Bochner's theorem [23, 1.4.3], $B(\Gamma)$ is closed with respect to uniform convergence on Γ. So we established that $h \in B(\Gamma) \cap L^1_\Gamma(d\gamma) \cap L^2_\Gamma(d\gamma)$, $\|\hat{f}_n - h\|_1 \to 0$, $\|\hat{f}_n - h\|_\infty \to 0$, and hence $h(\gamma) = 0$ $\forall \gamma \in \Gamma \setminus \mathrm{supp}\tilde{\mu}$. Let

$$f(x) := \int_\Gamma h(\gamma)(x,\gamma)\, d\gamma.$$

By the inversion theorem $f \in L^1_G(dx) \cap B(G)$ and for each $x \in G$ one can write

$$|f(x) - f_n(x)| = | \int_\Gamma (h(\gamma) - \hat{f}_n(x))(x,\gamma)\, d\gamma|$$

$$\le \|h - \hat{f}_n\|_1 \to 0.$$

Again by the inversion theorem one can see that $\hat{f} = h$. Since $h \in L^1_\Gamma(d\gamma) \cap L^2_\Gamma(d\gamma)$ it follows that $\|f\|_2 < \infty$, as a consequence of the Plancherel theorem, [23, 1.6.1].

Let us prove now that f is the $\| \cdot \|$ - limit of $(f_n)_n$. For arbitrary fixed $\epsilon > 0$ consider n_0 a positive integer such that

$$\|f_m - f_n\| < \frac{\epsilon}{2} \qquad \forall m, n \ge n_0.$$

One has that

$$\|f - f_n\|^2 = \int_\Gamma \liminf_{k\to\infty} |\hat{f}_{m_k}(\gamma) - \hat{f}_n(\gamma)|^2 d\tilde{\mu}(\gamma) \le$$

$$\le \liminf_{k\to\infty} \int_\Gamma |\hat{f}_{m_k}(\gamma) - \hat{f}_n(\gamma)|^2 d\tilde{\mu}(\gamma) \le \left(\frac{\epsilon}{2}\right)^2 < \epsilon^2$$

whenever $n \geq n_0$. Above we used Fatou's lemmma [24, 1.28] and the existence of a subsequence $(\hat{f}_{m_k})_k$ of $(\hat{f}_n)_n$ convergent to $h = \hat{f}$ $d\gamma$-a.e. on Γ. So $\mathcal{H}_0$ is a Hilbert space and since $\mathcal{H}_0 \subseteq L^1_G(dx) \cap B(G)$ one deduces that the reproducing property holds on $\mathcal{H}_0$, i.e.

$$< f, k_x >= f(x) \qquad x \in G.$$

The above equality is a direct consequence of (9). Thus $\mathcal{H}_0$ is an RKHS with kernel

$$K(x,y) =< k_y, k_x >= \int_{\text{supp}\tilde{\mu}} (-y,\gamma)\varphi(\gamma)(x,\gamma)\varphi(\gamma)\frac{1}{\varphi(\gamma)}\, d\gamma = \hat{\mu}(y-x)$$

by Lemma 3. Given the uniqueness of the RKHS associated to a given reproducing kernel it follows that $\mathcal{H}_\mu = \mathcal{H}_0$. □

The statement in the Theorem 2 in the particular case $G = \Gamma = \mathbb{R}$ appears in [13]. Theorem 2 can be formulated in a special way if G is compact. Before stating it in that context we need to make some simple observations and introduce more notations. Recall that if G is compact, then Γ is a complete orthonormal subset of $L^2_G(dx)$, [18]. For each $f \in L^1_G(dx)$ and each $\gamma \in \Gamma$ we denote by $c_\gamma(f)$ the Fourier coefficient of f of index γ, i.e.

$$c_\gamma(f) = \int_G f(x)(-x,\gamma)dx.$$

Denote by $\mathcal{C}(G)$ the algebra of all complex-valued continuous functions on G. One can give the following characterization to the space $\mathcal{H}_\mu$.

Theorem 3. *Let G be a compact, abelian topological group. Let Γ denote its dual group, and let μ be a finite, positive, regular Borel measure on Γ. Then $f \in \mathcal{H}_\mu$ if and only if $f \in \mathcal{C}(G)$,*

$$c_\gamma(f) = 0 \qquad \forall \gamma \in \Gamma \setminus \text{supp}\tilde{\mu} \tag{12}$$

and

$$\sum_{\gamma \in \text{supp}\tilde{\mu}} \frac{|c_\gamma(f)|^2}{\mu(\{\gamma\})} < \infty. \tag{13}$$

Proof. Since dx is a finite measure any function in $\mathcal{C}(G)$ is an $L^1_G(dx)$-function. On the other hand, conditions (12) and (13) are exactly (10) and (11) in our context, since Γ is a discrete topological group, [23, 1.7.3]. Also, $\hat{\mu}$ is automatically in $L^1_G(dx)$ when G is compact because $\hat{\mu}$ is continuous on G and hence bounded. □

Theorem 3, in the particular case $G = \mathbb{T}$ and $\Gamma = \mathbb{Z}$ appears in [19, page 84]. In the following we will illustrate the utility of Theorem 2 by calculating the reproducing kernel space associated to some stochastic processes. Let $(X_t)_{t \in I}$ be a stochastic process. The RKHS generated by the covariance kernel of a stochastic process is often a valuable instrument. We will designate the aforementioned reproducing kernel Hilbert space as the RKHS associated to the stochastic process.

Definition 2. *A process $(X(t))_{t\in S}$ is called wide-sense stationary, if it has constant mean and the autocorrelation function $K(s,t) = \mathbf{E}[X(s)X(t)]$, $s,t \in S$, depends only on the difference $t-s$.*

The index set S is assumed to be the subset of a group. Theorem 2 can be used to calculate the RKHS associated to wide-sense stationary processes. We give several examples in the following. Recall that if K is a reproducing kernel on X, then for each nonempty subset E of X, the restriction of K to $E \times E$ is a reproducing kernel on E and the RKHS induced by this second kernel is simply the space of the restrictions to E of all functions in $\mathcal{H}_K$. We will use this fact in some of the examples without mentioning it each time.

Example 1. *Let $(N(t))_{t\geq 0}$ be a Poisson process, and define its increment process as follows $X(t) := N(t+1) - N(t), t \geq 0$. Denote by dx the Lebesgue measure on the real line. The* RKHS *associated to $(X(t))_{t\geq 0}$ is the space of all functions g which are restrictions to $[0,\infty)$ of functions $f \in \mathcal{C}_0(\mathbb{R}) \cap L^1_{\mathbb{R}}(dx)$ with the following properties*

$$\hat{f}(2k\pi) = 0 \qquad \forall k \in \mathbb{Z}, k \neq 0 \tag{14}$$

and

$$\int_{-\infty}^{\infty} \frac{|\hat{f}(x)|^2 x^2}{\sin^2(x/2)}\, dx < \infty. \tag{15}$$

Proof. Let $G = \mathbb{R}$ and $\Gamma = \mathbb{R}$, (see [23] for the fact that they are duals of each other). Let $d\mu = \sin^2(x/2)/(\sqrt{2\pi}(x/2)^2)dx$. Clearly $(\sin^2(x/2)/(x/2)^2) \in L^1_{\mathbb{R}}(dx)$. Working as usual with the normalized Haar measure $dx/\sqrt{2\pi}$, one obtains by a straightforward calculation

$$\hat{\mu}(x) = \begin{cases} 1-|x| & \text{if} \quad |x| \leq 1 \\ 0 & \text{if} \quad |x| > 1 \end{cases} .$$

This is a continuous, compactly supported function and hence belongs to $L^1_{\mathbb{R}}(dx)$. Therefore Theorem 2 can be applied to the kernel $K(t,s) = \nu\hat{\mu}(s-t)$ whose restriction to $[0,\infty)$ is the covariance kernel of the process $(X(t))_{t\geq 0}$, [20]. The positive constant ν is sometimes called the intensity of the Poisson process, [20]. Since the function φ in Theorem 2 is in our case $\varphi(x) = \sin^2(x/2)/(\sqrt{2\pi}(x/2)^2)$ whose set of zeros is $\{2k\pi : k \in \mathbb{Z}, k \neq 0\}$, one gets the characterization above. □

Our next example is concerned with the following stochastic process.

Definition 3. *The stationary Ornstein-Uhlenbeck process is the unique Gaussian process with mean zero and covariance kernel $K(t,s) = (\sigma^2/(2\beta))e^{-\beta|t-s|}$ $t,s \in \mathbb{R}$.*

Its associated RKHS is described in the next example.

Example 2. *The* RKHS *associated to the stationary Ornstein-Uhlenbeck process consists of all functions $f \in \mathcal{C}_0(\mathbb{R}) \cap L^1_{\mathbb{R}}(dx)$ which satisfy the following condition*

$$\int_{-\infty}^{\infty} |\hat{f}(x)|^2 (\beta^2 + x^2)\, dx < \infty. \tag{16}$$

Proof. Using again the notations of Theorem 2, consider $\varphi(x) = \sigma^2/(\sqrt{2\pi}(\beta^2 + x^2))$. It is easy to check that $\hat{\varphi}(x) = (\sigma^2/(2\beta))e^{-\beta|x|}$, [23, 1.5.3]. Clearly φ satisfies the assumptions in Theorem 2 and condition (10) is vacuously satisfied by all functions $f \in \mathcal{C}_0(\mathbb{R}) \cap L^1_{\mathbb{R}}(dx)$ since φ has no zeros. □

Let us consider some examples of discrete processes now.

Example 3. *Let r and σ be real constants, $0 < r < 1$, $(A_n)_{n\geq 0}$ a sequence of zero mean uncorrelated random variables such that*

$$\mathrm{Var}(A_0) = \frac{\sigma^2}{1-r^2} \quad \text{and} \quad \mathrm{Var}(A_n) = \sigma^2 \qquad \forall n > 0.$$

The first order autoregressive process AR(1) $(X_n)_{n\geq 0}$ *is defined as follows.* $X_0 = A_0$, $X_n = rX_{n-1} + A_n$. *The* RKHS $\mathcal{H}$ *associated to this process has the following description.* $\mathcal{H}$ *consists of all absolutely-summable sequences $(w_n)_{n\geq 0}$ of complex numbers which are restrictions to the set $\mathbb{N}$ of non-negative integers of sequences $(z_n)_{n\in\mathbb{Z}}$ of complex numbers with the following property*

$$\sum_{n=-\infty}^{\infty} \left((1+r^2)|z_n|^2 - r(\bar{z}_{n+1} + \bar{z}_{n-1})z_n \right) < \infty.$$

Proof. A straightforward computation leads to the formula

$$\mathrm{Cov}(X_n, X_m) = \frac{\sigma^2}{1-r^2} r^{|m-n|}.$$

Let

$$\varphi(e^{i\theta}) := \frac{\sigma^2}{1-r^2} \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}.$$

Let $G = \mathbb{Z}$ and $\Gamma = \mathbb{T}$ (see [23] for the fact that they are each other' s dual). The Fourier transform of φ is the sequence of its Fourier coefficients calculated with respect to the standard orthonormal basis $\{e^{in\theta} : n \in \mathbb{Z}\}$ of $\Gamma = \mathbb{T}$. Obviously, if $d\mu := \varphi(e^{i\theta})\frac{d\theta}{2\pi}$ (where $d\theta$ is the arc-length measure on $\mathbb{T}$), one has that

$$\hat{\varphi} = \left(\frac{r^{|n|}\sigma^2}{1-r^2} \right)_{n\in\mathbb{Z}}.$$

Therefore $\hat{\varphi}(m-n) = \mathrm{Cov}(X_n, X_m)$ $\forall m, n \in \mathbb{Z}$ and hence Theorem 2 can be used to calculate the RKHS induced by the covariance kernel of $(X_n)_{n\geq 0}$. Indeed, since $0 < r < 1$, the sequence $\hat{\varphi}$ is absolutely summable, i.e. $\hat{\mu} \in L^1_G(dx)$. A straightforward computation leads to the following simpler representation of φ

$$\varphi(e^{i\theta}) = \frac{\sigma^2}{|1 - re^{i\theta}|^2}.$$

Applying Theorem 2 to the groups $G = \mathbb{Z}$, $\Gamma = \mathbb{T}$ and the measure $d\mu = \varphi(e^{i\theta})\frac{d\theta}{2\pi}$ one obtains that $\mathcal{H}_\mu$ consists of those absolutely summable sequences $(z_n)_{n\in\mathbb{Z}}$ of complex numbers with the property that

$$\int_{-\pi}^{\pi} \left| \sum_{n=-\infty}^{\infty} z_n e^{in\theta} \right|^2 |1 - re^{i\theta}|^2 d\theta < \infty. \tag{17}$$

Note that $\varphi(e^{i\theta})$ is never zero, so (17) is the only condition $(z_n)_{n\in\mathbb{Z}}$ must satisfy besides being absolutely summable. The absolute summability of $(z_n)_{n\in\mathbb{Z}}$ and a straightforward computation lead to the following alternative expression of (17)

$$\sum_{n=-\infty}^{\infty} ((1 + r^2)|z_n|^2 - r(\bar{z}_{n-1} + \bar{z}_{n+1})z_n) < \infty.$$

□

Example 4. *Let q be a positive integer. Consider a sequence $(A_n)_{n\geq 0}$ of random variables with the properties $E[A_n] = E[A_0]$, $\forall n \geq 0$ and* $\mathrm{Var}(A_n) = \sigma^2 > 0$, *$\forall n \geq 0$. Let $(X_n)_{n\geq q}$ be the following moving average process of order q,* MA(q)

$$X_n = \frac{1}{q+1} \sum_{k=0}^{q} A_{n-k} \quad \forall n \geq q.$$

The RKHS *$\mathcal{H}$ associated to this process has the following description. $\mathcal{H}$ consists of those absolutely summable sequences $(w_n)_{n\geq q}$ of complex numbers which are restrictions to the set $\{n \in \mathbb{Z} : n \geq q\}$ of sequences $(z_n)_{n\in\mathbb{Z}}$ for which the following conditions hold*

$$\sum_{n=-\infty}^{\infty} z_n e^{\frac{2nk\pi i}{q+1}} = 0 \qquad \forall k \in \mathbb{Z}, 0 < |k| \leq \frac{q+1}{2} \tag{18}$$

and

$$\int_{-\pi}^{\pi} \frac{\left|\sum_{n=-\infty}^{\infty} z_n e^{in\theta}\right|^2 \sin^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2} + \sin\frac{(q+2)\theta}{2}\sin\frac{q\theta}{2}} d\theta < \infty. \tag{19}$$

Proof. The covariance kernel of $(X_n)_{n\geq q}$ is described by the following

$$\mathrm{Cov}(X_n, X_m) = \begin{cases} \frac{\sigma^2}{(q+1)^2}(q + 1 - |m - n|) & \text{if} \quad |m - n| \leq q \\ 0 & \text{if} \quad |m - n| > q \end{cases}$$

Consider the function $\varphi(e^{i\theta})$ whose sequence of Fourier coefficients is given by

$$c_n = \begin{cases} \frac{\sigma^2}{(q+1)^2}(q+1-|n|) & \text{if} \quad |n| \le q \\ 0 & \text{if} \quad |n| > q \end{cases}$$

Clearly

$$\varphi(e^{i\theta}) = \frac{\sigma^2}{(q+1)^2} \sum_{k=-q}^{q} (q+1-|k|)e^{ik\theta}.$$

Straightforward computations lead to the following simpler representation of φ

$$\varphi(e^{i\theta}) = \frac{\sigma^2}{(q+1)^2} \left[1 + \frac{\sin\frac{(q+2)\theta}{2} \sin\frac{q\theta}{2}}{\sin^2\frac{\theta}{2}} \right].$$

Applying Theorem 2 to $G = \mathbb{Z}$, $\Gamma = \mathbb{T}$, $d\mu = \varphi(e^{i\theta})\frac{d\theta}{2\pi}$ one gets that $\mathcal{H}_\mu$ consists of those absolutely summable sequences $(z_n)_{n\in\mathbb{Z}}$ of complex numbers for which conditions (18) and (19) hold. □

References

[1] Aronszajn N., *Theory of Reproducing Kernels*, Trans. AMS, 68(1950), 337–404.

[2] Chevet S., *Kernel Associated with a Cylindrical Measure*, Lecture notes in Mathematics 860(1981), 51-84, Springer-Verlag, New York, 1981.

[3] Duncan T.E., *Some Applications of Fractional Brownian Motion to Linear Systems*, Kluwer Internat. Ser. Engr. and Comput. Sci., 518(2000), 97-105, Kluwer Acad. Publ., Boston, MA, 2000.

[4] Duttweiler D.L. and Kailath T., RKHS *Approach to Detection and Estimation Problems* III: *Generalized Innovations Representations and a Likelihood-Ratio Formula*, IEEE Trans. Information Theory 18(1972), 730-745.

[5] Duttweiler D.L. and Kailath T., RKHS *Approach to Detection and Estimation Problems* IV: *Non-Gaussian Detection*, IEEE Trans. Information Theory 19(1972), 19-28.

[6] Duttweiler D.L. and Kailath T., RKHS *Approach to Detection and Estimation Problems* V: *Parameter Estimation*, IEEE Trans. Information Theory 19(1972), 29-36.

[7] Halmos P.R., *Measure Theory*, Springer-Verlag, New York, Heidelberg, Berlin, 1974.

[8] Halmos P.R., *A Hilbert Space Problem Book*, 2-d edition, Springer-Verlag, New York, Heidelberg, Berlin, 1980.

[9] Hewitt E. and Ross K. A., *Abstract Harmonic Analysis*, Springer-Verlag, Berlin-New York, 1963.

[10] Hida T., *Stationary Stochastic Processes*, Princeton University Press, Princeton, NJ, 1970.

[11] Jain N.C. and Kallianpur G., *A Note on Uniform Convergence of Stochastic Processes*, Ann. Math. Statist. 41(1970), 1360-1362.

[12] Jain N.C. and Monrad D., *Gaussian Measures in Certain Function Spaces*, Lecture Notes in Math. 860(1981), 246-256, Springer-Verlag, New York, 1981.

[13] Kailath T., RKHS *Approach to Detection and Estimation Problems* I: *Deterministic Signals in Gaussian Noise* , IEEE Trans. Information Theory 17(1971), 530-549.

[14] Kailath T. and Wienert H.L., RKHS *Approach to Detection and Estimation Problems* II: *Gaussian Signal Detection*, IEEE Trans. Information Theory 21(1975), 15-23.

[15] Karlin S. and Taylor H.M., *An Introduction to Stochastic Modelling*, 3-d edition, Academic Press, New York, 1998.

[16] Koski T. and Sundar P., *Two Applications of Reproducing Kernel Hilbert Spaces in Stochastic Analysis*, Stochastic in Finite and Infinite Dimensions, 195-206, Trends Math., Birkhäuser Boston, Boston, MA,2001.

[17] Le Page R., *Subgroups of Paths and Reproducing Kernels*, Ann. of Probability, 1(1973), 345-347.

[18] Loomis L.H., *An Introduction to Abstract Harmonic Analysis*, D. Van Nostrand, Princeton, 1953.

[19] Neveu J., *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco, 1965.

[20] Parzen E., *Stochastic Processes*, Holden-Day, San Francisco, 1962.

[21] Parzen E., *Probability Density Functionals and Reproducing Kernel Hilbert Spaces*, Proceedings of the Symposium on Time Series Analysis (1962-Brown University, Editor M. Rosenblatt), pp.155-169, Wiley, New York, 1963.

[22] Patarasarathy K.R., *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

[23] Rudin W., *Fourier Analysis on Groups*, Wiley, New York, 1962.

[24] Rudin W., *Real and Complex Analysis*, McGraw-Hill, New York, 1966.

[25] Stein E.M. and Weiss G., *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton 1971.

Department of Mathematics, University of Nebraska, Omaha, NE 68182-0243

dmatache@mail.unomaha.edu

vmatache@mail.unomaha.edu