J. Math. Anal. Appl. 349 (2009) 201–210

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

# Orthogonality from disjoint support in reproducing kernel Hilbert spaces

Haizhang Zhang

*Department of Mathematics, Syracuse University, 215 Carnegie, Syracuse, NY 13244, USA*

ARTICLE INFO

*Article history:*
Received 15 April 2008
Available online 23 August 2008
Submitted by M. Milman

*Keywords:*
Orthogonality
Disjoint support
Reproducing kernels
Reproducing kernel Hilbert spaces
Translation invariant kernels
Sobolev spaces

ABSTRACT

We investigate reproducing kernel Hilbert spaces (RKHS) where two functions are orthogonal whenever they have disjoint support. Necessary and sufficient conditions in terms of feature maps for the reproducing kernel are established. We also present concrete examples of finite dimensional RKHS and RKHS with a translation invariant reproducing kernel. In particular, it is shown that a Sobolev space has the orthogonality from disjoint support property if and only if it is of integer index.



## 1. Introduction

Let $C(X)$ denote the set of all the continuous functions on a prescribed topological space $X$. The support $\operatorname{supp} f$ of a function $f \in C(X)$ is the closure in $X$ of the subset

$$\Omega_f := \{x \in X\colon\ f(x) \neq 0\}.$$

Assume that we have a Hilbert space $\mathcal{H}$ all of whose elements are continuous functions on $X$. We say that disjoint support implies orthogonality in $\mathcal{H}$ if for all $f, g \in \mathcal{H}$ satisfying $\operatorname{supp} f \cap \operatorname{supp} g = \emptyset$ there holds $(f, g)_{\mathcal{H}} = 0$, where $(\cdot,\cdot)_{\mathcal{H}}$ denotes the inner product on $\mathcal{H}$. If disjoint support implies orthogonality in $\mathcal{H}$ we also say that $\mathcal{H}$ has the orthogonality from disjoint support property.

One may take it for granted that a given Hilbert space of functions has the orthogonality from disjoint support property. Although such an expectation is true for most commonly used Hilbert spaces, there exists a large class that does not possess this property. We point this out by presenting two typical examples below.

**Example 1.1** *(Finite topological spaces).* Let $X$ be a finite set with the discrete topology. Without loss of generality, assume that $X = \mathbb{N}_m := \{1, 2, \ldots, m\}$ for some $m \in \mathbb{N}$. Under the discrete topology, every function $a$ on $X$ is continuous, and can be identified as a vector $a := \{a_j\colon\ j \in \mathbb{N}_m\} \in \mathbb{C}^m$. Thus, the space $\mathcal{H}$ of continuous functions on $X$ is identified as $\mathbb{C}^m$. We make it a Hilbert space by choosing an $m \times m$ hermitian and strictly positive definite matrix $A$ and endowing $\mathcal{H}$ with the inner product

$$(a, b)_{\mathcal{H}} := \sum_{j \in \mathbb{N}_m} \sum_{k \in \mathbb{N}_m} a_j \overline{b_k} A_{kj}, \quad a, b \in \mathcal{H}. \tag{1.1}$$

*E-mail address:* hzhang12@syr.edu.

0022-247X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmaa.2008.08.030

Note that for each $a\in\mathcal{H}$, $\operatorname{supp} a=\{j\in\mathbb{N}_m:\ a_j\neq 0\}$. Thus, it can be seen that disjoint support implies orthogonality in $\mathcal{H}$ if and only if $A$ is diagonal.

**Example 1.2** *(Sobolev spaces).* Let $X:=\mathbb{R}$, $s>1/2$ and

$$\mathcal{H}^s(\mathbb{R}):=\left\{f\in L^2(\mathbb{R}):\ \int_{\mathbb{R}}\left|\hat f(\xi)\right|^2\left(1+\xi^2\right)^s d\xi<+\infty\right\} \tag{1.2}$$

with inner product

$$(f,g)_{\mathcal{H}^s(\mathbb{R})}:=\int_{\mathbb{R}}\hat f(\xi)\overline{\hat g(\xi)}\left(1+\xi^2\right)^s d\xi. \tag{1.3}$$

Here $\hat h$ is the Fourier transform of $h$ defined for $h\in L^1(\mathbb{R})$ as

$$\hat h(\xi):=\int_{\mathbb{R}} h(x)e^{-i2\pi x\xi}\,dx,\quad \xi\in\mathbb{R},$$

and for $h\in L^2(\mathbb{R})$ by a standard approximation process [9]. By the Sobolev imbedding theorem (see, for example, [1, p. 97]), $\mathcal{H}^s(\mathbb{R})\subseteq C(\mathbb{R})$. As a corollary to the main result of Section 3, we shall see that disjoint support implies orthogonality in $\mathcal{H}^s(\mathbb{R})$ if and only if $s\in\mathbb{N}$. The set $\mathbb{N}$ of positive integers has zero Lebesgue measure in $(1/2,+\infty)$. In this sense, there are very few Sobolev spaces with the orthogonality from disjoint support property.

We are interested in the Hilbert spaces $\mathcal{H}$ of functions on a prescribed topological space $X$ where point evaluations are always continuous. In other words, we require that for each $x\in X$, the point evaluation $\delta_x:\mathcal{H}\to\mathbb{C}$ given as

$$\delta_x(f):=f(x),\quad f\in\mathcal{H},$$

be a continuous linear functional on $\mathcal{H}$. Such spaces are called *Reproducing Kernel Hilbert Spaces* (RKHS). They are an essential component in the theory of mathematical learning, [5,8,11,16,21,22,27]. They are also responsible for many interesting applications in various fields (see, for example, [4,6,12,20,23]).

The purpose of this paper is to investigate RKHS where disjoint support implies orthogonality. This question occurred to the author in the study of refinable kernels [29]. We shall establish characterizations in terms of feature maps for the corresponding reproducing kernel in the next section. Concrete examples of finite dimensional RKHS and RKHS with a translation invariant reproducing kernel will be presented in Sections 2 and 3, respectively.

## 2. Characterizations in terms of feature maps

Let $\mathcal{H}$ be an RKHS on a topological space $X$. By the Riesz representation theorem applied to the continuous linear functional $\delta_x$ for every $x\in X$, there exists a unique function $K:X\times X\to\mathbb{C}$ such that

$$K(\cdot,x)\in\mathcal{H},\quad \text{for all } x\in X, \tag{2.1}$$

and

$$f(x)=\left(f,K(\cdot,x)\right)_{\mathcal{H}},\quad \text{for all } x\in X \text{ and } f\in\mathcal{H}. \tag{2.2}$$

It can be verified by (2.1) and (2.2) that for any finite subset $\{x_j:\ j\in\mathbb{N}_m\}\subseteq X$, the $m\times m$ matrix

$$\left[K(x_j,x_k)\right]_{j,k\in\mathbb{N}_m}$$

is hermitian and positive semi-definite. In other words, $K$ is a *positive definite kernel* (or *kernel* for short) on $X$, [2]. We interpret Eq. (2.2) as that functions in $\mathcal{H}$ can be reproduced through their inner products with the kernel $K$. Thus, $K$ is usually referred to as the *reproducing kernel* of $\mathcal{H}$. Conversely, if $K$ is a kernel on $X$ then there exists a unique RKHS on $X$ that takes $K$ as its reproducing kernel. It is constructed explicitly as the completion of the linear space

$$\tilde{\mathcal{H}}:=\operatorname{span}\left\{K(\cdot,x):\ x\in X\right\}$$

under an inner product defined for all $x_j,y_k\in X$ and $c_j,d_k\in\mathbb{C}$, $j\in\mathbb{N}_p$, $k\in\mathbb{N}_q$ as

$$\left(\sum_{j\in\mathbb{N}_p}c_jK(\cdot,x_j),\sum_{k\in\mathbb{N}_q}d_kK(\cdot,y_k)\right)_{\tilde{\mathcal{H}}}:=\sum_{j\in\mathbb{N}_p}\sum_{k\in\mathbb{N}_q}c_j\overline{d_k}K(y_k,x_j).$$

To signify the bijective correspondence between RKHS and their reproducing kernels, we shall denote $\mathcal{H}$ as $\mathcal{H}_K$ and call it the RKHS of $K$.

It is well known that $K$ is a positive definite kernel on $X$ if and only if there is a function $\Phi$ from $X$ to a Hilbert space $\mathcal{W}$ such that

$$\overline{\operatorname{span}}\,\Phi(X) = \mathcal{W} \tag{2.3}$$

and

$$K(x, y) = \big(\Phi(x), \Phi(y)\big)_{\mathcal{W}}, \quad x, y \in X. \tag{2.4}$$

The function $\Phi$ is called a *feature map* for $K$. There is a well-known characterization [15,17,21,22,29] of the RKHS $\mathcal{H}_K$ of $K$ in terms of its feature maps.

**Lemma 2.1.** *Suppose that $\Phi : X \to \mathcal{W}$ is a feature map for $K$ satisfying* (2.3) *and* (2.4)*. Then*

$$\mathcal{H}_K = \big\{\big(\Phi(\cdot), u\big)_{\mathcal{W}}\colon u \in \mathcal{W}\big\}$$

*with the inner product*

$$\big(\big(\Phi(\cdot), u\big)_{\mathcal{W}}, \big(\Phi(\cdot), v\big)_{\mathcal{W}}\big)_{\mathcal{H}_K} = (v, u)_{\mathcal{W}}, \quad u, v \in \mathcal{W}. \tag{2.5}$$

We assume that $\Phi$ is continuous from $X$ to $\mathcal{W}$. By (2.4), this is equivalent to saying that $K$ is a continuous function on $X \times X$. By the Schwartz inequality, we get for $f = (\Phi(\cdot), u)_{\mathcal{W}}$ where $u \in \mathcal{W}$ that

$$\big|f(x) - f(y)\big| \leqslant \|u\|_{\mathcal{W}} \big\|\Phi(x) - \Phi(y)\big\|_{\mathcal{W}} = \|f\|_{\mathcal{H}_K} \big\|\Phi(x) - \Phi(y)\big\|_{\mathcal{W}}.$$

It follows that $f$ is continuous on $X$. Since by Lemma 2.1, $(\Phi(\cdot), u)_{\mathcal{W}}$, $u \in \mathcal{W}$, exhaust all the functions in $\mathcal{H}_K$, $\mathcal{H}_K \subseteq C(X)$.

We are in a position to present a necessary and sufficient condition for disjoint support to imply orthogonality in $\mathcal{H}_K$. For each subset $U \subseteq X$, we denote by $\Phi(U)$ the image of $U$ under $\Phi$, and $\Phi(U)^\perp$ the set of elements in $\mathcal{W}$ that are orthogonal to $\Phi(U)$. We also make the convention that $\Phi(\emptyset) := \{0\}$.

**Theorem 2.2.** *Suppose that $\Phi$ is a continuous function from $X$ to a Hilbert space $\mathcal{W}$ that satisfies* (2.3)*, and $K$ is a kernel on $X$ defined by* (2.4)*. Then disjoint support implies orthogonality in $\mathcal{H}_K$ if and only if for all open subsets $U, V$ of $X$ satisfying $U \cup V = X$ there holds*

$$\Phi(U)^\perp \perp \Phi(V)^\perp, \tag{2.6}$$

*or equivalently, $\Phi(U)^\perp \subseteq \overline{\operatorname{span}}\,\Phi(V)$, the closure of* $\operatorname{span}\Phi(V)$ *in $\mathcal{W}$.*

**Proof.** Suppose that (2.6) holds true for all open $U, V \subseteq X$ such that $U \cup V = X$. Let $f, g \in \mathcal{H}_K$ with $\operatorname{supp} f \cap \operatorname{supp} g = \emptyset$. By Lemma 2.1, there exist $u, v \in \mathcal{W}$ such that

$$f = \big(\Phi(\cdot), u\big)_{\mathcal{W}}, \qquad g = \big(\Phi(\cdot), v\big)_{\mathcal{W}} \tag{2.7}$$

and

$$(f, g)_{\mathcal{H}_K} = (v, u)_{\mathcal{W}}. \tag{2.8}$$

Set $U := X \setminus \operatorname{supp} f$ and $V := X \setminus \operatorname{supp} g$. Then $U, V$ are open in $X$ and

$$X \setminus (U \cup V) = (X \setminus U) \cap (X \setminus V) = \operatorname{supp} f \cap \operatorname{supp} g = \emptyset.$$

Thus, $U \cup V = X$. Since $f$ vanishes on $X \setminus \operatorname{supp} f$, $u$ is orthogonal to $\Phi(U)$ in $\mathcal{W}$. Similarly, $v \perp \Phi(V)$. By (2.6), $u$ is orthogonal to $v$ in $\mathcal{W}$. This follows by (2.8) that $(f, g)_{\mathcal{H}_K} = 0$. Therefore, disjoint support implies orthogonality in $\mathcal{H}_K$.

Conversely, suppose that disjoint support implies orthogonality in $\mathcal{H}_K$. Let $U, V$ be open subsets of $X$ satisfying that $U \cup V = X$. Set $u \in \Phi(U)^\perp$ and $v \in \Phi(V)^\perp$. We introduce two functions $f, g \in \mathcal{H}_K$ as defined by (2.7). Since $u \perp \Phi(U)$, $f$ vanishes on $U$. Thus, $\operatorname{supp} f \subseteq X \setminus U$. Likewise, $\operatorname{supp} g \subseteq X \setminus V$. Since

$$\operatorname{supp} f \cap \operatorname{supp} g \subseteq (X \setminus U) \cap (X \setminus V) = X \setminus (U \cup V) = \emptyset,$$

functions $f$ and $g$ have disjoint support. By our assumption that disjoint support implies orthogonality in $\mathcal{H}_K$, $(f, g)_{\mathcal{H}_K} = 0$, which follows by (2.8) that $(u, v)_{\mathcal{W}} = 0$. Hence, (2.6) holds true. The proof is completed by noting that (2.6) is equivalent to $\Phi(U)^\perp \subseteq \overline{\operatorname{span}}\,\Phi(V)$ since $(\Phi(V)^\perp)^\perp = \overline{\operatorname{span}}\,\Phi(V)$. $\square$

Sometimes, the support of a continuous function may not well reflect the place where $f$ is really "supported." For instance, by slightly modifying the construction of the Cantor set, one can obtain for each $\varepsilon > 0$ a function $f \in C(\mathbb{R})$ such that $\operatorname{supp} f = \mathbb{R}$ while the Lebesgue measure $|\Omega_f|$ of $\Omega_f$ is not greater than $\varepsilon$ (see [19, p. 58]). In some applications, one may hence have to use $\Omega_f$ instead of $\operatorname{supp} f$ (see, for example, [29, Theorem 23]). We next present a necessary and sufficient condition ensuring that $f, g \in \mathcal{H}_K$ are orthogonal whenever $\Omega_f \cap \Omega_g = \emptyset$.

**Corollary 2.3.** *Suppose that $K$ is a kernel on $X$ defined by* (2.4) *through a function $\Phi$ from $X$ to a Hilbert space $\mathcal{W}$ that satisfies* (2.3)*. Then $(f, g)_{\mathcal{H}_K} = 0$ for all $f, g \in \mathcal{H}_K$ with $\Omega_f \cap \Omega_g = \emptyset$ if and only if for all subsets $U \subseteq X$ there holds*

$$\Phi(U)^\perp \perp \Phi(X \setminus U)^\perp, \tag{2.9}$$

*or equivalently,*

$$\Phi(U)^\perp \subseteq \overline{\operatorname{span}}\, \Phi(X \setminus U). \tag{2.10}$$

**Proof.** We reassign $X$ the discrete topology. Then every subset of $X$ is both open and closed under this topology. Thus we have that $\mathcal{H}_K \subseteq C(X)$ and that for each $f \in \mathcal{H}_K$, $\Omega_f = \operatorname{supp} f$. Therefore, it suffices to prove under the discrete topology that disjoint support implies orthogonality in $\mathcal{H}_K$ if and only if (2.9) holds true for all subsets $U$ of $X$.

Suppose first that disjoint support implies orthogonality in $\mathcal{H}_K$. Set $U \subseteq X$ and $V := X \setminus U$. Then $U \cup V = X$ and they are both open. By Theorem 2.2, (2.9) holds true. On the other hand, suppose that there holds (2.10) for all subsets $U \subseteq X$. Set $U, V \subseteq X$ with $U \cup V = X$. Then we have $X \setminus U \subseteq V$. It follows by (2.10) that

$$\Phi(U)^\perp \subseteq \overline{\operatorname{span}}\, \Phi(X \setminus U) \subseteq \overline{\operatorname{span}}\, \Phi(V).$$

Therefore, (2.6) holds true for all subsets $U, V \subseteq X$ with $U \cup V = X$. By Theorem 2.2, disjoint support implies orthogonality in $\mathcal{H}_K$. The proof is complete. □

For applications of Theorem 2.2 and Corollary 2.3, we next investigate finite dimensional RKHS. Let $X$ be a discrete topological space, and $K$ a kernel on $X$ for which $\mathcal{H}_K$ is of finite dimension $m \in \mathbb{N}$. By a theorem on page 347 of [2], there exists an $m \times m$ hermitian and strictly positive definite matrix $A$, and $m$ linearly independent functions $\phi_j$, $j \in \mathbb{N}_m$, on $X$ such that

$$K(x, y) = \big(A\phi(x), \phi(y)\big)_{\mathbb{C}^m}, \quad x, y \in X, \tag{2.11}$$

where $\phi(x) := \{\phi_j(x)\colon\ j \in \mathbb{N}_m\} \in \mathbb{C}^m$, and $(\cdot,\cdot)_{\mathbb{C}^m}$ denotes the standard inner product on $\mathbb{C}^m$. By Lemma 2.1, we have

$$\mathcal{H}_K = \left\{ \sum_{j \in \mathbb{N}_m} c_j \phi_j\colon\ c = \{c_j\colon\ j \in \mathbb{N}_m\} \in \mathbb{C}^m \right\}$$

with inner product

$$\left( \sum_{j \in \mathbb{N}_m} c_j \phi_j, \sum_{j \in \mathbb{N}_m} d_j \phi_j \right)_{\mathcal{H}_K} = \left( \left(A^T\right)^{-1} c, d \right)_{\mathbb{C}^m}.$$

We shall consider the case where $\phi_j$, $j \in \mathbb{N}_m$, have disjoint support, that is,

$$\operatorname{supp} \phi_j \cap \operatorname{supp} \phi_k = \emptyset, \quad j \neq k,\ \ j, k \in \mathbb{N}_m. \tag{2.12}$$

**Proposition 2.4.** *Suppose that $K$ is a kernel on a discrete topological space $X$ defined by* (2.11) *via a hermitian and strictly positive definite matrix $A$ and linearly independent functions $\phi_j$, $j \in \mathbb{N}_m$, on $X$ satisfying* (2.12)*. Then disjoint support implies orthogonality in $\mathcal{H}_K$ if and only if $A$ is diagonal.*

**Proof.** We first identify by (2.11) a feature map $\Phi : X \to \mathbb{C}^m$ for $K$ as

$$\Phi(x) := A^{1/2} \phi(x), \quad x \in X. \tag{2.13}$$

Suppose that $A$ is diagonal. Then $A^{1/2}$ is diagonal as well. Let $a_j$, $j \in \mathbb{N}_m$, be the diagonal components of $A^{1/2}$. Assume that $U, V$ are disjoint open subsets of $X$ such that $U \cup V = X$. Set

$$I_U := \{j \in \mathbb{N}_m\colon\ U \cap \operatorname{supp} \phi_j \neq \emptyset\}, \qquad I_V := \{j \in \mathbb{N}_m\colon\ V \cap \operatorname{supp} \phi_j \neq \emptyset\}.$$

Let $c \in \mathbb{C}^m$ be orthogonal to $\Phi(U)$, namely,

$$\sum_{j \in \mathbb{N}_m} a_j c_j \overline{\phi_j(x)} = 0, \quad x \in U.$$

By (2.12), we have for each $j \in I_U$ that

$$c_j\overline{\phi_j(x)} = 0, \quad x \in \operatorname{supp}\phi_j \cap U,$$

which follows that $c_j = 0$, $j \in I_U$. Similarly, for $d \perp \Phi(V)$ we have $d_j = 0$, $j \in I_V$. Since $U \cup V = X$, $I_U \cup I_V = \mathbb{N}_m$. Therefore, for $c \in \Phi(U)^\perp$ and $d \in \Phi(V)^\perp$, there holds $c_j d_j = 0$, $j \in \mathbb{N}_m$. Consequently, $c \perp d$. By condition (2.6) in Theorem 2.2, disjoint support implies orthogonality in $\mathcal{H}_K$.

Conversely, suppose that $A$ is not diagonal. Then $A^{-1}$ is not diagonal. In other words, there exists $j \neq k \in \mathbb{N}_m$ such that $(A^{-1})_{jk} \neq 0$. Set $U = X \setminus \operatorname{supp}\phi_j$ and $V = X \setminus \operatorname{supp}\phi_k$. By (2.12), $U \cup V = X$. We also observe by definition (2.13) that

$$\Phi(U)^\perp = \left\{c \in \mathbb{C}^m\colon \left(A^{1/2}c\right)_l = 0,\ l \neq j\right\}$$

and

$$\Phi(V)^\perp = \left\{c \in \mathbb{C}^m\colon \left(A^{1/2}c\right)_l = 0,\ l \neq k\right\}.$$

Thus $A^{-1/2}e_j \in \Phi(U)^\perp$ and $A^{-1/2}e_k \in \Phi(V)^\perp$, where for $l \in \mathbb{N}_m$, $e_l$ denotes the vector in $\mathbb{C}^m$ whose $l$th component is one and whose $l'$th component is zero for $l' \neq l$. We verify that

$$\left(A^{-1/2}e_k, A^{-1/2}e_j\right)_{\mathbb{C}^m} = \left(A^{-1}e_k, e_j\right)_{\mathbb{C}^m} = \left(A^{-1}\right)_{jk} \neq 0.$$

Thus (2.6) is not satisfied. By Theorem 2.2, $\mathcal{H}_K$ does not possess the orthogonality from disjoint support property. The proof is complete. □

We remark that the first example in the introduction may be viewed as a direct consequence of the above proposition. To see this, let $\mathcal{H}$ be the RKHS on $X := \mathbb{N}_m$ with inner product (1.1). Next we set $\phi_j := e_j$, $j \in \mathbb{N}_m$, to observe that the kernel for $\mathcal{H}$ has the form

$$K = \left(\left(A^T\right)^{-1}\phi(\cdot), \phi(\cdot)\right)_{\mathbb{C}^m}.$$

Therefore, by Proposition 2.4, disjoint support implies orthogonality in $\mathcal{H}$ if and only if $A$ is diagonal.

To study orthogonality from disjoint support in a more concrete context, we shall investigate translation invariant kernels in the next section.

## 3. Translation invariant reproducing kernels

In this section, we specify the topological space $X$ to be the real line $\mathbb{R}$ and study RKHS on $\mathbb{R}$ with a reproducing kernel $K$ that is *translation invariant* in the sense that

$$K(x-a, y-a) = K(x, y), \quad \text{for all } x, y, a \in \mathbb{R}.$$

Translation invariant kernels have been extensively investigated for their particular importance in the theory of learning (see, for example, [21,22,24,25,27,29]). Denote by $\mathcal{B}_+(\mathbb{R})$ the set of all the finite positive Borel measures on $\mathbb{R}$. It was established by Bochner [3] that $K$ is a continuous translation invariant reproducing kernel on $\mathbb{R}$ if and only if there exists a $\mu \in \mathcal{B}_+(\mathbb{R})$ such that

$$K(x, y) = \int_{\mathbb{R}} e^{i2\pi(x-y)\xi}\, d\mu(\xi), \quad x, y \in \mathbb{R}. \tag{3.1}$$

Our purpose is to characterize the Borel measure $\mu$ such that for the kernel $K$ defined above, disjoint support implies orthogonality in $\mathcal{H}_K$. To this end, we apply the Lebesgue–Radon–Nikodym theorem (see [19, p. 121]) to $\mu$ to get that there exists a nonnegative $\phi \in L^1(\mathbb{R})$ and a Borel measure $\nu \in \mathcal{B}_+(\mathbb{R})$ that is singular with respect to the Lebesgue measure such that for each Borel subset $U \subseteq \mathbb{R}$

$$\mu(U) = \int_U \phi(\xi)\, d\xi + \nu(U). \tag{3.2}$$

The above decomposition yields a corresponding one for the kernel $K$, that is, $K = K_c + K_s$, where

$$K_c(x, y) := \int_{\mathbb{R}} e^{i2\pi(x-y)\xi}\phi(\xi)\, d\xi, \quad x, y \in \mathbb{R}, \tag{3.3}$$

$$K_s(x, y) := \int_{\mathbb{R}} e^{i2\pi(x-y)\xi}\, d\nu(\xi), \quad x, y \in \mathbb{R}. \tag{3.4}$$

We shall establish a result that can somewhat allow us to reduce the question to the two spaces $\mathcal{H}_{K_c}$ and $\mathcal{H}_{K_s}$. To prepare for it, we set $\mathcal{W}_c$ the Hilbert space of Borel measurable functions $u$ on $\mathbb{R}$ such that $\int_{\mathbb{R}} |u(\xi)|^2 \phi(\xi)\,d\xi < +\infty$ with the inner product $(u,v)_{\mathcal{W}_c} := \int_{\mathbb{R}} u(\xi)\overline{v(\xi)}\phi(\xi)\,d\xi$, $u, v \in \mathcal{W}_c$. Similarly, we let $\mathcal{W}_s$ be the Hilbert space of Borel measurable functions $u$ on $\mathbb{R}$ such that $\int_{\mathbb{R}} |u(\xi)|^2\,d\nu(\xi) < +\infty$ with the inner product $(u,v)_{\mathcal{W}_s} := \int_{\mathbb{R}} u(\xi)\overline{v(\xi)}\,d\nu(\xi)$, $u, v \in \mathcal{W}_s$. Finally, we define $\mathcal{W}$ to be the direct sum of $\mathcal{W}_c$ and $\mathcal{W}_s$. In other words, $\mathcal{W} = \{(u,v)\colon\ u \in \mathcal{W}_c,\ v \in \mathcal{W}_s\}$ with inner product

$$\big((u_1,v_1),(u_2,v_2)\big)_{\mathcal{W}} = (u_1,u_2)_{\mathcal{W}_c} + (v_1,v_2)_{\mathcal{W}_s}.$$

By (3.3) and (3.4), $\Phi_c : \mathbb{R} \to \mathcal{W}_c$ and $\Phi_s : \mathbb{R} \to \mathcal{W}_s$ given for each $x \in \mathbb{R}$ as $\Phi_c(x)(\xi) = \Phi_s(x)(\xi) := e^{i2\pi x\xi}$, $\xi \in \mathbb{R}$, are a feature map for $K_c$ and $K_s$, respectively. Similarly, we observe by (3.1) that a feature map $\Phi : \mathbb{R} \to \mathcal{W}$ for $K$ is

$$\Phi(x) := \big(\Phi_c(x), \Phi_s(x)\big), \quad x \in \mathbb{R}. \tag{3.5}$$

**Proposition 3.1.** *If $\mathcal{H}_K$ has the orthogonality from disjoint support property then so does $\mathcal{H}_{K_c}$ and $\mathcal{H}_{K_s}$.*

**Proof.** Assume that disjoint support implies orthogonality in $\mathcal{H}_K$. Let $U, V$ be open subsets of $X$ with $U \cup V = X$. Set $u_1 \in \Phi_c(U)^\perp$ and $u_2 \in \Phi_c(V)^\perp$. Then it is clear that $(u_1, 0) \in \Phi(U)^\perp$ and $(u_2, 0) \in \Phi(V)^\perp$. Since disjoint support implies orthogonality in $\mathcal{H}_K$, by Theorem 2.2, (2.6) holds true. Consequently, we have

$$(u_1,u_2)_{\mathcal{W}_c} = \big((u_1,0),(u_2,0)\big)_{\mathcal{W}} = 0,$$

which implies that $\Phi_c(U)^\perp \perp \Phi_c(V)^\perp$. Therefore, by Theorem 2.2, $\mathcal{H}_{K_c}$ has the orthogonality from disjoint support property. The result for $\mathcal{H}_{K_s}$ can be proved in a similar way. □

By Proposition 3.1, we shall discuss orthogonality from disjoint support respectively in $\mathcal{H}_{K_c}$ and $\mathcal{H}_{K_s}$. Let us start with $\mathcal{H}_{K_c}$ by recalling some basic facts on distributions [7,9].

Let $\mathcal{D}(\mathbb{R})$ be the space of infinitely differentiable functions on $\mathbb{R}$ with compact support and $\mathcal{S}(\mathbb{R})$ the Schwartz class on $\mathbb{R}$. For the topology on $\mathcal{D}(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$, see, for example, [7]. Denote by $\mathcal{D}'(\mathbb{R})$ and $\mathcal{S}'(\mathbb{R})$ the dual of $\mathcal{D}(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$, respectively. We have the inclusion relations that for all $p \in [1, +\infty]$,

$$\mathcal{D}(\mathbb{R}) \subseteq \mathcal{S}(\mathbb{R}) \subseteq L^p(\mathbb{R}) \cap C(\mathbb{R}), \qquad L^p(\mathbb{R}) \cup C(\mathbb{R}) \subseteq L^1_{\mathrm{loc}}(\mathbb{R}) \subseteq \mathcal{D}'(\mathbb{R}),$$
$$L^p(\mathbb{R}) \cup \{\text{polynomials}\} \subseteq \mathcal{S}'(\mathbb{R}) \subseteq \mathcal{D}'(\mathbb{R}).$$

Elements in $\mathcal{D}'(\mathbb{R})$ are called *distributions* and those in $\mathcal{S}'(\mathbb{R})$ are called *temperate distributions*. Borel measurable functions $f$ on $\mathbb{R}$ for which there exists a polynomial $p$ such that

$$\big|f(\xi)\big| \leqslant \big|p(\xi)\big|, \quad \xi \in \mathbb{R}, \tag{3.6}$$

are temperate distributions.

The support $\operatorname{supp} T$ of a distribution $T$ consists of all the points $x_0 \in \mathbb{R}$ such that for any open neighborhood $U$ of $x_0$ there exists $\varphi \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp}\varphi \subseteq U$ so that $\langle T, \varphi\rangle := T(\varphi) \neq 0$. This definition is consistent with the ordinary one when $T$ is a continuous function. The derivative $T'$ of a distribution $T$ is again a distribution defined by

$$\langle T', \varphi\rangle = -\langle T, \varphi'\rangle.$$

One can see that $\operatorname{supp} T' \subseteq \operatorname{supp} T$ for each $T \in \mathcal{D}'(\mathbb{R})$. For a temperate distribution $T$, its Fourier transform $\hat{T}$ is defined to be a temperate distribution by the dual equation

$$\langle \hat{T}, \varphi\rangle = \langle T, \hat{\varphi}\rangle. \tag{3.7}$$

This definition is also consistent with the one when $T \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$. A result connecting the Fourier transform and support of distributions is the Paley–Wiener theorem (see, for example, [9, p. 293]), which states that if $T \in \mathcal{S}'(\mathbb{R})$ has compact support then $\hat{T}$ extends to an analytic function on $\mathbb{C}$. A consequence is that for compactly supported $T \in \mathcal{S}'(\mathbb{R})$, $\operatorname{supp}\hat{T} = \mathbb{R}$ unless $T$ is the zero distribution.

Let us return to the Hilbert space $\mathcal{H}_{K_c}$, which is identified by Lemma 2.1 as

$$\mathcal{H}_{K_c} = \left\{ f \in C(\mathbb{R})\colon\ \big|\Omega_{\hat{f}} \setminus \Omega_\phi\big| = 0,\ \int_{\mathbb{R}} \big|\hat{f}(\xi)\big|^2 \frac{1}{\phi(\xi)}\,d\xi < +\infty \right\} \tag{3.8}$$

with inner product

$$(f,g)_{\mathcal{H}_{K_c}} = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} \frac{1}{\phi(\xi)}\,d\xi, \quad f, g \in \mathcal{H}_{K_c}. \tag{3.9}$$

It may happen that we cannot even have two nontrivial functions in $\mathcal{H}_{K_c}$ with disjoint support. Such a case occurs when $\phi$ is compactly supported. A typical example is the sinc kernel

$$\operatorname{sinc}(x,y) := \frac{\sin\pi(x-y)}{\pi} = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi(x-y)\xi}\,d\xi, \quad x,y\in\mathbb{R},$$

which plays a fundamental role in the sampling theory [10,13,18]. It can be seen that the $\phi$ for the sinc kernel is the characteristic function of interval $[-\frac{1}{2},\frac{1}{2}]$. In general, suppose that $\operatorname{supp}\phi$ is compact. Note that the condition $|\Omega_{\hat f}\setminus\Omega_\phi|=0$ in (3.8) implies that $\operatorname{supp}\hat f\subseteq\operatorname{supp}\phi$ for each $f\in\mathcal{H}_{K_c}$. By the Paley–Wiener theorem, functions in $\mathcal{H}_{K_c}$ are real-analytic on $\mathbb{R}$. There hence do not exist two nontrivial functions in $\mathcal{H}_{K_c}$ with disjoint support.

Same thing happens if $\frac{1}{\phi(\xi)}$ increases too fast as $|\xi|$ goes to infinity. For example, if $1/\phi$ has an exponential growth, that is, there exist $C,\lambda>0$ such that

$$\frac{1}{\phi(\xi)} \geqslant Ce^{\lambda|\xi|}, \quad \text{for } |\xi| \text{ big enough}$$

then every function $f\in\mathcal{H}_{K_c}$ is the restriction on $\mathbb{R}$ of

$$F(z) := \int_{\mathbb{R}} e^{i2\pi z\xi}\hat f(\xi)\,d\xi, \quad z\in\mathbb{C},\ \bigl|\operatorname{Im}(z)\bigr| < \frac{\lambda}{4\pi},$$

which is analytic over $\{z\in\mathbb{C}\colon\ |\operatorname{Im}(z)|<\lambda/4\pi\}$. Thus, every $f\in\mathcal{H}_{K_c}$ is also real-analytic on $\mathbb{R}$. A concrete example of this case is the important Gaussian kernels in learning theory (see, for example, [14,21,26,28]), which are defined as

$$G_\delta(x,y) := e^{-\delta(x-y)^2}, \quad x,y\in\mathbb{R},\ \delta>0.$$

By the Fourier transform of Gaussian functions,

$$G_\delta(x,y) = \int_{\mathbb{R}} e^{i2\pi(x-y)\xi}\sqrt{\frac{\pi}{\delta}}e^{-\frac{\pi^2}{\delta}\xi^2}\,d\xi, \quad x,y\in\mathbb{R}.$$

Thus, for the Gaussian kernels,

$$\frac{1}{\phi(\xi)} = \sqrt{\frac{\delta}{\pi}}e^{\frac{\pi^2}{\delta}\xi^2}, \quad \xi\in\mathbb{R},$$

which increases much faster than any exponential function.

By the above discussion, to make the question of orthogonality from disjoint support in $\mathcal{H}_{K_c}$ interesting, we have to impose some restrictions on the support of $\phi$ and the growth of $1/\phi$. We shall assume that $\Omega_\phi=\mathbb{R}$ and $1/\phi$ does not increase too fast. A natural choice is to require that $1/\phi$ has at most a polynomial growth. Since by (3.6) locally integrable functions with at most a polynomial growth are temperate distributions, to reach more generality, we assume that

$$\frac{1}{\phi} \in \mathcal{S}'(\mathbb{R}). \tag{3.10}$$

This requirement means that $1/\phi\in L^1_{\mathrm{loc}}(\mathbb{R})$ and $\varphi\to\int_{\mathbb{R}}\frac{1}{\phi(\xi)}\varphi(\xi)\,d\xi$ defines a continuous linear functional on $\mathcal{S}(\mathbb{R})$.

**Theorem 3.2.** *Suppose that $K_c$ is defined by (3.3) through nonnegative $\phi\in L^1(\mathbb{R})$ satisfying $\Omega_\phi=\mathbb{R}$ and the growth condition (3.10). Then disjoint support implies orthogonality in $\mathcal{H}_{K_c}$ if and only if $1/\phi$ is a polynomial.*

**Proof.** Suppose that $1/\phi$ is a polynomial. Since $\phi$ is positive everywhere on $\mathbb{R}$ and $\phi\in L^1(\mathbb{R})$, there exists $m\in\mathbb{N}$ and constants $c_j$, $0\leqslant j\leqslant 2m$, where $c_0,c_{2m}>0$ such that

$$\frac{1}{\phi(\xi)} = \sum_{j=0}^{2m} c_j\xi^j, \quad \xi\in\mathbb{R}.$$

By (3.8), functions in $\mathcal{H}_{K_c}$ are contained by the Sobolev space $\mathcal{H}^m(\mathbb{R})$. Thus, for every $f\in\mathcal{H}_{K_c}$, the distributional derivatives $f^{(j)}$, $0\leqslant j\leqslant 2m$, are in $L^2(\mathbb{R})$. This together with (3.9) and Parseval's identity yields for all $f,g\in\mathcal{H}_{K_c}$ that

$$(f,g)_{\mathcal{H}_{K_c}} = \sum_{j=0}^{2m} c_j\int_{\mathbb{R}} \hat f(\xi)\overline{\hat g(\xi)}\xi^j\,d\xi = \sum_{j=0}^{2m} c_j\frac{i^{j-2\bar j}}{(2\pi)^j}\int_{\mathbb{R}} f^{(\bar j)}(x)\overline{g^{(j-\bar j)}(x)}\,dx, \tag{3.11}$$

where $\bar{j}$ is defined to be $j/2$ if $j$ is even and $(j+1)/2$ if $j$ is odd. Suppose that $f, g \in \mathcal{H}_{K_c}$ have disjoint support. Then since $\operatorname{supp} T' \subseteq \operatorname{supp} T$ for any distribution $T$, $f^{(\bar{j})}$ and $g^{(j-\bar{j})}$ have disjoint support for each $j$. Thus, we get by (3.11) that $(f, g)_{\mathcal{H}_{K_c}} = 0$. Therefore, disjoint support implies orthogonality in $\mathcal{H}_{K_c}$.

Conversely, suppose that $\mathcal{H}_{K_c}$ has the orthogonality from disjoint support property. The assumption that $1/\phi \in \mathcal{S}'(\mathbb{R})$ ensures that for each $\varphi \in \mathcal{S}(\mathbb{R})$

$$\int_{\mathbb{R}} |\varphi(\xi)|^2 \frac{1}{\phi(\xi)}\, d\xi < +\infty.$$

This follows by (3.8) that $\mathcal{D}(\mathbb{R}) \subseteq \mathcal{H}_{K_c}$. Choose $f, g \in \mathcal{D}(\mathbb{R})$ with disjoint support. Since disjoint support implies orthogonality in $\mathcal{H}_{K_c}$, we have by (3.9) that

$$\left\langle \left(\bar{\hat{f}}\frac{1}{\phi}\right)^{\hat{}}, g\right\rangle = \left\langle \bar{\hat{f}}\frac{1}{\phi}, \hat{g}\right\rangle = \int_{\mathbb{R}} \overline{\hat{f}(\xi)}\hat{g}(\xi)\frac{1}{\phi(\xi)}\, d\xi = (g, f)_{\mathcal{H}_{K_c}} = 0.$$

Note that the above equality holds true for all $g \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp} g \subseteq \mathbb{R} \setminus \operatorname{supp} f$. Therefore, we must have

$$\operatorname{supp}\left(\bar{\hat{f}}\frac{1}{\phi}\right)^{\hat{}} \subseteq \operatorname{supp} f. \tag{3.12}$$

The above inclusion holds true for all $f \in \mathcal{D}(\mathbb{R})$. We hence claim that

$$\operatorname{supp}\left(\frac{1}{\phi}\right)^{\hat{}} \subseteq \{0\}. \tag{3.13}$$

Otherwise, assume that we have $a \in \operatorname{supp}(\frac{1}{\phi})^{\hat{}} \setminus \{0\}$. Choose a neighborhood $U$ of $a$ that is free of the origin. By definition, there exists some $\varphi \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp}\varphi \subseteq U$ such that $\langle(\frac{1}{\phi})^{\hat{}}, \varphi\rangle \neq 0$. We set $f := \overline{\varphi(-\cdot)}$ and check by the relation between convolution and the Fourier transform that

$$\left(\bar{\hat{f}}\frac{1}{\phi}\right)^{\hat{}}(0) = \left\langle \left(\frac{1}{\phi}\right)^{\hat{}}, \overline{f(-\cdot)}\right\rangle = \left\langle \left(\frac{1}{\phi}\right)^{\hat{}}, \varphi\right\rangle \neq 0,$$

which implies that $0 \in \operatorname{supp}(\bar{\hat{f}}\frac{1}{\phi})^{\hat{}}$. However, since $\operatorname{supp} f \subseteq -U$, $0 \notin \operatorname{supp} f$. We reach a contradiction with (3.12). Therefore, (3.13) holds. We complete the proof by pointing out the fact that the Fourier transform of a temperate distribution is supported at the origin if and only if the distribution is a polynomial (see [7, p. 103]). □

**Corollary 3.3.** *For $\frac{1}{2} < s < +\infty$, disjoint support implies orthogonality in the Sobolev space $\mathcal{H}^s(\mathbb{R})$ if and only if $s \in \mathbb{N}$.*

**Proof.** By the Sobolev imbedding theorem, $\mathcal{H}^s(\mathbb{R})$ is a RKHS. Let us figure out its reproducing kernel. Set $f \in \mathcal{H}^s(\mathbb{R})$ and $x \in \mathbb{R}$. We get by the inverse Fourier transform that

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi)e^{i2\pi x\xi}\, d\xi = \int_{\mathbb{R}} \hat{f}(\xi)\frac{e^{i2\pi x\xi}}{(1+\xi^2)^s}\left(1+\xi^2\right)^s d\xi = \left(f, G(\cdot, x)\right)_{\mathcal{H}^s(\mathbb{R})},$$

where by (1.3) and Parseval's identity the kernel $G$ is identified as

$$G(t, x) := \int_{\mathbb{R}} e^{i2\pi(t-x)\xi}\frac{1}{(1+\xi^2)^s}\, d\xi, \quad t \in \mathbb{R}.$$

Clearly, the $\phi$ for $G$ is $1/(1+\xi^2)^s$. By Theorem 3.2, disjoint support implies orthogonality in $\mathcal{H}^s(\mathbb{R})$ if and only if $1/\phi = (1+\xi^2)^s$ is a polynomial, which happens if and only if $s \in \mathbb{N}$. □

Our next task is to establish similar results for $\mathcal{H}_{K_s}$. Let $\nu \in \mathcal{B}_+(\mathbb{R})$ be singular with respect to the Lebesgue measure. Denote by $\mathbb{Z}$ the set of all integers. Similar to the continuous case, we assume that $\operatorname{supp}\nu = \mathbb{Z}$ and

$$T_\nu := \sum_{n\in\mathbb{Z}} \frac{\delta_n}{\nu_n} \in \mathcal{S}'(\mathbb{R}), \tag{3.14}$$

where $\nu_n := \nu(\{n\})$, $n \in \mathbb{Z}$, and for all $a \in \mathbb{R}$, $\delta_a$ is the delta distribution at $a$ defined for all $\varphi \in \mathcal{D}(\mathbb{R})$ by $\langle\delta_a, \varphi\rangle := \varphi(a)$. Condition (3.14) is equivalent to that $1/\nu_n$ is slowly increasing in the sense that there exists a polynomial $p$ such that $1/\nu_n \leqslant |p(n)|$, $n \in \mathbb{Z}$ (see [9, p. 340]). Under these assumptions, we identify $\mathcal{H}_{K_s}$ by Lemma 2.1 as

$$\mathcal{H}_{K_s} = \left\{ f \in C(\mathbb{R})\colon\ f \text{ is periodic with period } 1,\ \sum_{n\in\mathbb{Z}} \left|c_n(f)\right|^2 \frac{1}{\nu_n} < +\infty \right\} \tag{3.15}$$

with inner product

$$(f, g)_{\mathcal{H}_{K_s}} = \sum_{n\in\mathbb{Z}} \frac{c_n(f)\overline{c_n(g)}}{\nu_n}, \quad f, g \in \mathcal{H}_{K_s},$$

where $c_n(f)$ is the $n$th Fourier coefficient of $f$ defined as

$$c_n(f) := \int_0^1 f(t)e^{-i2\pi nt}\,dt, \quad n \in \mathbb{Z}.$$

**Theorem 3.4.** *Suppose that $K_s$ is defined by* (3.4) *through $\nu \in \mathcal{B}_+(\mathbb{R})$ satisfying* $\operatorname{supp}\nu = \mathbb{Z}$ *and the growth condition* (3.14)*. Then disjoint support implies orthogonality in $\mathcal{H}_{K_s}$ if and only if there exists a polynomial $q$ such that $1/\nu_n = q(n)$, $n \in \mathbb{Z}$.*

**Proof.** The arguments are similar to those in the proof of Theorem 3.2. The key point is to show that if $\mathcal{H}_{K_s}$ has the orthogonality from disjoint support property then the Fourier transform of $T_\nu$ defined by (3.14) is supported on $\mathbb{Z}$. And the main tool is the Fourier transform of periodic distributions. □

Note that the discussion above much relies on the periodicity of functions in $\mathcal{H}_{K_s}$. Translation invariant kernels defined by a general singular measure deserve further attention.

**Corollary 3.5.** *Let $1/2 < p < +\infty$ and $\mathcal{H}^p[0,1]$ the Sobolev space of all functions $f \in L^2[0,1]$ satisfying*

$$\sum_{n\in\mathbb{Z}} \left(1+n^2\right)^p \left|c_n(f)\right|^2 < +\infty$$

*with inner product*

$$(f, g)_{\mathcal{H}^p[0,1]} := \sum_{n\in\mathbb{Z}} \left(1+n^2\right)^p c_n(f)\overline{c_n(g)}, \quad f, g \in \mathcal{H}^p[0,1].$$

*Then disjoint support implies orthogonality in $\mathcal{H}^p[0,1]$ if and only if $p \in \mathbb{N}$.*

**Proof.** The result follows from Theorem 3.4 and similar arguments as those in the proof of Corollary 3.3. □

By Theorems 3.2 and 3.4, we have obtained a class of $K_c$ and $K_s$ whose RKHS have the orthogonality from disjoint support property. However, the RKHS of the sum of such $K_c$ and $K_s$ does not have the property.

**Proposition 3.6.** *Suppose that $K_c$ and $K_s$ are defined respectively by* (3.3) *and* (3.4)*, where nonnegative $\phi$ satisfies* $\operatorname{supp}\phi = \mathbb{R}$ *and* (3.10)*, singular Borel measure $\nu \in \mathcal{B}_+(\mathbb{R})$ satisfies* $\operatorname{supp}\nu = \mathbb{Z}$ *and* (3.14)*. Then for $K = K_c + K_s$, $\mathcal{H}_K$ does not have the orthogonality from disjoint support property.*

**Proof.** Suppose to the contrary that disjoint support implies orthogonality in $\mathcal{H}_K$. Select an arbitrary nontrivial function $\varphi \in \mathcal{D}(\mathbb{R})$ such that $\operatorname{supp}\varphi \subseteq [0,1]$. By (3.8), $\varphi \in \mathcal{H}_{K_c}$. Define

$$\psi(x) := \varphi(x) - \sum_{n\in\mathbb{Z}} c_n(\varphi)e^{i2\pi nx}, \quad x \in \mathbb{R}.$$

Then by (3.15), $\psi - \varphi \in \mathcal{H}_{K_s}$ and $\operatorname{supp}\varphi \cap \operatorname{supp}\psi = \emptyset$. We write $\varphi, \psi$ in terms of the feature map (3.5) as

$$\varphi(x) = \int_{\mathbb{R}} e^{i2\pi x\xi} \frac{\hat{\varphi}(\xi)}{\phi(\xi)} \phi(\xi)\,d\xi, \quad x \in \mathbb{R},$$

and

$$\psi(x) = \int_{\mathbb{R}} e^{i2\pi x\xi} \frac{\hat{\varphi}(\xi)}{\phi(\xi)} \phi(\xi)\,d\xi - \sum_{n\in\mathbb{Z}} e^{i2\pi nx} \frac{c_n(\varphi)}{\nu_n} \nu_n, \quad x \in \mathbb{R}.$$

Therefore, we get by Lemma 2.1 that $\varphi, \psi \in \mathcal{H}_K$. Since they have disjoint support, we must have

$$\int_{\mathbb{R}} |\hat{\varphi}(\xi)|^2 \frac{1}{\phi(\xi)} d\xi = (\varphi, \psi)_{\mathcal{H}_K} = 0,$$

which yields that $\hat{\varphi} \equiv 0$, contradicting that $\varphi$ is nontrivial. □

Combining Theorems 3.2, 3.4 and the above proposition, we may conclude that there are not many RKHS with a translation invariant reproducing kernel where disjoint support implies orthogonality. We end the paper by remarking that one can obtain similar results for the multidimensional case by carefully dealing with polynomials of several variables.

## References

[1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
[2] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 68 (1950) 337–404.
[3] S. Bochner, Hilbert distances and positive definite functions, Ann. of Math. (2) 42 (1941) 647–656.
[4] V. Bolotnikov, Interpolation for multipliers on reproducing kernel Hilbert spaces, Proc. Amer. Math. Soc. 131 (2003) 1373–1383.
[5] F. Cucker, S. Smale, On the mathematical foundations of learning, Bull. Amer. Math. Soc. 39 (2002) 1–49.
[6] X. Dong, D.X. Zhou, Learning gradients by a gradient descent algorithm, J. Math. Anal. Appl. 341 (2008) 1018–1027.
[7] W.F. Donoghue, Distributions and Fourier Transforms, Academic Press, New York, 1969.
[8] T. Evgeniou, M. Pontil, T. Poggio, Regularization networks and support vector machines, Adv. Comput. Math. 13 (2000) 1–50.
[9] C. Gasquet, P. Witomski, Fourier Analysis and Applications, Springer-Verlag, New York, 1999.
[10] D. Jagerman, Bounds for truncation error of the sampling expansion, SIAM J. Appl. Math. 14 (1966) 714–723.
[11] G. Kimeldorf, G. Wahba, Some results on Tchebycheffian spline functions, J. Math. Anal. Appl. 33 (1971) 82–95.
[12] J. Maltz, R. De Mello Koch, A. Willis, Reproducing kernel Hilbert space method for optimal interpolation of potential field data, IEEE Trans. Image Process. 7 (1998) 1725–1730.
[13] R.J. Marks II, Introduction to Shannon Sampling and Interpolation Theory, Spring-Verlag, New York, 1991.
[14] C.A. Micchelli, M. Pontil, Learning the kernel function via regularization, J. Mach. Learn. Res. 6 (2005) 1099–1125.
[15] C.A. Micchelli, M. Pontil, On learning vector-valued functions, Neural Comput. 17 (2005) 177–204.
[16] C.A. Micchelli, Y. Xu, H. Zhang, Universal kernels, J. Mach. Learn. Res. 7 (2006) 2651–2667.
[17] R. Opfer, Multiscale kernels, Adv. Comput. Math. 25 (2006) 357–380.
[18] L. Qian, On the regularized Whittaker–Kotel'nikov–Shannon sampling formula, Proc. Amer. Math. Soc. 131 (2003) 1169–1176.
[19] W. Rudin, Real and Complex Analysis, third ed., McGraw–Hill, New York, 1987.
[20] S. Saitoh, Representations of inverse functions, Proc. Amer. Math. Soc. 125 (1997) 3633–3639.
[21] B. Schölkopf, A.J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, MIT Press, Cambridge, MA, 2002.
[22] J. Shawe-Taylor, N. Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, Cambridge, 2004.
[23] R. Shonkwiler, On the solution of moment problems by reproducing kernel methods, J. Math. Anal. Appl. 130 (1988) 271–299.
[24] S. Smale, D.X. Zhou, Estimating the approximation error in learning theory, Appl. Anal. 1 (2003) 17–41.
[25] S. Smale, D.X. Zhou, Shannon sampling and function reconstruction from point values, Bull. Amer. Math. Soc. 41 (2004) 279–305.
[26] I. Steinwart, C. Scovel, Fast rates for support vector machines using Gaussian kernels, in: Proc. 18th Annual Conference on Learning Theory (COLT 05), Bertinoro, 2005, pp. 279–294.
[27] V.N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[28] Y. Ying, D.X. Zhou, Learnability of Gaussians with flexible variances, J. Mach. Learn. Res. 8 (2007) 249–276.
[29] Y. Xu, H. Zhang, Refinable kernels, J. Mach. Learn. Res. 8 (2007) 2083–2120.