Constr. Approx. (1999) 15: 175–208

CONSTRUCTIVE APPROXIMATION

Variational Principles and Sobolev-Type Estimates for Generalized Interpolation on a Riemannian Manifold

N. Dyn, F. J. Narcowich, and J. D. Ward

Abstract. The purpose of this paper is to study certain variational principles and Sobolev-type estimates for the approximation order resulting from using strictly positive definite kernels to do generalized Hermite interpolation on a closed (i.e., no boundary), compact, connected, orientable, m-dimensional C^∞ Riemannian manifold $\mathcal{M}$, with C^∞ metric g_{ij}. The rate of approximation can be more fully analyzed with rates of approximation given in terms of Sobolev norms. Estimates on the rate of convergence for generalized Hermite and other distributional interpolants can be obtained in certain circumstances and, finally, the constants appearing in the approximation order inequalities are explicit. Our focus in this paper will be on approximation rates in the cases of the circle, other tori, and the 2-sphere.

0. Introduction

Over the last few years, there has been an increasing interest in approximation methods based on variational techniques in a reproducing kernel Hilbert space setting. Perhaps the first paper on this topic was that of Golomb and Weinberger [10]. More recently, Duchon [5], [6] and Madych and Nelson [15], [16] have extensively investigated approximation properties of functions having nonnegative Fourier transforms or, more generally, conditionally positive definite functions of order n on $\mathbf{R}^s$. Wu and Schaback [30] used kriging methods, which are based on variational techniques, in their investigation of the approximation properties for this class of functions. In this framework, the approximation rates apply to functions in a certain "induced" space.

The object of this paper is to investigate variational principles and Sobolev-type estimates for the approximation order resulting from using strictly positive definite kernels to do generalized Hermite interpolation on a closed (i.e., no boundary), compact, connected, orientable, m-dimensional C^∞ Riemannian manifold $\mathbf{M}^m$, with C^∞ metric g_{ij} [1], [8], [11]. Generalized Hermite interpolation in $\mathbf{R}^s$ was introduced in [20] and [29], and on a manifold in [19]. We should also mention that Xu and Cheney [31] gave conditions for Schoenberg's spherical positive definite functions [24] to be *strictly positive*, thereby providing a wide class of basis functions with which to do interpolation on $\mathbf{S}^m$.

Date received: April 10, 1996. Dates revised: March 26, 1997; August 26, 1997. Date accepted: September 12, 1997. Communicated by Ronald A. DeVore.
AMS classification: Primary: 41A05, 42A82, 42C10; Secondary: 42A10.
Key words and phrases: Approximation order, Positive definite kernels, Sobolev spaces, Riemannian manifolds.

It frequently happens that the induced spaces contain certain Sobolev spaces. The rate of approximation can therefore be analyzed in terms of Sobolev norms. Estimates on the rate of convergence for generalized Hermite and other distributional interpolants can be obtained in certain circumstances, and finally the constants appearing in the approximation order inequalities are explicit. Of special interest in this paper will be the approximation rates obtained in the setting of the 2-sphere. Some work in this direction has already been done [7], [17], [22], [26], [27]. A crucial ingredient for our estimates is a recent sampling theorem for band-limited functions on the 2-sphere given in [4].

The outline for the paper is as follows: In Section 1, relevant notation and certain basic results about Sobolev spaces on Riemannian manifolds are given. In Section 2, materials about complex-valued kernels are presented. Special attention is given to those kernels which are positive definite and strictly positive definite within our framework. Also the abstract interpolation problem is presented. In Section 3, a variational principle is derived which shows that the solution to the abstract interpolation problem satisfies a certain minimal norm criterion which in turn gives rise to the "rate of approximation" estimates. In Section 4, the notion of conditionally positive definite kernels on manifolds is introduced. In Section 5, rate of approximation estimates are derived in the case of the circle and other tori for both equally spaced and scattered points. In Section 6, approximation rates for certain kernels on $\mathbf{S}^2$ are derived.

1. Notation and Background

We will introduce notation and discuss pertinent matters related to manifolds. For the most part, we will use the notation for manifolds and distributions that is used in Friedlander [8, §2.8] and in [19]. In particular, in local coordinates the invariant volume element $d\mu$ induced by the Riemannian metric g_{ij} is

$$d\mu(x) = g(x)^{1/2}\,dx \qquad \text{where} \quad g(x) := \det g_{ij}(x) > 0$$
$$\text{and} \quad dx := dx^1 \cdots dx^m.$$

As usual, we denote the components of the inverse of g_{ij} by g^{ij}. By $C^k(\mathbf{M}^m)$ we will mean the collection of all k-times continuously differentiable functions on $\mathbf{M}^m$; the compactness of the closed manifold $\mathbf{M}^m$ implies that $C^k(\mathbf{M}^m) = C_0^k(\mathbf{M}^m)$, where $C_0^k(\mathbf{M}^m)$ comprises all C^k functions whose support is contained in a compact subset of $\mathbf{M}^m$.

Much of what we do here is related to expansions involving eigenfunctions of the Laplace–Beltrami operator, which in local coordinates has the form [8, p. 11]

$$\Delta w := g^{-1/2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial}{\partial x^i}\left(g^{1/2} g^{ij} \frac{\partial w}{\partial x^j}\right) \qquad \text{where} \quad g = \det g_{ij}.$$

It is possible to show that Δ is self-adjoint relative to $L^2(\mathbf{M}^m, g)$ [1, p. 54] and that it is elliptic [28, p. 250]. The eigenfunctions mentioned above arise in connection with the eigenvalue problem

$$\Delta F + \lambda F = 0, \qquad F \in C^2(\mathbf{M}^m),$$

concerning which we have the following well-known result:

Proposition 1.1. *The eigenvalues are discrete, nonnegative, and form an increasing sequence with $+\infty$ as the only accumulation point. The eigenspaces corresponding to a given eigenvalue are finite-dimensional subspaces of $C^\infty(\mathbf{M}^m)$. The eigenspaces are orthogonal in $L^2(\mathbf{M}^m, g)$, and the direct sum of these spaces is all of $L^2(\mathbf{M}^m, g)$.*

Proof. See [1, Chap. 3] and [28, Chap. 6]. ■

We remark that much of what we do here could also be done with other eigenfunction expansions, provided we replace the Sobolev spaces by weighted eigenfunction expansions.

Occasionally we need to label the eigenvalues. A common choice for the index j comes from labeling in increasing order the eigenvalues λ_j, with appropriate repetitions for degenerate cases. The compactness of $\mathbf{M}^m$ implies $\lambda_1 = 0$ is the lowest eigenvalue; it is easy to show that it has multiplicity 1, so $\lambda_j > 0$ if $j > 1$. Other indexing schemes are, of course, possible. We will let F_j be the eigenfunction corresponding to the eigenvalue λ_j. Of course, we can choose the F_j's so that they form an orthonormal basis for $L^2(\mathbf{M}^m, g)$, and we do so here. We also want to point out that the theorem above implies that each F_j is actually a C^∞ function.

In [19] we discussed in detail the definition of a distribution on a Riemannian manifold, and we refer the reader to that paper for the necessary background material. We shall denote the set of all distributions on the manifold $\mathbf{M}^m$ by $\mathcal{D}'(\mathbf{M}^m)$. To evaluate any distribution, we may use any partition of unity subordinated to a finite covering of $\mathbf{M}^m$ to evaluate (u, φ). Let $\mathcal{P} = \{(\Omega_\nu, \pi_\nu)\}$ be a finite set of coordinate charts that cover $\mathbf{M}^m$, and let $\mathcal{X} = \{\chi_\nu\}$ be a partition of unity subordinated to this covering. We have

$$(u, \varphi) = \sum_\nu \int_{\pi_\nu(\Omega_\nu)} u \circ \pi_\nu^{-1}(x)(\chi_\nu \varphi) \circ \pi_\nu^{-1}(x)\, d\mu(x).$$

This is the precise meaning of the more standard expression

$$(u, \varphi) = \int_{\mathbf{M}^m} u(p)\varphi(p)\, d\mu(p).$$

We will make heavy use of Sobolev spaces [9, Chap. I; 13, §1.7] in our analysis. In terms of the eigenvalue problem mentioned above, these spaces can be characterized in a relatively simple way [13, §1.7.3]. If s is an arbitrary real number, then

$$H_s(\mathbf{M}^m) = \left\{ u \in \mathcal{D}'(\mathbf{M}^m) : \sum_{j=1}^\infty \lambda_j^s |\widehat{u}(j)|^2 < \infty \right\}, \qquad \widehat{u}(j) := (u, \bar{F}_j),$$

where F_j being in $C^\infty(\mathbf{M}^m)$ implies that $(u, \bar{F}_j)$ is well defined. The norm on this space will be taken to be

$$\|u\|_s = \left(\sum_{j=1}^\infty (1 + \lambda_j)^s |\widehat{u}(j)|^2 \right)^{1/2}.$$

Below, we collect some important standard results about Sobolev spaces:

Lemma 1.2. *If s is an arbitrary real number, $H_{-s} = H_s^*$. If $s > k + m/2$, then $H_s(\mathbf{M}^m) \subset C^k(\mathbf{M}^m)$.*

Proof. See Gilkey's book [9, p. 35]. ■

We need to consider tensor products of distributions. Let κ be a C^∞ function $\kappa : \mathbf{M}^m \times \mathbf{M}^m \to \mathbf{C}$, and let u and v be in $\mathcal{D}'(\mathbf{M}^m)$. Using [8, Theorem 2.3.2, p. 42] together with the partitions of unity discussed above, we may define the tensor product of u and v as the distribution $u \otimes v$ that is given via

$$\begin{aligned}(u \otimes v, \kappa) &= \int_{\mathbf{M}^m} u(p) \left\{ \int_{\mathbf{M}^m} v(q)\kappa(p,q)\, d\mu(q) \right\} d\mu(p) \\ &= \int_{\mathbf{M}^m} v(q) \left\{ \int_{\mathbf{M}^m} u(p)\kappa(p,q)\, d\mu(p) \right\} d\mu(q).\end{aligned}$$

It is important to note that the following result holds for the tensor products of distributions:

Proposition 1.3. *Let s and t be positive real numbers. If $u \in H_{-s}(\mathbf{M}^m)$ and $v \in H_{-t}(\mathbf{M}^m)$, then $u \otimes v \in H_{-s-t}(\mathbf{M}^m \times \mathbf{M}^m)$. Moreover,*

$$\|u \otimes v\|_{-s-t} \le \|u\|_{-s} \|v\|_{-t}.$$

Proof. The Laplace–Beltrami operator for the product manifold $\mathbf{M}^m \times \mathbf{M}^m$ has eigenvalues $\lambda_j + \lambda_k$ and corresponding eigenfunctions $F_j \otimes F_k$. Clearly, $(u \otimes v, \overline{F_j \otimes F_k}) = (u, \bar{F}_j)(v, \bar{F}_k)$. Consequently,

$$\begin{aligned}\|u \otimes v\|_{-s-t}^2 &= \sum_{j,k=1}^{\infty} (1 + \lambda_j + \lambda_k)^{-s-t} |(u \otimes v, \overline{F_j \otimes F_k})|^2 \\ &= \sum_{j,k=1}^{\infty} (1 + \lambda_j + \lambda_k)^{-s-t} |(u, \bar{F}_j)(v, \bar{F}_k)|^2 \\ &= \sum_{j,k=1}^{\infty} \frac{(1+\lambda_j)^s}{(1+\lambda_j+\lambda_k)^s} \frac{(1+\lambda_k)^t}{(1+\lambda_j+\lambda_k)^t} \\ &\quad \times (1+\lambda_j)^{-s} |(u, \bar{F}_j)|^2 (1+\lambda_k)^{-t} |(v, \bar{F}_k)|^2 \\ &\le \left(\sum_{j=1}^{\infty} (1+\lambda_j)^{-s} |(u, \bar{F}_j)|^2 \right) \left(\sum_{k=1}^{\infty} (1+\lambda_k)^{-t} |(v, \bar{F}_k)|^2 \right) \\ &= \|u\|_{-s}^2 \|v\|_{-t}^2.\end{aligned}$$

Taking square roots yields the norm-inequality in the lemma. ■

We will close our introductory remarks with a brief discussion of the mapping properties associated with a kernel $\kappa \in H_{2s} \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$. Clearly, we can define the linear transformation $\Psi_\kappa : C^0(\mathbf{M}^m) \to C^0(\mathbf{M}^m)$ via the expression

$$\Psi_\kappa[f](p) = \int_{\mathbf{M}^m} \kappa(p,q) f(q)\, d\mu(q).$$

Using the same kind of proof as in the last proposition, we can obtain the following result:

Proposition 1.4. *If* $0 \le t \le 2s$ *and if* $\kappa \in H_{2s} \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$, *then the linear map* Ψ_κ *defined above continuously maps* H_{-t} *into* H_{2s-t}. *Moreover, we have that*

$$\|\Psi_\kappa\| \le \|\kappa\|_{2s}.$$

Proof. Omitted. ■

In order to emphasize the kernel generating Ψ_κ, we will adopt the convention that

$$\kappa * u := \Psi_\kappa[u].$$

The symbol "$*$" is used to suggest convolution, although in using the symbol we should keep in mind that $*$ is *not* commutative.

Any continuous kernel $\kappa(p, q)$ defined on $\mathbf{M}^m$ has an eigenfunction expansion of the form

$$\kappa(p, q) = \sum_{j,k} \alpha_{j,k} F_k(p) \bar{F}_j(q).$$

Later on, we will work with the subclass of kernels for which $\alpha_{j,k} = a_k \delta_{j,k}$. Specifically, these kernels have the form

$$\kappa(p, q) := \sum_{k=1}^{\infty} a_k F_k(p) \bar{F}_k(q) \qquad \text{where} \qquad \sum_{k=1}^{\infty} (1 + 2\lambda_k)^{2s} |a_k|^2 < \infty. \tag{1.1}$$

The condition on the a_k's is equivalent to assuming that κ is in $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m)$. Kernels of the form (1.1) are analogous to convolution kernels and are similar to those treated in Section 3 of [19]. Concerning such kernels, we have the following:

Proposition 1.5. *Let s be a positive real number. If* $\kappa(p, q)$ *has the form* (1.1), *then* κ *is a self-adjoint kernel,* $\bar{\kappa}(p, q) = \kappa(q, p)$. *In addition, if u and v be in* $H_{-s}(\mathbf{M}^m)$, *then*

$$(u \otimes v, \kappa) = \sum_{k=1}^{\infty} a_k (u, F_k)(v, \bar{F}_k). \tag{1.2}$$

Proof. That κ is self-adjoint is obvious from (1.1). By Proposition 1.3, $u \otimes v$ is in $H_{-2s}(\mathbf{M}^m \times \mathbf{M}^m)$. Since $\kappa \in H_{2s}(\mathbf{M}^m \times \mathbf{M}^m)$ has the expansion (1.1), which is obviously convergent in $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m)$, we can replace κ by (1.1) on the left in (1.2), and then interchange the sum with the continuous linear functional $u \otimes v$ to obtain the series expansion on the left in (1.2). ■

2. Strictly Positive Definite Kernels

A complex-valued kernel $\kappa \in C(\mathbf{M}^m \times \mathbf{M}^m)$ is termed *positive definite on* $\mathbf{M}^m$ if $\bar{\kappa}(q, p) = \kappa(p, q)$ and if for every finite set of points $\mathcal{C} = \{p_1, \ldots, p_n\}$ in $\mathbf{M}^m$, the self-adjoint, $N \times N$ matrix with entries $\kappa(p_j, p_k)$ is positive semidefinite. The case in which κ is C^∞ was studied in detail in [19]. Positive definite kernels have arisen in many contexts [25]. The definition that we use here is motivated by Schoenberg's [24], and is related to the ones studied by M. G. Krein [12], Yu. M. Berezanskii [2], and K. Maurin [14]. We will be especially interested in H_{2s} positive definite kernels, for which we have the following result, which is similar to one for the C^∞ case [19, Theorem 2.1]:

Theorem 2.1. *Let the kernel κ be in $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m) \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$, and let κ satisfy $\bar{\kappa}(q, p) = \kappa(p, q)$. Then, κ is positive definite if and only if for every $u \in H_{-s}(\mathbf{M}^m)$,*

$$(\bar{u} \otimes u, \kappa) \geq 0. \tag{2.1}$$

Proof. The proof is similar to that given in [19, Theorem 2.1], and we will omit it. ■

Later on, we will construct a number of examples of positive definite kernels, but first we need to define precisely the generalized Hermite interpolation problem that we wish to study. Let $\{u_j\}_{j=1}^N \subset H_{-s}$ be a linearly independent set of distributions on $\mathbf{M}^m$, and let $f \in H_s$. We may think of f as being "measured" and producing data in the form:

$$\int_{\mathbf{M}^m} \bar{u}_j(p) f(p)\, d\mu(p) = d_j \qquad \text{for} \quad j = 1, \ldots, N$$

where the d_j's are complex numbers. The condition that the set $\{u_j\}_{j=1}^N$ be linearly independent is equivalent to there being no redundancy in the data.

Problem 2.2. *Given a linearly independent set $\{u_j\}_{j=1}^N \subset H_{-s}(\mathbf{M}^m)$, complex numbers d_j, $j = 1, \ldots, N$, and an $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m) \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$ positive definite kernel κ on $\mathbf{M}^m$, find $u \in \mathrm{Span}\{u_1, \ldots, u_N\}$ such that*

$$\int_{\mathbf{M}^m} \bar{u}_j(p)(\kappa * u)(p)\, d\mu(p) = d_j \qquad \textit{for} \quad j = 1, \ldots, N.$$

Equivalently, find $u \in \mathrm{Span}\{u_1, \ldots, u_N\}$ such that the function $\kappa * u$ is a $H_s(\mathbf{M}^m)$ interpolant for the data generated by the function f above.

The problem itself encompasses the problem of scattered data interpolation on a manifold; it also includes almost any problem involving data generated by derivatives, differences, and even continuous distributions. We will say that such a problem is *well-poised* if there exists a solution and the solution is unique. The next theorem characterizes the types of positive definite kernels κ on $\mathbf{M}^m$ relative to which generalized Hermite problems are well-poised. Before we can state it, we need a definition: Let κ be a $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$ positive definite kernel on $\mathbf{M}^m$. If the equation $(\bar{u} \otimes u, \kappa) = 0$ implies that the distribution $u = 0$, then we will say that κ is *strictly* positive definite (SPD) on $\mathbf{M}^m$. The results given below have proofs nearly identical to the corresponding results [19, Theorem 2.4 and Corollary 2.5]; they will be stated without proof.

Theorem 2.3. *Problem* 2.2 *is well-poised if κ is* strictly *positive definite on* $\mathbf{M}^m$.

Corollary 2.4. *If κ is a strictly positive definite kernel on $\mathbf{M}^m$, and if the set of distributions $\{u_j\}_{j=1}^N$ is linearly independent, then the interpolation matrix A with entries given by*

$$A_{j,k} := \int_{\mathbf{M}^m} \bar{u}_j(p)(\kappa * u_k)(p)\, d\mu(p) \qquad \textit{where} \quad j, k = 1, \ldots, N, \tag{2.2}$$

is self-adjoint and positive definite.

Theorem 2.3 establishes that Problem 2.2 will be well-poised *provided* we have a strictly positive definite kernel κ defined on $\mathbf{M}^m$, and Corollary 2.4 gives us that the corresponding interpolation matrices will be positive definite. The next result gives us a large class of strictly positive definite kernels.

Corollary 2.5. *Let s be a positive real number. If $\kappa(p,q)$ has the form* (1.1)*, then κ is strictly positive definite if and only if $a_k > 0$ for all $k \geq 1$.*

Proof. By Proposition 1.5, we have that $\bar{\kappa}(p,q) = \kappa(q,p)$, and that, if $u \in H_{-s}(\mathbf{M}^m)$,

$$\begin{aligned}(\bar{u} \otimes u, \kappa) &= \sum_{k=1}^{\infty} a_k(\bar{u}, F_k)(u, \bar{F}_k) \\ &= \sum_{k=1}^{\infty} a_k |(u, \bar{F}_k)|^2 \\ &= \sum_{k=1}^{\infty} a_k |\widehat{u}(k)|^2,\end{aligned} \tag{2.3}$$

where $\widehat{u}(k) = (u, \bar{F}_k)$ is the kth coefficient in the eigenfunction expansion for u. Since $a_k \geq 0$, the expansion (2.3) clearly establishes that κ is a positive definite kernel. Moreover, choosing $u = F_j$ in (2.3) gives $(\bar{u} \otimes u, \kappa) = a_j$. Thus, if κ *is* strictly positive definite, we see that all the a_k's are positive.

We want to show that the converse is also true. Assume that $a_k > 0$ for all k. If $(\bar{u} \otimes u, \kappa) = 0$ for some $u \in H_{-s}(\mathbf{M}^m)$, then since $a_k > 0$ we have that $\widehat{u}(k) = (u, \bar{F}_k) = 0$ for all k, from which it follows that $u = 0$. Hence, κ is strictly positive. ■

Many manifolds of interest—tori of dimension 2 or greater, for example— are products of compact, connected Riemannian manifolds. Choosing the Riemannian metric on $\mathbf{M}^m := \mathbf{M}_1^m \times \cdots \times \mathbf{M}_d^m$ to be the one induced by the Riemannian metrics on each of the $\mathbf{M}_\gamma^m$'s, we get a Laplace–Beltrami operator that is the sum of those coming from the $\mathbf{M}_\gamma^m$'s. For such an operator, we easily see that the eigenvalues and eigenfunctions have, respectively, the forms

$$\lambda_\alpha = \lambda_{j_1}^1 + \cdots + \lambda_{\alpha_d}^d \qquad \text{and} \qquad F_\alpha(p_1, \ldots, p_d) = F_{\alpha_1}^1(p_1) \cdots F_{\alpha_d}^d(p_d), \tag{2.4}$$

where $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index. The eigenfunctions being products allows us to obtain strictly positive definite kernels on $\mathbf{M}_1^m \times \cdots \times \mathbf{M}_d^m$ from products of kernels that are strictly positive definite on each $\mathbf{M}_\gamma^m$ and that have the form (1.1).

Corollary 2.6. *If the factors in the product*

$$\kappa(p_1, \ldots, p_d, q_1, \ldots, q_d) := \kappa_1(p_1, q_1) \cdots \kappa_d(p_d, q_d) \tag{2.5}$$

are strictly positive definite kernels in $H_{2s}(\mathbf{M}_\gamma^m \times \mathbf{M}_\gamma^m)$ and have the form (1.1)*, then the kernel κ defined by* (2.5) *is in $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m)$, has the form* (1.1)*, and is strictly positive definite on the product manifold $\mathbf{M}_1^m \times \cdots \times \mathbf{M}_d^m$. Moreover,*

$$\|\kappa\|_{2s} \leq \prod_{\gamma=1}^{d} \|\kappa_\gamma\|_{2s}. \tag{2.6}$$

Proof. The proof that κ has an expansion of the form (1.1) with positive coefficients is identical to the one for [19, Corollary 3.10] in the C^∞ case. If

$$\kappa_\gamma = \sum_{k=1}^{\infty} a_k^\gamma F_k(p_\gamma)\bar{F}_k(q_\gamma),$$

then the coefficient multiplying the term $F_\alpha(p_1, \dots, p_d)\bar{F}_\alpha(q_1, \dots, q_d)$ is

$$a_\alpha = \prod_{\gamma=1}^{d} a^\gamma_{\alpha_\gamma}. \tag{2.7}$$

From the inequality $(1 + 2\lambda^1_{\alpha_1} + \cdots + 2\lambda^d_{\alpha_d})^{2s} \le \prod_{\gamma=1}^d (1 + 2\lambda^\gamma_{\alpha_\gamma})^{2s}$, it follows that

$$\begin{aligned}\|\kappa\|^2_{2s} &= \sum_\alpha (1 + 2\lambda^1_{\alpha_1} + \cdots + 2\lambda^d_{\alpha_d})^{2s}(a_\alpha)^2 \\ &\le \prod_{\gamma=1}^{d}\left(\sum_{\alpha_\gamma=1}^{\infty}(1+2\lambda^\gamma_{\alpha_\gamma})^{2s}(a^\gamma_{\alpha_\gamma})^2\right) \\ &= \prod_{\gamma=1}^{d}\|\kappa_\gamma\|^2_{2s}.\end{aligned}$$

Taking square roots in the last inequality establishes (2.6) and completes the proof. ∎

3. A Variational Principle

Let κ be a strictly positive definite kernel in $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m)$. We can use κ to define an inner product on the space $H_{-s}(\mathbf{M}^m)$. If u and v are in $H_{-s}(\mathbf{M}^m)$, then $\bar{v} \otimes u$ is in $H_{-2s}(\mathbf{M}^m \times \mathbf{M}^m)$. Thus, the form

$$\lfloor u, v \rfloor := (\bar{v} \otimes u, \kappa) \tag{3.1}$$

is well defined, and by Proposition 1.3 satisfies the inequality

$$|\lfloor u, v \rfloor| \le \|u\|_{-s}\|v\|_{-s}\|\kappa\|_{2s}. \tag{3.2}$$

The self-adjointness and strict positivity of κ imply that (3.1) is an inner product on $H_{-s}(\mathbf{M}^m)$. The norm associated with this inner product is

$$[\!]u[\!] := \sqrt{(\bar{u} \otimes u, \kappa)}. \tag{3.3}$$

Of course, the inner product space that we obtain this way is *not* a Hilbert space. We could complete it to be one, but the space that we would obtain would be too large to be useful, at least in the $H_\infty = C^\infty$ case. Finally, by means of the inner product defined in (3.1), we can recast Problem 2.2.

Problem 2.2′. *Given a linearly independent set* $\{u_j\}_{j=1}^N \subset H_{-s}(\mathbf{M}^m)$, *complex numbers* d_j, $j = 1, \dots, N$, *and a positive definite kernel* $\kappa \in H_{2s}(\mathbf{M}^m) \cap C^0(\mathbf{M}^m)$, *find* $u \in \mathcal{U} := \mathrm{Span}\{u_1, \dots, u_N\}$ *such that* $\lfloor u, u_j \rfloor = d_j$ *for* $j = 1, \dots, N$.

Since we are assuming that κ is strictly positive definite, we know by Theorem 2.3 that this problem is well-poised. The unique solution to it is $u = \sum_{j=1}^{N} c_j u_j$, where $c = (c_1 \cdots c_N)^T \in \mathbf{R}^N$ is the unique solution to $Ac = d$. Here, $d = (d_1 \cdots d_N)^T$. The interpolant we arrive at in this way is $\kappa * u$. The distribution u satisfies the following minimization principle:

Theorem 3.1. *Let $u \in H_{-s}(\mathbf{M}^m)$ be the unique distribution in $\mathcal{U}$ for which $\kappa * u$ solves Problem* 2.2′. *If $v \in H_{-s}(\mathbf{M}^m)$ is any other distribution for which $\kappa * v$ also interpolates the data in Problem* 2.2′*—that is, $[\![v, u_j]\!] = d_j$ for $j = 1, \ldots, N$—then $v - u$ is orthogonal to* $\mathrm{Span}\{u_1, \ldots, u_N\}$ *with respect to the inner product* (3.1). *In addition,*

$$\|v\|^2 = \|v - u\|^2 + \|u\|^2.$$

Finally, if $v \neq u$,

$$\|u\| < \|v\|.$$

Proof. Since both u and v interpolate the data, we have that $[\![v - u, u_j]\!] = 0$ for $j = 1 \ldots N$; thus, $v - u$ is orthogonal to $\mathcal{U}$. The rest of the theorem follows from standard linear algebra. ∎

We remark that this is quite similar to the variational principle derived by Madych and Nelson [15], [16] in the radial basis function (RBF) case. (For a review of RBFs, see [21].) The last inequality in the theorem amounts to saying that among all $v \in H_{-s}(\mathbf{M}^m)$ for which $\kappa * v$ interpolates the data, the distribution u minimizes the norm (3.3). We now turn to the estimates that we mentioned earlier.

We again suppose that u and v are as above, that the functions

$$f := \kappa * v \qquad \text{and} \qquad \tilde{f} := \kappa * u \tag{3.4}$$

both interpolate the data, and that $u \in \mathcal{U}$ is the unique solution to Problem 2.2′. If we regard f as having generated the data in the first place, we can ask how well the interpolant $\tilde{f}$ approximates the original function f.

Proposition 3.2. *Let f and $\tilde{f}$ be as in* (3.4), *and let w be an arbitrary distribution in $H_{-s}(\mathbf{M}^m)$. If g and g_w are defined by*

$$g := f - \tilde{f} = \kappa * (v - u) \qquad \textit{and} \qquad g_w := (\bar{w}, g), \tag{3.5}$$

then

$$|g_w| \le \mathrm{dist}(v, \mathcal{U})\, \mathrm{dist}(w, \mathcal{U}), \tag{3.6}$$

where the distances are computed relative to the norm (3.3).

Proof. As we noted earlier, $v - u$ is orthogonal to $\mathcal{U} = \mathrm{Span}\{u_1, \ldots, u_N\}$. Thus, if $c_1, \ldots, c_N$ are arbitrary complex numbers, then

$$g_w := \int_{\mathbf{M}^m} \bar{w}(p) g(p)\, d\mu(p) \tag{3.7}$$

$$
\begin{aligned}
&= \llbracket v-u, w \rrbracket \\
&= \left\llbracket v-u, w-\sum_{j=1}^{N} c_j u_j \right\rrbracket .
\end{aligned}
$$

Applying Schwarz's inequality to (3.7) yields

$$
|g_w| \le \llbracket v-u \rrbracket \left\llbracket w-\sum_{j=1}^{N} c_j u_j \right\rrbracket . \tag{3.8}
$$

Since $v-u$ is orthogonal to $\mathcal{U}$, u is the orthogonal projection of v onto $\mathcal{U}$. Hence,

$$
\llbracket v-u \rrbracket = \operatorname{dist}(v,\mathcal{U}). \tag{3.9}
$$

Also, we have that

$$
\operatorname{dist}(v,\mathcal{U}) = \min_{c_1,\ldots,c_N} \left\llbracket w-\sum_{j=1}^{N} c_j u_j \right\rrbracket . \tag{3.10}
$$

Coupling these facts with (3.8) yields (3.6). ∎

Remark 3.3. Formula (3.6) is a version of the hypercircle inequality [3, p. 230].

The utility in (3.6) is that, because the right side separates into a product of a term depending only on g and a term involving only w, we can estimate the effect of w independently of the original g. Another feature is that the distributions w and v appear symmetrically on the right side of (3.6), in factors that involve only distances to $\mathcal{U}$. Thus our strategy for obtaining estimates on g_w is to get upper bounds on the distance from a distribution to $\mathcal{U}$, given that the norm we employ is (3.3).

Remark 3.4. The distance $\operatorname{dist}(v,\mathcal{U}) = \llbracket v-u \rrbracket$ depends on the interpolant $\kappa * u$. We can, however, bound $\operatorname{dist}(v,\mathcal{U})$ independently of u. From Theorem 3.1, we see that

$$
\operatorname{dist}(v,\mathcal{U}) = \llbracket v-u \rrbracket \le \llbracket v \rrbracket, \tag{3.11}
$$

which is independent of $\kappa * u$. ∎

Up to now in this section we have treated general SPD kernels. For the rest of it, the types of kernels that we will work with are ones having the form (1.1), with $a_k > 0$ for all k. By Corollary 2.5, these are strictly positive definite kernels. Moreover, this class includes a wide variety of C^k kernels, including the C^∞ kernels treated in Section 3 of [19].

If we impose additonal restictions on the rate of decay of a_k's, then for these kernels we can obtain estimates on $\operatorname{dist}(w,\mathcal{U})$ and $\llbracket v \rrbracket$ in terms of more standard norms. We have the following result:

Proposition 3.5. *Let κ and s be as in* (1.1). *If there is a $t > s$ for which the a_k's satisfy a bound of the form*

$$
a_k \ge c^{-1}(1+\lambda_k)^{-(s+t)/2} \tag{3.12}
$$

for all k and some constant c, then every $f \in H_t(\mathbf{M}^m)$ can be written as $f = \kappa * v$ for some v in $H_{-s}(\mathbf{M}^m)$. Moreover,

$$\|v\|_{-s} \le c\|f\|_t \qquad \text{and} \qquad [\![v]\!] \le \|\kappa\|_{2s}^{1/2} c\|f\|_t. \tag{3.13}$$

Proof. Use the eigenfunctions F_k to expand f,

$$f(p) = \sum_{k=1}^{\infty} \widehat{f}(k) F_k(p).$$

Let $\widehat{v}(k) = \widehat{f}(k)/a_k$, and observe that for all k,

$$(1+\lambda_k)^{-s}|\widehat{v}(k)|^2 \le c^2(1+\lambda_k)^t|\widehat{f}(k)|^2.$$

Summing both sides over all $k \ge 0$ yields

$$\sum_{k=1}^{\infty}(1+\lambda_k)^{-s}|\widehat{v}(k)|^2 \le c^2\|f\|_t^2.$$

From this, it follows that

$$v(p) := \sum_{k=1}^{\infty} \widehat{v}(k) F_k(p)$$

defines a distribution in $H_{-s}(\mathbf{M}^m)$ that satisfies $\|v\|_{-s} \le c\|f\|_t$. This inequality when coupled with (3.2) yields the inequality $[\![v]\!] \le \|\kappa\|_{2s}^{1/2} c\|f\|_t$. ■

Our next result sets up a framework that will enable us to discuss how well our interpolants approximate a given function. The idea is to estimate the distance $\mathrm{dist}(w, \mathcal{U})$, given that the kernels used to generate the inner product on $\mathcal{U}$ are of the form (1.1).

Proposition 3.6. *Let M be a positive integer, let $s > 0$, and let κ have the form* (1.1) *with $a_k > 0$ for all $k \ge 1$. If there are coefficients $c_1, \ldots, c_N$ such that, for $k = 1 \ldots M$, $(w - \sum_{j=1}^N c_j u_j, \bar{F}_k) = 0$, and if there is a sequence $b_k > 0$, $k = M+1, M+2, \ldots$, for which $|(w - \sum_{j=1}^N c_j u_j, \bar{F}_k)|^2 \le b_k$ when $k \ge M+1$, then*

$$\mathrm{dist}(w, \mathcal{U}) \le \left[\!\!\left[w - \sum_{j=1}^{N} c_j u_j \right]\!\!\right] \le \left(\sum_{k=M+1}^{\infty} a_k b_k \right)^{1/2}. \tag{3.14}$$

Proof. From (3.3) and (2.3), we have that

$$\left[\!\!\left[w - \sum_{j=1}^{N} c_j u_j \right]\!\!\right]^2 = \sum_{k=1}^{\infty} a_k \left| \left(w - \sum_{j=1}^{N} c_j u_j, \bar{F}_j \right) \right|^2, \tag{3.15}$$

from which (3.14) follows immediately. ■

Remark 3.7. It is not necessary to choose the ordering of the eigenfunctions to be that obtained from the labeling of the eigenvalues in increasing order. In fact, the ordering is arbitrary, and may be chosen to be whatever is convenient.

We can combine the results above to obtain a variety of estimates on how well the interpolant $\kappa * u$ fits the original function $\kappa * v$. Perhaps the most useful case is when the function generating the data belongs to one of the Sobolev spaces. This is the situation that we address below.

Theorem 3.8. *Let κ and s be as in* (1.1), *and let $t > s$ satisfy the conditions in Proposition* 3.5. *If f is in $H_t(\mathbf{M}^m)$, if w is a distribution in H_{-s}, and if $\tilde{f}$ is the solution to Problem* 2.2′ *for data generated by f, then*

$$|(\bar{w},\ f - \tilde{f})| \le \|\kappa\|_{2s}^{1/2} c \|f\|_t \operatorname{dist}(w, \mathcal{U}). \tag{3.16}$$

In addition, if w and the sequence $\{b_k\}_{k=M+1}^{\infty}$ satisfy the conditions of Proposition 3.6, *then*

$$|(\bar{w},\ f - \tilde{f})| \le \|\kappa\|_{2s}^{1/2} c \|f\|_t \left(\sum_{k=M+1}^{\infty} a_k b_k \right)^{1/2}. \tag{3.17}$$

Proof. From Proposition 3.5, we have that $f = \kappa * v$ for some v in $H_{-s}(\mathbf{M}^m)$. Since $\tilde{f} = \kappa * u$, we also have $g = \kappa * (v - u)$, so $(\bar{w},\ f - \tilde{f}) = (\bar{w}, g) = g_w$. Thus, from Proposition 3.2, Remark 3.4, and (3.13), we obtain (3.16). Combining (3.16) with Proposition 3.6 then yields (3.17). ■

There are a number of other estimates we can obtain using the results of this section. The ones given in Theorem 3.8 should be considered representative.

We now turn to the case in which $\mathbf{M}^m$ is the product manifold $\mathbf{M}_1^m \times \cdots \times \mathbf{M}_d^m$, and the kernels are of the form (2.5). Consider the space of distributions

$$\mathcal{U} := \mathcal{U}^1 \otimes \cdots \otimes \mathcal{U}^d, \tag{3.18}$$

where each space $\mathcal{U}^\gamma$ is a finite-dimensional subset of $H_{-s_\gamma}(\mathbf{M}_\gamma^m)$ and where $s_1 + \cdots + s_d = s$.

Standard formulas involving tensor products apply to the inner product induced by the kernel (2.5). In particular, we have

$$\llbracket w^1 \otimes w^2 \otimes \cdots \otimes w^d, z^1 \otimes z^2 \otimes \cdots \otimes z^d \rrbracket = \prod_{\gamma=1}^{d} \llbracket w^\gamma, z^\gamma \rrbracket_\gamma, \tag{3.19}$$

where w^γ, z^γ are in $H_{-s_\gamma}(\mathbf{M}_\gamma^m)$ and the inner products are of the form (3.1), with the kernel being κ_γ.

Proposition 3.9. *If $w := w^1 \otimes \cdots \otimes w^d$, where $w^\gamma \ne 0$ is in $H_{-s_\gamma}(\mathbf{M}_\gamma^m)$ for $\gamma = 1, \ldots, d$ and if the space $\mathcal{U}$ is given by* (3.18), *then we have*

$$\operatorname{dist}(w, \mathcal{U}) = \llbracket w \rrbracket \sqrt{1 - \prod_{\gamma=1}^{d} \left(1 - \operatorname{dist}(w^\gamma, \mathcal{U}^\gamma)^2 / \llbracket w^\gamma \rrbracket_\gamma^2\right)} \tag{3.20}$$

and in addition

$$\operatorname{dist}(w, \mathcal{U}) \le \llbracket w \rrbracket \left(\sum_{\gamma=1}^{d} \frac{\operatorname{dist}(w^\gamma, \mathcal{U}^\gamma)^2}{\llbracket w^\gamma \rrbracket_\gamma^2} \right)^{1/2}. \tag{3.21}$$

Proof. Let $\tilde{u}^\gamma$ be the orthogonal projection of w^γ onto $\mathcal{U}^\gamma$ relative to the inner product $[\![\cdot,\cdot]\!]_\gamma$, and set $\tilde{u} = \tilde{u}^1 \otimes \cdots \otimes \tilde{u}^d$. Note that, from (3.19) and the well-known properties of orthogonal projections, we can easily show that $[\![w - \tilde{u}, \tilde{u}]\!] = 0$; consequently, $\tilde{u}$ is the orthogonal projection of w onto the subspace $\mathcal{U}$. This and the fact that $\tilde{u}^\gamma$ is the orthogonal projection of w^γ onto $\mathcal{U}^\gamma$ imply that

$$\begin{aligned}
\operatorname{dist}(w,\mathcal{U})^2 &= [\![w]\!]^2 - [\![\tilde{u}]\!]^2 \\
&= [\![w]\!]^2 \left(1 - \prod_{\gamma=1}^{d} [\![\tilde{u}^\gamma]\!]_\gamma^2 / [\![w^\gamma]\!]^2\right) \\
&= [\![w]\!]^2 \left(1 - \prod_{\gamma=1}^{d} (1 - \operatorname{dist}(w^\gamma,\mathcal{U}^\gamma)^2/[\![w^\gamma]\!]_\gamma^2)\right).
\end{aligned}$$

Equation (3.20) then follows from taking square roots in the previous equation. To establish (3.21), use induction on d to show that the inequality $\prod_{\gamma=1}^{d}(1 - x_\gamma^2) \geq 1 - \sum_{\gamma=1}^{d} x_\gamma^2$ holds for all x_γ's that satisfy $0 \leq x_\gamma < 1$. Apply this to the right side of (3.20) to get (3.21). ■

The quality of an approximation method is measured by the error estimates, by its stability, and by the amount of computation it requires. It is known for methods based on radial functions that better rates of approximation are coupled with worse stability constants [23]. Numerical results indicate a similar trade-off in the cases where we deal with the torus or the sphere, and it should be possible to quantify such a trade-off in the situation described here.

We will discuss specific examples in Sections 5 and 6. We now turn to dealing with kernels that are analogous to RBFs of order 1 or higher.

4. Conditionally Positive Definite Kernels on Manifolds

In practical applications we often want interpolants that reproduce exactly some fixed, finite-dimensional space of functions. For example, in Euclidean spaces we would like to use interpolants that reproduce exactly the space of polynomials of degree $n-1$ or less. On the circle, the functions to be reproduced exactly would be all trigonometric polynomials of degree less than n, and on the sphere they would be all spherical harmonics, again with degree less than some fixed number. Trigonometric polynomials are linear combinations of $e^{ik\theta}$'s and spherical harmonics are linear combinations of $Y_{\ell,m}$'s (see Section 6); the $e^{ik\theta}$'s and the $Y_{j,k}$'s are eigenfunctions of the Laplace–Beltrami operators for the circle and the sphere, respectively. It is thus natural to consider interpolation problems on $\mathbf{M}^m$ in which we want interpolants that reproduce exactly a space of functions with a basis comprising finitely many eigenfunctions of the Laplace–Beltrami operator.

To formulate and solve such problems, we will introduce a class of *conditionally* positive definite kernels. Let $\mathcal{I}$ be a finite set of indices in $\mathbf{Z}_+$, and consider the finite-dimensional space defined by $\mathcal{S}_\mathcal{I} := \operatorname{Span}\{F_i : i \in \mathcal{I}\}$. In addition, define the space of distributions $\mathcal{S}_\mathcal{I}^\perp := \{u \in H_{-s}(\mathbf{M}^m) : (u, \bar{F}) = 0,\ F \in \mathcal{S}_\mathcal{I}\}$. We will say that a self-adjoint kernel κ in $H_{2s}(\mathbf{M}^m \times \mathbf{M}^m) \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$ is *conditionally* positive definite

(CPD) with respect to a finite set of indices $\mathcal{I} \subset \mathbf{Z}_+$ if for every $u \in \mathcal{S}_{\mathcal{I}}^{\perp}$ we have

$$(\bar{u} \otimes u, \bar{\kappa}) \geq 0.$$

If the inequality above is strict for $u \neq 0$, then we will say that κ is *conditionally* strictly positive definite (CSPD) on $\mathbf{M}^m$.

Since the proofs of the elementary properties of the CPD and CSPD kernels are similar to those of the corresponding properties for PD and SPD kernels, in the propositions below we will list them without proofs.

Proposition 4.1. *If κ is a conditionally strictly positive definite kernel on $\mathbf{M}^m$, and if the set of distributions $\{v_j\}_{j=1}^N \subset \mathcal{S}_{\mathcal{I}}^{\perp}$ is linearly independent, then the interpolation matrix A having entries $A_{j,k} := (\bar{v}_j, \kappa * v_k)$, where $j, k = 1, \ldots, N$, is self-adjoint and positive definite.*

We remark that, as in the case of SPD kernels, we have a whole class of CSPD kernels arising from self-adjoint kernels of the form (1.1).

Proposition 4.2. *Let $\kappa \in H_{2s}(\mathbf{M}^m \times \mathbf{M}^m) \cap C^0(\mathbf{M}^m \times \mathbf{M}^m)$ be a self-adjoint kernel of the form* (1.1). *κ is a CPD kernel if and only if $a_k \geq 0$ for all $k \notin \mathcal{I}$. In addition, κ is a CSPD kernel if and only if $a_k > 0$ for all $k \notin \mathcal{I}$.*

We will refer to such kernels at the end of the section. For the present, we do not need to require that κ be of the form (1.1).

Proposition 4.3. *Let κ be a conditionally strictly positive definite kernel on $\mathbf{M}^m$. If u and v are in $H_{-s}(\mathbf{M}^m)$, and if $u_{\mathcal{I}} = u - \sum_{i \in \mathcal{I}}(u, \bar{F}_i)F_i$ and $v_{\mathcal{I}} = v - \sum_{i \in \mathcal{I}}(v, \bar{F}_i)F_i$, then the Hermitian form*

$$[\![u, v]\!]_{\mathcal{I}} := (\bar{v}_{\mathcal{I}} \otimes u_{\mathcal{I}}, \kappa), \tag{4.1}$$

defines a semi-inner product on $H_{-s}(\mathbf{M}^m)$, and an inner product on $\mathcal{S}_{\mathcal{I}}^{\perp}$.

We will denote the associated seminorm (or norm, if we are on $\mathcal{S}_{\mathcal{I}}^{\perp}$) by

$$[\![u]\!]_{\mathcal{I}} := [\![u, v]\!]_{\mathcal{I}}^{1/2}. \tag{4.2}$$

We now want to consider the following problem, which is analogous to Problem 2.2:

Problem 4.4. *Given a linearly independent set $\{u_j\}_{j=1}^N \subset H_{-s}(\mathbf{M}^m)$, complex numbers $\{d_j\}_{j=1}^N$, and a CPD kernel κ in $H_{2s}(\mathbf{M}^m) \cap C^0(\mathbf{M}^m)$, find $u \in \mathcal{U}_{\mathcal{I}} := \mathrm{Span}\{u_j : j = 1, \ldots, N\} \cap \mathcal{S}_{\mathcal{I}}^{\perp}$ and $F \in \mathcal{S}_{\mathcal{I}}$, such that*

$$(\bar{u}_j, \kappa * u + F) = d_j \qquad \textit{for} \quad j = 1, \ldots, N;$$

that is, the function $f := \kappa * u + F$ interpolates the data.

Concerning this problem, we have the result below.

Theorem 4.5. *Let κ be CSPD in $H_{2s}(\mathbf{M}^m) \cap C^0(\mathbf{M}^m)$, and let the matrix C with entries $C_{j,i} = (\bar{u}_j, F_i)$, where $j \in \{1, \dots, N\}$ and $i \in \mathcal{I}$, be of full rank. If* $\operatorname{rank}(C) = N$, *then Problem* 4.4 *has one or more solutions, all of the form $f = F \in \mathcal{S}_{\mathcal{I}}$. If the* $\operatorname{rank}(C) = |\mathcal{I}| < N$, *then Problem* 4.4 *has a unique solution*; *moreover, when the data set $\{d_j\}_{j=1}^N$ is generated by a function in $\mathcal{S}_{\mathcal{I}}$, then this solution reproduces the function*.

Proof. Problem 4.4 can be reformulated in terms of a system of equations for the unknowns $b \in \mathbf{C}^N$ and $e \in \mathbf{C}^{|\mathcal{I}|}$ for which $u = \sum_{j=1}^N b_j u_j$ and $F = \sum_{i \in \mathcal{I}} e_i F_i$. The system has the form

$$Bb + Ce = d, \qquad C^* b = 0, \tag{4.3}$$

where B is the matrix with elements

$$B_{j,k} := (\bar{u}_j, \kappa * u_k), \qquad j, k = 1, \dots, N. \tag{4.4}$$

The condition that $C^* b = 0$ comes from requiring that u be in $\mathcal{S}_{\mathcal{I}}^{\perp}$.

In the case $N \le |\mathcal{I}|$, we have $\operatorname{rank}(C) = N$; this implies both that the only vector b satisfying the second condition in (4.3) is $b = 0$ and that there is a vector $e \in \mathbf{C}^{|\mathcal{I}|}$ satisfying $Ce = d$. If $N < |\mathcal{I}|$, then C has a nontrivial null space, and e will not be unique. In that case, there are infinitely many solutions to Problem 4.4, all of the form $f = F = \sum_{i \in \mathcal{I}} e_i F_i$.

For $\operatorname{rank}(C) = |\mathcal{I}| < N$, the matrices C^* and C have nullities $N - |\mathcal{I}|$ and 0, respectively. We want to show that the rank of the system in (4.3) is $N + |\mathcal{I}|$, from which it follows that the (4.3) can be uniquely solved for b and e.

To find the nullity of this system, let $d = 0$ in (4.3), and observe that $C^* b = 0$ implies that $b^* Ce = 0$; hence, $b^* Bb = 0$. On the other hand, $b^* Bb = (\bar{u} \otimes u, \kappa)$, with $u = \sum_j b_j u_j$. Taken together, these yield $(\bar{u} \otimes u, \kappa) = 0$. Since κ is a CSPD kernel, $u = 0$. The linear independence of the u_j's then also gives us that $b = 0$. From this and (4.3), we then get $Ce = 0$, and because the nullity of C is 0, $e = 0$. Thus the nullity of (4.3) is 0.

The statement concerning the reproduction of functions in $\mathcal{S}_{\mathcal{I}}$ follows from the existence and uniqueness result just proved. ∎

The solution of the generalized Hermite problem satisfies a variational principle relative to the seminorm (4.2). Since the proof is straightforward, we omit it.

Theorem 4.6. *Let κ be a CSPD kernel in $H_{2s}(\mathbf{M}^m) \cap C^0(\mathbf{M}^m)$, let* $\operatorname{rank}(C) = |\mathcal{I}| < N$, *and let $f = \kappa * u + F$ be the unique solution to Problem* 4.4. *If*

$$\tilde{f} = \kappa * \tilde{u} + \tilde{F}, \qquad \tilde{u} \in \mathcal{S}_{\mathcal{I}}^{\perp} \quad \text{and} \quad \tilde{F} \in \mathcal{S}_{\mathcal{I}}, \tag{4.5}$$

also satisfies $(\bar{u}_j, \tilde{f}) = d_j$, $j = 1, \dots, N$, then $\tilde{u} - u$ is orthogonal to the subspace $\mathcal{U}_{\mathcal{I}}$, relative to the inner product in (4.1). *Moreover*

$$[\tilde{u}]_{\mathcal{I}}^2 = [\tilde{u} - u]_{\mathcal{I}}^2 + [u]_{\mathcal{I}}^2, \tag{4.6}$$

from which we conclude that the solution to Problem 4.4 *minimizes the seminorm* (4.2) *among all other functions that are of the form* (4.5) *and satisfy the interpolation conditions*.

We want to obtain an error estimate for distributions in $H_{-s}(\mathbf{M}^m)$ applied to the difference $\tilde{f} - f$. That is, if $w \in H_{-s}(\mathbf{M}^m)$, we want bounds on $(\bar{w}, \tilde{f} - f)$ similar to those given in Section 3. To do this, we will prove the following lemma:

Lemma 4.7. *Adopt the notation and assumptions of Theorem* 4.6. *If* $w \in H_{-s}(\mathbf{M}^m)$, *then there exists a set* $\{\alpha_j\}_{j=1}^N \subset \mathbf{C}$ *for which* $w - \sum_{j=1}^N \alpha_j u_j \in \mathcal{S}_{\mathcal{I}}^{\perp}$. *Moreover, for any such set*

$$(4.7)\qquad (\bar{w}, \tilde{f} - f) = \left(\bar{w} - \sum_{j=1}^N \bar{\alpha}_j \bar{u}_j,\ \tilde{f} - f\right) = \left[\!\left[\ \tilde{u} - u, w - \sum_{j=1}^N \alpha_j u_j\ \right]\!\right]_{\mathcal{I}}.$$

Proof. Since $\tilde{f}$ and f both satisfy the interpolation conditions in Problem 4.4, we have that $(\bar{u}_j, \tilde{f} - f) = 0$ for $j = 1, \ldots, N$. Consequently, for every choice of α_j's, we see that $(\bar{w}, \tilde{f} - f) = (\bar{w} - \sum_{j=1}^N \bar{\alpha}_j \bar{u}_j, \kappa * (\tilde{u} - u) + \tilde{F} - F)$. If we can find α_j's such that $w - \sum_{j=1}^N \alpha_j u_j$ is in $\mathcal{S}_{\mathcal{I}}^{\perp}$, then, since $\tilde{F} - F$ is in $\mathcal{S}_{\mathcal{I}}$, we would have $(\bar{w} - \sum_{j=1}^N \bar{\alpha}_j \bar{u}_j, \tilde{F} - F) = 0$, and so

$$(\bar{w}, \tilde{f} - f) = \left(\bar{w} - \sum_{j=1}^N \bar{\alpha}_j \bar{u}_j, \kappa * (\tilde{u} - u)\right).$$

From this equation, the definition of $[\![\cdot, \cdot]\!]_{\mathcal{I}}$, and that both $\tilde{u} - u$ and $w - \sum_{j=1}^N \alpha_j u_j$ are in $\mathcal{S}_{\mathcal{I}}^{\perp}$, we obtain (4.7).

Thus, the whole result will follow if we can show the existence of α_j's for which $w - \sum_{j=1}^N \alpha_j u_j$ is in $\mathcal{S}_{\mathcal{I}}^{\perp}$. Doing this is equivalent to solving the system of equations below for all $i \in \mathcal{I}$:

$$(w, \bar{F}_i) = \sum_{j=1}^N \alpha_j \underbrace{(u_j, \bar{F}_i)}_{\bar{C}_{j,i}} = [C^* \alpha]_i.$$

Solving this is possible because $\operatorname{rank}(C) = |\mathcal{I}|$. ∎

If we combine the results in Theorems 4.5, 4.6, and Lemma 4.7, we easily obtain the following variant of the hypercircle inequality [3, p. 230]:

Corollary 4.8. *With the notation and assumptions of Theorem* 4.6, *we have*

$$|(\bar{w}, \tilde{f} - f)| \le \operatorname{dist}(\tilde{u}, \mathcal{U}_{\mathcal{I}}) \operatorname{dist}\left(w - \sum_{j=1}^N \alpha_j u_j, \mathcal{U}_{\mathcal{I}}\right),$$

where distances are computed relative to the norm (4.2).

Remark 4.9. With the exception of results that make specific use of kernels of the form (1.1), all the results in this section hold if we regard $\mathcal{S}_{\mathcal{I}}$ as a finite-dimensional subspace of smooth functions and the set $\{F_i : i \in \mathcal{I}\}$ is interpreted as an orthonormal basis for $\mathcal{S}_{\mathcal{I}}$.

Remark 4.10. With minor modifications, most results involving Problem 2.2 and based on Proposition 3.6 will also be true for Problem 4.4, provided the CSPD kernels employed have the form (1.1).

5. The Torus

In this section and Section 6 below, we will apply Proposition 3.6, Theorem 3.8, and other results from previous sections to several types of interpolation problems. For the circle and higher-dimensional tori, we will obtain rates for scattered data point interpolation and for a restricted class of Hermite problems that involve interpolating a fixed linear combination of function values and derivatives at scattered sites. In the next section we will deal with a point-evaluation problem for the 2-sphere. To apply Proposition 3.6 in any of these cases requires two things. First, we must show that the coefficients $\{c_j\}_{j=1}^N$ exist. Second, we need to bound the sequence $\{b_k\}$ so that the estimates in (3.14) and (3.17) are useful. Doing these two things is, in essence, our goal here and in the next section. In all cases the rates that we obtain reflect the smoothness of both the underlying data-generating function and the functions in the approximating subspace.

Let $\mathbf{T}^m$ be the m torus. Consider a kernel of the form

$$\kappa(\varphi, \varphi') = \prod_{\gamma=1}^{m} P_\gamma(\varphi^\gamma - \varphi'^\gamma), \tag{5.1}$$

where φ and φ' are standard periodic coordinates. Suppose that for some $s \geq 0$ each $P_\gamma \in H_{2s}(\mathbf{T}^1)$ is continuous and has the Fourier series expansion

$$P_\gamma(\varphi) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} a_k^\gamma e^{ik\varphi} \qquad \text{where} \quad a_k^\gamma > 0. \tag{5.2}$$

Noting that the eigenfunctions for $\mathbf{T}^m$ are just the functions $F_{k_1,\ldots,k_m}(\varphi^1, \ldots, \varphi^m) = (2\pi)^{-m/2} e^{i(k_1\varphi^1 + \cdots + k_m\varphi^m)}$ and using Corollary 2.6, we see that the kernel $\kappa(\varphi, \varphi')$ is strictly positive definite, continuous, has the form (1.1), and also belongs to $H_{2s}(\mathbf{T}^m)$. C^∞ kernels of this type were described in [19] and [5].

The simplest case is the circle, $\mathbf{T}^1$. The kernel in (5.1) is then a single function, $P(\varphi - \varphi')$, and $F_k(\varphi) = e^{ik\varphi}/\sqrt{2\pi}$. We want to apply the estimates from Section 3 to the Hermite Problem 2.2′ in which the distributions u_j are

$$u_j(\varphi) := \sum_{n=0}^{\nu} \upsilon_n \delta^{(n)}(\varphi - \varphi_j), \qquad j = 0 \ldots N-1, \tag{5.3}$$

where the φ_j's are distinct angles in the interval $[0, 2\pi)$. To get rates of approximation, we will have to put further restrictions on these angles; we will do this later. Also, we have labeled the distributions in the natural way, $j = 0 \ldots N-1$ rather than $j = 1 \ldots N$. In (5.3), ν is some finite integer, and the coefficients υ_n are in $\mathbf{C}$. In addition, we require

$$\nu < s - \tfrac{1}{4}, \quad \upsilon_0 \neq 0, \quad \upsilon_\nu \neq 0, \qquad \text{and} \qquad \sum_{n=0}^{\nu} \upsilon_n i^n k^n \neq 0 \quad \text{for all} \quad k \in \mathbf{Z}. \tag{5.4}$$

For later use, we remark that these conditions imply that

$$0 < c_{\mathcal{U}}^{-1} \le \frac{(1+k^2)^{\nu/2}}{|\sum_{n=0}^{\nu} \upsilon_n i^n k^n|} \le C_{\mathcal{U}} < \infty \qquad \text{for all} \quad k \in \mathbf{Z}. \tag{5.5}$$

To solve the Hermite problem using the distributions u_j, we need P to be at least 2ν times continuously differentiable. From Lemma 1.2, we can achieve this if we assume $s > \nu + \frac{1}{4}$; this is the reason for the condition $\nu < s - \frac{1}{4}$. As in Section 3, we will take $\mathcal{U}$ to be the span of the u_j's. We remark that the basis functions that we use to solve our Hermite problem are

$$\begin{aligned} P_j(\varphi) &:= P * u_j(\varphi) \\ &= \int_0^{2\pi} P(\varphi - \varphi')u_j(\varphi')\, d\varphi' \\ &= \sum_{n=0}^{\nu} \upsilon_n P^{(n)}(\varphi - \varphi_j). \end{aligned} \tag{5.6}$$

Our basic tool will be Proposition 3.6. Using it requires the following two lemmas:

Lemma 5.1. *Let the u_j's be as in* (5.3), *with the φ_j's being distinct angles in the interval* $[0, 2\pi)$, *let the c_j's be arbitrary complex numbers, and suppose that* (5.4) *holds. If $w \in H_{-s}(\mathbf{T}^1)$ and if $W_k := (w, \bar{F}_k) - \sum_{j=0}^{N-1} c_j(u_j, \bar{F}_k)$, then for all $k \in \mathbf{Z}$ we have*

$$W_k = \frac{1}{\sqrt{2\pi}} \left\{ (w, e^{-ik\varphi}) - \left(\sum_{n=0}^{\nu} \upsilon_n i^n k^n \right) \tilde{c}_k \right\}, \tag{5.7}$$

where

$$\tilde{c}_k = \sum_{j=0}^{N-1} c_j e^{-ik\varphi_j}. \tag{5.8}$$

Moreover, there exist unique scalars $c_1, \ldots, c_N$ such that $W_k = 0$ for all integers $k \in \mathcal{J}_N := [-[N/2], [(N-1)/2]] \cap \mathbf{Z}$.

Proof. We have the following chain of equations:

$$\begin{aligned} W_k &= (w, e^{-ik\varphi}/\sqrt{2\pi}) - \sum_{j=0}^{N-1} c_j(u_j, e^{-ik\varphi}/\sqrt{2\pi}) \\ &= (w, e^{-ik\varphi}/\sqrt{2\pi}) - \sum_{j=0}^{N-1} c_j \sum_{n=0}^{\nu} \upsilon_n \left(\delta^{(n)}(\varphi - \varphi_j), e^{-ik\varphi}/\sqrt{2\pi} \right) \\ &= \frac{1}{\sqrt{2\pi}} \left\{ (w, e^{-ik\varphi}) - \left(\sum_{n=0}^{\nu} \upsilon_n i^n k^n \right) \left(\sum_{j=0}^{N-1} c_j e^{-ik\varphi_j} \right) \right\}. \end{aligned}$$

Combining this and (5.8) yields (5.7). For $k \in \mathcal{J}_N$, we wish to solve the system of equations

$$\sum_{j=0}^{N-1} c_j e^{-ik\varphi_j} = \frac{(w, e^{-ik\varphi})}{\sum_{n=0}^{\nu} \upsilon_n i^n k^n}. \tag{5.9}$$

The determinant of the coefficient matrix for this system is a Vandermonde determinant. Since the φ_j's are distinct angles in $[0, 2\pi)$, the exponentials $e^{-i\varphi_j}$ correspond to distinct points on the unit circle. Inspecting the usual formula for a Vandermonde determinant shows that it is not zero; the system thus has a unique solution. With this choice of c_j's, we have

$$\tilde{c}_k = \frac{(w, e^{-ik\varphi})}{\sum_{n=0}^{\nu} \upsilon_n i^n k^n} \qquad \text{for all} \quad k \in \mathcal{J}_N. \tag{5.10}$$

Inserting this expression in (5.7) then yields $W_k = 0$ for $k \in \mathcal{J}_N$. ∎

Lemma 5.2. *With the notation and assumptions of Lemma* 5.1, *if* $k \notin \mathcal{J}_N$, *we have*

$$|W_k| \le (1 + k^2)^{s/2} \left(\|w\|_{-s} + \frac{c_{\mathcal{U}} |\tilde{c}_k|}{\sqrt{2\pi}(1 + ([N/2])^2)^{(s-\nu)/2}} \right). \tag{5.11}$$

In addition, if the c_j*'s are taken to be the unique coefficients for which* $W_k = 0$, $k \in \mathcal{J}_N$, *then*

$$\sup_{k \in \mathcal{J}_N} |\tilde{c}_k| \le \sqrt{2\pi} C_{\mathcal{U}} (1 + [N/2]^2)^{(s-\nu)/2} \|w\|_{-s}. \tag{5.12}$$

Proof. The distribution w is in $H_{-s}(\mathbf{T}^1)$; its Fourier coefficients thus satisfy a bound of the form

$$|(w, e^{-ik\varphi})| \le \sqrt{2\pi} \|w\|_{-s} (1 + k^2)^{s/2}, \qquad k \in \mathbf{Z}, \tag{5.13}$$

If we combine (5.13) with (5.7) and (5.5), we arrive at the estimate

$$|W_k| \le (1 + k^2)^{s/2} \left(\|w\|_{-s} + \frac{c_{\mathcal{U}} |\tilde{c}_k|}{\sqrt{2\pi}(1 + k^2)^{(s-\nu)/2)}} \right). \tag{5.14}$$

Note that $s > \nu$ and that if $k \notin \mathcal{J}_N$, then $|k| \ge [N/2]$. Consequently, $(1 + k^2)^{(s-\nu)/2} \ge (1 + ([N/2])^2)^{(s-\nu)/2}$. Using this in (5.14) yields (5.11). Next, take the c_j's to be those for which $W_k = 0$ holds when $k \in \mathcal{J}_N$. The $\tilde{c}_k$'s then satisfy (5.10). If we combine (5.10), (5.5), and (5.13), we get the following inequality:

$$|\tilde{c}_k| \le \sqrt{2\pi} C_{\mathcal{U}} \|w\|_{-s} (1 + k^2)^{(s-\nu)/2} \qquad \text{for all} \quad k \in \mathcal{J}_N. \tag{5.15}$$

Again, since $s > \nu$ and $|k| \le [N/2]$ if $k \in \mathcal{J}_N$, we see that $(1 + k^2)^{(s-\nu)/2} \le (1 + ([N/2])^2)^{(s-\nu)/2}$. Employing this in conjunction with (5.15) then yields (5.12). ∎

Thus far, we have made no real restrictions on the φ_j's used in the u_j's. At this point, we will assume that these angles are distributed quasi-uniformly in the following sense. Specifically, let

$$\varphi_j = \frac{2\pi}{N}(j + \varepsilon_j), \tag{5.16}$$

where the ε_j's are real numbers that satisfy

$$\sup_j |\varepsilon_j| = L, \qquad 0 \le L < \tfrac{1}{4}. \tag{5.17}$$

Using a discrete analogue of a theorem of Kadec [32, p. 43], we are able to estimate the distance from w to $\mathcal{U}$, provided the norm (3.3) is used to compute the distances.

Theorem 5.3. *Let the u_j's be as in* (5.3), *with the φ_j's being given by* (5.16), *and let $\mathcal{J}_N$ be as in Lemma* 5.1. *If $P \in H_{2s}(\mathbf{T}^1)$, and if $\sum_{k\in\mathbf{Z}}(1+k^2)^s a_k$ converges, then, for any $w \in H_{-s}(\mathbf{T}^1)$, we have*

$$\operatorname{dist}(w,\mathcal{U}) \le \|w\|_{-s}\sigma(L,N)\left(\sum_{k\notin\mathcal{J}_N}(1+k^2)^s a_k\right)^{1/2}, \tag{5.18}$$

where the a_k's are the Fourier coefficients of P, and

$$\sigma(L,N) := 1 + c_{\mathcal{U}}C_{\mathcal{U}}\begin{cases}1, & \text{if } L=0;\\ \sqrt{N/2}\,\csc(\pi/4-\pi L) & \text{if } 0<L<\frac{1}{4}.\end{cases} \tag{5.19}$$

Here, $c_{\mathcal{U}}$ and $C_{\mathcal{U}}$ are given in (5.5).

Proof. By Lemma 5.1, there are unique c_j's for which $W_k = 0$ when $k \in \mathcal{J}_N$. With this choice of c_j's, when k is in $\mathcal{J}_N$, $\tilde{c}_k$ is given by (5.10) and satisfies (5.12). There are now two cases. If $L = 0$ in (5.17), then the φ_j's are equally spaced angles, and $\tilde{c}_k = \hat{c}_k$, the discrete Fourier transform of the c_j's. Since the $\hat{c}_k$'s are periodic in k with period N, we see that $\sup_{k\in\mathbf{Z}}|\tilde{c}_k| = \sup_{k\in\mathcal{J}_N}|\tilde{c}_k|$. The second case is the one in which $0 < L < \frac{1}{4}$. From Theorem A.2, we have

$$|\tilde{c}_k| \le \sqrt{N/2}\csc\left(\frac{\pi}{4}-\pi L\right)\sup_{j\in\mathcal{J}_N}|\tilde{c}_j|, \qquad k\notin\mathcal{J}_N \quad (0<L<\tfrac{1}{4}).$$

Combining the estimates for $|\tilde{c}_k|$ from the two cases and using (5.19), we get

$$|\tilde{c}_k| \le \sigma(L,N)\sup_{j\in\mathcal{J}_N}|\tilde{c}_j|, \qquad k\notin\mathcal{J}_N.$$

Now use (5.12) to replace $\sup_{j\in\mathcal{J}_N}|\tilde{c}_j|$ in the inequality above, and then insert the result in (5.11) to get

$$|W_k| \le (1+k^2)^{s/2}\|w\|_{-s}\sigma(L,N), \qquad k\notin\mathcal{J}_N \quad (0\le L<\tfrac{1}{4}). \tag{5.20}$$

Clearly, Proposition 3.6 and Remark 3.7 now apply, provided that in (3.14) we take b_k to be the square of the right side of (5.20). Doing so then yields (5.18). ■

Suppose that $u \in \mathcal{U}$ is the unique solution to Problem 2.2′, given that the data is generated by a function $f := \kappa * v$, where $v \in H_{-s}(\mathbf{T}^1)$. Here, κ is of the form (5.1), P is as in (5.2), and the u_j's are as in (5.3). In addition, the interpolant generated is $\tilde{f} := \kappa * u$. We obtain the following:

Corollary 5.4. *With the notation used above and the assumptions from Theorem* 5.3, *if $w \in H_{-s}(\mathbf{T}^1)$, then*

$$|(w, f-\tilde{f})| \le \|v\|_{-s}\|w\|_{-s}\sigma(L,N)^2\left(\sum_{k\notin\mathcal{J}_N}(1+k^2)^s a_k\right). \tag{5.21}$$

Proof. Combine Proposition 3.2 and Theorem 5.3 to get (5.21). ■

We now wish to obtain a result for the case in which f is in $H_t(\mathbf{T}^1)$, where $t > s$. To do this, we will assume that (3.12) holds. If we also assume that the series $\sum_{k\in\mathbf{Z}}(1+k^2)^s a_k$ converges, then we must have $\sum_{k\in\mathbf{Z}}(1+k^2)^{(s-t)/2}$ convergent, which implies the additional constraint that t has to be larger than $s+1$.

Corollary 5.5. *Let $t > s+1$ and let $f \in H_t(\mathbf{T}^1)$ have $\tilde{f}$ as the interpolant described above, and suppose that $\sum_{k\in\mathbf{Z}}(1+k^2)^s a_k$ converges. With the assumptions from Theorem* 5.3, *if* (3.12) *holds for the a_k's and if $w \in H_{-s}(\mathbf{T}^1)$, then*

$$|(w, f-\tilde{f})| \leq c\|f\|_t\|w\|_{-s}\sigma(L,N)^2\left(\sum_{k\notin\mathcal{J}_N}(1+k^2)^s a_k\right), \tag{5.22}$$

where c is as in (3.12).

Proof. Since (3.12) holds and since $t > s+1 > s$, Proposition 3.5 applies. Thus, we see that $f = \kappa * v$, with $\|v\|_{-s} \leq c\|f\|_t$. Using this in (5.21) yields (5.22). ■

Remark 5.6. In Corollary 5.5, the right side of (5.22) is the tail end of a convergent series, so we would like to have the a_k's decay as quickly as possible. On the other hand, (3.12) limits the rate at which the coefficients can decay. Now, the rate at which the a_k's decay roughly translates into the smoothness class to which P belongs. Let $s' \geq s$. With a little work we can show that if $P \in H_{s+s'}(\mathbf{T}^1)$, then $s \leq s' < t - \frac{1}{2}$. We also remark that it should be true that if $w \in H_{-r}$, $r < s$, then we should get better estimates. Unfortunately, the Hilbert space techniques used here do not provide a suitable means for establishing this. Finally, we note the mesh size $h_N = \max_{0\leq j\leq N-1}(\varphi_{j+1}-\varphi_j)$, where φ_j is given by (5.16) and $\varphi_N := \varphi_0 + 2\pi$, satisfies $\pi/N \leq h_N \leq 3\pi/N$. Consequently, the estimates we derive can be related to classical estimates employing the mesh size. ■

We now turn to higher-dimensional tori, $\mathbf{T}^m$, with $m \geq 2$. We will use a kernel of the form (5.1), and our space of distributions will be

$$\mathcal{U} := \mathcal{U}^1 \otimes \cdots \otimes \mathcal{U}^m, \tag{5.23}$$

where each space $\mathcal{U}^\gamma$ is a finite-dimensional subset of $H_{-s_\gamma}(\mathbf{T}^1)$ and where $s_1+\cdots+s_m = s$. In addition, we will assume that each $\mathcal{U}^\gamma$ is spanned by N_γ distributions of the form (5.3), with angles involved being not only distinct but also of the form (5.16), with N replaced by N_γ, and the corresponding ε_j's there satisfying (5.17) with L replaced by L_γ.

Corollary 5.7. *If $\sum_{k\in\mathbf{Z}}(1+k^2)^{s_\gamma}a_k^\gamma$ converges for $\gamma = 1,\ldots,m$, and if $w := w^1 \otimes \cdots \otimes w^m$, where $w^\gamma \neq 0$ is in $H_{-s_\gamma}(\mathbf{T}^1)$, then*

$$\begin{aligned}\operatorname{dist}(w,\mathcal{U}) \leq &\left(\prod_{\gamma=1}^m \|\kappa_\gamma\|_{2s_\gamma}\|w^\gamma\|_{-s_\gamma}\right)\\ &\times\left(\sum_{\gamma=1}^m \frac{\sigma(L_\gamma,N_\gamma)^2}{\|\kappa_\gamma\|_{2s_\gamma}^2}\left(\sum_{k\notin\mathcal{I}_{N_\gamma}}(1+k^2)^{s_\gamma}a_k^\gamma\right)\right)^{1/2},\end{aligned} \tag{5.24}$$

where the a_k^γ's are the Fourier coefficients of P_γ, the set $\mathcal{I}_{N_\gamma}$ is as in Lemma 5.1, *and the factor $\sigma(\cdot,\cdot)$ is given in* (5.19).

Proof. By Theorem 5.3, we have that

$$\operatorname{dist}(w^\gamma, \mathcal{U}^\gamma)^2 \le \|w^\gamma\|_{-s_\gamma}^2 \sigma(L_\gamma, N_\gamma)^2 \left(\sum_{k \notin \mathcal{J}_N} (1+k^2)^{s_\gamma} a_k^\gamma \right).$$

From (3.19), we also see that

$$\llbracket w \rrbracket^2 = \prod_{\beta=1}^m \llbracket w^\beta \rrbracket_\beta^2.$$

Multiply both sides of the previous inequality by $\llbracket w \rrbracket^2 / \llbracket w^\gamma \rrbracket_\gamma^2$ and use the last equation to obtain

$$\frac{\llbracket w \rrbracket^2 \operatorname{dist}(w^\gamma, \mathcal{U}^\gamma)^2}{\llbracket w^\gamma \rrbracket_\gamma^2} \le \|w^\gamma\|_{-s_\gamma}^2 \sigma(L_\gamma, N_\gamma)^2 \left(\sum_{k \notin \mathcal{J}_N} (1+k^2)^{s_\gamma} a_k^\gamma \right) \prod_{\beta \ne \gamma}^m \llbracket w^\beta \rrbracket_\beta^2.$$

Recall that

$$\llbracket w^\beta \rrbracket_\beta^2 \le \|\kappa_\beta\|_{2s_\beta}^2 \|w^\beta\|_{-s_\beta}^2.$$

Combining this with the previous inequality yields

$$\frac{\llbracket w \rrbracket^2 \operatorname{dist}(w^\gamma, \mathcal{U}^\gamma)^2}{\llbracket w^\gamma \rrbracket_\gamma^2} \le \frac{\sigma(L_\gamma, N_\gamma)^2}{\|\kappa_\gamma\|_{2s_\gamma}^2} \left(\sum_{k \notin \mathcal{J}_N} (1+k^2)^{s_\gamma} a_k^\gamma \right) \left(\prod_{\beta=1}^m \|\kappa_\beta\|_{2s_\beta}^2 \|w^\beta\|_{-s_\beta}^2 \right).$$

The inequality (5.24) follows from using the inequality above in conjunction with (3.20) of Proposition 3.9. ■

We can now use this result to obtain rate estimates for the case of $\mathbf{T}^m$.

Theorem 5.8. *With the notation and assumptions of Corollary* 5.7, *if $u \in H_{-s}(\mathbf{T}^m)$ is the unique distribution in $\mathcal{U}$ for which $\kappa * u$ solves Problem* 2.2′, *and if $v \in H_{-s}(\mathbf{T}^m)$ is any other distribution for which $\kappa * v$ also interpolates the data in Problem* 2.2′, *then*

$$|(\bar{w}, f - \tilde{f})| \le K_s \|v\|_{-s} \left(\prod_{\gamma=1}^m \|w^\gamma\|_{-s_\gamma} \right) \times \left(\sum_{\gamma=1}^m \frac{\sigma(L_\gamma, N_\gamma)^2}{\|\kappa_\gamma\|_{2s_\gamma}^2} \left(\sum_{k \notin \mathcal{I}_{N_\gamma}} (1+k^2)^{s_\gamma} a_k^\gamma \right) \right)^{1/2}, \tag{5.25}$$

where

$$K_s := \|\kappa\|_{2s} \prod_{\gamma=1}^m \|\kappa_\gamma\|_{2s_\gamma}. \tag{5.26}$$

Proof. By Remark 3.4 and the inequality $[\![v]\!] \le \|\kappa\|_{2s}\|v\|_{-s}$, we see that $\operatorname{dist}(v, \mathcal{U}) \le \|\kappa\|_{2s}\|v\|_{-s}$. The result then follows on combining this with Proposition 3.2 and Corollary 5.7. ■

If each of the P_γ's in the product (5.1) has coefficients that satisfy a bound of the form

$$a_k^\gamma \ge c_\gamma^{-1}(1+k^2)^{-(s_\gamma+t_\gamma)/2}, \tag{5.27}$$

then $a_{k_1,\ldots,k_m} = \prod_{\gamma=1}^m a_{k_\gamma}^\gamma$ satisfies

$$a_{k_1,\ldots,k_m} \ge c^{-1}\left(1+\sum_{\gamma=1}^m k_\gamma^2\right)^{-(s+t)/2} = c^{-1}(1+\lambda_{k_1,\ldots,k_m})^{-(s+t)/2},$$

where $c = \prod_{\gamma=1}^m c_\gamma$, $s = \sum_{\gamma=1}^m s_\gamma$, and $t = \sum_{\gamma=1}^m t_\gamma$. This, coupled with Theorems 3.8 and 5.8, yields the following result:

Corollary 5.9. *If the a_k^γ's satisfy* (5.27), *and if $f = \kappa * v$ is in $H_t(\mathbf{T}^m)$, then we may replace* (5.25) *by*

$$|(\bar{w}, f - \tilde{f})| \le K_s c\|f\|_t \left(\prod_{\gamma=1}^m \|w^\gamma\|_{-s_\gamma}\right) \times \left(\sum_{\gamma=1}^m \frac{\sigma(L_\gamma, N_\gamma)^2}{\|\kappa_\gamma\|_{2s_\gamma}^2}\left(\sum_{k\notin \mathcal{I}_{N_\gamma}} (1+k^2)^{s_\gamma} a_k^\gamma\right)\right)^{1/2}. \tag{5.28}$$

We close this section by remarking that the higher-dimensional tori estimates we obtained in Theorem 5.8 and Corollary 5.9 are not optimal. We conjecture that estimates similar to ones for the circle can be found.

6. The 2-Sphere

Our next example deals with the 2-sphere, $\mathbf{S}^2$. The interpolation problem that we will deal with in this section will be one in which the distributions in Problem 2.2′ are point evaluations at points $p_{j,k}$ that we will describe below; that is,

$$u_{j,k} = \delta_{p_{j,k}} \qquad \text{and} \qquad \mathcal{U} := \operatorname{Span}\{u_{j,k}\}. \tag{6.1}$$

The results we obtain in this section depend on recent results of Driscoll and Healy [4], and we will adopt the convention for spherical coordinates that is used by them and that is customary in physics: the angle $\theta \in [0, \pi]$ is measured off the positive z-axis and the angle $\varphi \in [0, 2\pi)$ is measured off the x-axis in the x–y plane. In addition, take Λ to be a fixed positive integer, then let

$$\begin{cases} \theta_j = \dfrac{\pi j}{2\Lambda}, & j = 0, \ldots, 2\Lambda - 1, \\[2mm] \varphi_k = \dfrac{\pi k}{\Lambda}, & k = 0, \ldots, 2\Lambda - 1, \\ (\theta_j, \varphi_k) = \text{coordinates of } p_{j,k}. \end{cases} \tag{6.2}$$

The eigenfunctions of the Laplace–Beltrami operator for $\mathbf{S}^2$ are the $Y_{\ell,m}$'s, where $\ell = 0, 1, 2, \ldots$ and where, for ℓ fixed, $m = -\ell, \ldots, \ell$; these functions are described in detail in [4(§2)], [18], and [19], and in spherical coordinates are related to associated Legendre functions. We will not need their explicit form here, although we will need some of their properties. The $Y_{\ell,m}$'s form an orthonormal basis for $L^2(\mathbf{S}^2)$ with respect to the standard measure on the sphere, $d\mu$, which has the form $d\mu = \sin(\theta)\, d\theta\, d\varphi$ in the coordinates above. As in [4], we will denote the expansion coefficient for a distribution h relative to $Y_{\ell,m}$ by $\widehat{h}(\ell, m)$; that is, set

$$\widehat{h}(\ell, m) = \int_{\mathbf{S}^2} h\overline{Y}_{\ell,m}\, d\mu.$$

We will say that h is *band-limited* if the orthogonal expansion for h contains only finitely many terms. Driscoll and Healy have established the following "sampling theorem."

Theorem 6.1 (4, Theorem 3). *If $f(p)$ is a band-limited function on $\mathbf{S}^2$ for which $\widehat{f}(\ell, m) = 0$ for $\ell \geq \Lambda$, then there exist coefficients $\alpha_j^{(\Lambda)}$ that are independent of f such that*

$$\widehat{f}(\ell, m) = \frac{\sqrt{2\pi}}{2\Lambda} \sum_{j=0}^{2\Lambda-1} \sum_{k=0}^{2\Lambda-1} \alpha_j^{(\Lambda)} f(p_{j,k})\overline{Y}_{\ell,m}(p_{j,k})$$

for $\ell < \Lambda$ and $|m| \leq \ell$.

We are now ready to proceed with the "rate of approximation" estimates, which we will obtain by applying Proposition 3.6. To apply this proposition, we will need the lemmas below.

Lemma 6.2. *Let $w \in H_{-s}(\mathbf{S}^2)$ and let the expansion coefficients for w be $\widehat{w}(\ell, m) = (w, \overline{Y}_{\ell,m})$. If w_Λ is the band-limited function defined by*

$$w_\Lambda(p) := \sum_{\ell=0}^{\Lambda-1} \sum_{m=-\ell}^{\ell} \widehat{w}(\ell, m) Y_{\ell,m}(p), \tag{6.3}$$

and if

$$c_{j,k} := \frac{\sqrt{2\pi}}{2\Lambda} w_\Lambda(p_{j,k})\alpha_j^\Lambda, \qquad 0 \leq j, \quad k \leq 2\Lambda - 1, \tag{6.4}$$

then

$$\widehat{w}(\ell, m) - \sum_{j=0}^{2\Lambda-1} \sum_{k=0}^{2\Lambda-1} c_{j,k}(u_{j,k}, \overline{Y}_{\ell,m}) = 0, \qquad 0 \leq \ell < \Lambda, \quad |m| \leq \ell. \tag{6.5}$$

Proof. By Theorem 6.1,

$$\widehat{w}(\ell, m) = \frac{\sqrt{2\pi}}{2\Lambda} \sum_{j=0}^{2\Lambda-1} \sum_{k=0}^{2\Lambda-1} w_\Lambda(p_{j,k})\alpha_j^\Lambda \overline{Y_{\ell,m}(p_{j,k})}, \qquad 0 \leq \ell < \Lambda, \quad |m| \leq \ell.$$

Since $\overline{Y_{\ell,m}(p_{j,k})} = (u_{j,k}, \overline{Y_{\ell,m}})$, choosing the $c_{j,k}$'s to be the coefficients given in (6.4) results in (6.5). ■

Next, we need to estimate various quantities related to ones appearing in the lemma above. For all ℓ and m, define

$$W_{\ell,m} := \widehat{w}(\ell,m) - \frac{\sqrt{2\pi}}{2\Lambda}\sum_{j=0}^{2\Lambda-1}\sum_{k=0}^{2\Lambda-1} w_\Lambda(p_{j,k})\alpha_j^\Lambda(u_{j,k}\overline{Y}_{\ell,m}), \tag{6.6}$$

where the α_j^Λ's are the coefficients from Theorem 6.1. By the previous lemma, we see that $W_{\ell,m} = 0$ for all $0 \le \ell < \Lambda,\ |m| \le \ell$. The results below provide a bound on this quantity in the case where $\ell \ge \Lambda$.

Lemma 6.3. *If $w_\Lambda(p)$ is given by* (6.3) *and if $\Lambda \ge 1$, then for all p*

$$|w_\Lambda(p)| \le \frac{\Lambda^{1+s}}{\sqrt{4\pi}}\|w\|_{-s}. \tag{6.7}$$

In addition, for all $\ell \ge \Lambda$, we have

$$\left|\frac{\sqrt{2\pi}}{2\Lambda}\sum_{j=0}^{2\Lambda-1}\sum_{k=0}^{2\Lambda-1} w_\Lambda(p_{j,k})\alpha_j^\Lambda(u_{j,k}\overline{Y}_{\ell,m})\right| \le \sqrt{\frac{2\ell+1}{8\pi}}\Lambda^{s+1}\|w\|_{-s}\sum_{j=0}^{2\Lambda-1}|\alpha_j^\Lambda|. \tag{6.8}$$

Proof. From Schwarz's inequality applied to the right-hand side of (6.3), we see that

$$|w_\Lambda(p)|^2 \le \left(\sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell}|\widehat{w}(\ell,m)|^2\right)\left(\sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell} Y_{\ell,m}(p)\bar{Y}_{\ell,m}(p)\right). \tag{6.9}$$

The first term on the right above satisfies this chain of inequalities:

$$\begin{aligned}\sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell}|\widehat{w}(\ell,m)|^2 &= \sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell}(1+\ell(\ell+1))^s\frac{|\widehat{w}(\ell,m)|^2}{(1+\ell(\ell+1))^s} \\ &\le (1+\Lambda(\Lambda-1))^s\sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell}\frac{|\widehat{w}(\ell,m)|^2}{(1+\ell(\ell+1))^s} \\ &\le \Lambda^{2s}\|w\|_{-s}^2.\end{aligned} \tag{6.10}$$

The second term on the right of (6.9) can be evaluated exactly via the Addition Theorem for spherical harmonics [18, p. 10]; the result is this:

$$\sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell} Y_{\ell,m}(p)\bar{Y}_{\ell,m}(p) = \sum_{\ell=0}^{\Lambda-1}\frac{2\ell+1}{4\pi}P_\ell(\underbrace{p\cdot p}_{1}),$$

where P_ℓ is the standard Legendre polynomial of degree ℓ, which satisfies the normalization condition that $P_\ell(1) = 1$. Since $\sum_{\ell=0}^{\Lambda-1}(2\ell+1) = \Lambda^2$, we see that

$$\sum_{\ell=0}^{\Lambda-1}\sum_{m=-\ell}^{\ell} Y_{\ell,m}(p)\bar{Y}_{\ell,m}(p) = \frac{\Lambda^2}{4\pi}. \tag{6.11}$$

Taking square roots after combining (6.9), (6.10), and (6.11) yields (6.7).

To establish (6.8), first note that from (6.7) we have

$$\left|\frac{\sqrt{2\pi}}{2\Lambda}\sum_{j=0}^{2\Lambda-1}\sum_{k=0}^{2\Lambda-1} w_\Lambda(p_{j,k})\alpha_j^\Lambda(u_{j,k},\overline{Y}_{\ell,m})\right| \leq \frac{\Lambda^s\|w\|_{-s}}{\sqrt{8}}\sum_{j=0}^{2\Lambda-1}\sum_{k=0}^{2\Lambda-1}|\alpha_j^\Lambda||(u_{j,k}\overline{Y}_{\ell,m})|. \tag{6.12}$$

As we noted in the previous proof, $(u_{j,k},\overline{Y}_{\ell,m}) = \overline{Y}_{\ell,m}(p_{j,k})$. Also, by [18, Lemma 8, p. 14], we have

$$|Y_{\ell,m}(p)| \leq \sqrt{\frac{2\ell+1}{4\pi}}.$$

Combining these two facts and using them in conjunction with (6.12), we find that

$$\begin{aligned}\left|\frac{\sqrt{2\pi}}{2\Lambda}\sum_{j=0}^{2\Lambda-1}\sum_{k=0}^{2\Lambda-1} w_\Lambda(p_{j,k})\alpha_j^\Lambda(u_{j,k},\overline{Y}_{\ell,m})\right| &\leq \frac{\Lambda^s\|w\|_{-s}\sqrt{2\ell+1}}{\sqrt{32\pi}}\sum_{j=0}^{2\Lambda-1}\sum_{k=0}^{2\Lambda-1}|\alpha_j^\Lambda| \\ &\leq \sqrt{\frac{2\ell+1}{8\pi}}\Lambda^{s+1}\|w\|_{-s}\sum_{j=0}^{2\Lambda-1}|\alpha_j^\Lambda|. \qquad \blacksquare\end{aligned}$$

We next wish to estimate the sum $\sum_{j=0}^{2\Lambda-1}|\alpha_j^\Lambda|$. This we can do when the integer Λ is a power of 2.

Lemma 6.4. *If* Λ *is a power of* 2, *then*

$$\sum_{j=0}^{2\Lambda-1}|\alpha_j^\Lambda| \leq \log_2(16\Lambda). \tag{6.13}$$

Proof. From [4, p. 216], if Λ is a power of 2, then

$$\alpha_j^\Lambda = \frac{2\sqrt{2}}{2\Lambda}\sin\left(\frac{\pi j}{2\Lambda}\right)\sum_{\ell=0}^{\Lambda-1}\frac{1}{2\ell+1}\sin\left(\frac{\pi j(2\ell+1)}{2\Lambda}\right), \qquad j = 0,\ldots,2\Lambda-1.$$

From this it follows that, for $j = 0,\ldots,2\Lambda-1$,

$$\begin{aligned}|\alpha_j^\Lambda| &\leq \frac{2\sqrt{2}}{2\Lambda}\sum_{\ell=0}^{\Lambda-1}\frac{1}{2\ell+1} \\ &\leq \frac{\sqrt{2}}{\Lambda}\left(1+\int_0^{\Lambda-1}\frac{dx}{2x+1}\right) \\ &\leq \frac{\sqrt{2}}{\Lambda}\left(1+\tfrac{1}{2}\ln(2\Lambda-1)\right) \\ &\leq \frac{1}{2\Lambda}\log_2(16\Lambda).\end{aligned}$$

Consequently, $|\alpha_j^\Lambda| \le \log_2(16\Lambda)/2\Lambda$ and so

$$\sum_{j=0}^{2\Lambda-1} |\alpha_j^\Lambda| \le \frac{\log_2(16\Lambda)}{2\Lambda}(2\Lambda) = \log_2(16\Lambda). \qquad \blacksquare$$

We can now bound the $W_{\ell,m}$'s defined in (6.6). The estimate that we obtain will allow us to use Proposition 3.6 to get rates of approximation.

Lemma 6.5. *Let* Λ *be a power of* 2. *If* $\ell \ge \Lambda$, *then*

$$|W_{\ell,m}| \le \left(1 + \frac{1}{2\sqrt{\pi}}\Lambda^{3/2}\log_2(16\Lambda)\right)(1+\ell(\ell+1))^{s/2}\|w\|_{-s}. \tag{6.14}$$

Proof. Note that, since w is in H_{-s}, its Fourier expansion coefficients satisfy

$$|\widehat{w}(\ell,m)| \le (1+\ell(\ell+1))^{s/2}\|w\|_{-s}.$$

This provides a bound for the first term in (6.6). A bound for the second term in (6.6) may be obtained by combining (6.8) and (6.13). This yields the following bound on $W_{\ell,m}$:

$$|W_{\ell,m}| \le (1+\ell(\ell+1))^{s/2}\|w\|_{-s} + \sqrt{\frac{2\ell+1}{8\pi}}\Lambda^{s+1}\|w\|_{-s}\log_2(16\Lambda).$$

Rearranging terms on the right above, we have

$$|W_{\ell,m}| \le (1+\ell(\ell+1))^{s/2}\|w\|_{-s}(1+A_s(\ell,\Lambda)), \tag{6.15}$$

where

$$A_s(\ell,\Lambda) = (1+\ell(\ell+1))^{-s/2}\sqrt{\frac{2\ell+1}{8\pi}}\Lambda^{s+1}\log_2(16\Lambda).$$

With a little algebra, we see that

$$A_s(\ell,\Lambda) = \frac{1}{2\sqrt{\pi}}\left(\frac{\Lambda}{\ell+\frac{1}{2}}\right)^{s-1/2}(1+\tfrac{3}{4}(\ell+\tfrac{1}{2})^{-2})^{-(s/2)},\ \Lambda^{3/2}\log_2(16\Lambda).$$

Since $\ell \ge \Lambda$, it follows that

$$A_s(\ell,\Lambda) \le \frac{1}{2\sqrt{\pi}}\Lambda^{3/2}\log_2(16\Lambda).$$

Using this bound in (6.15) results in (6.14). $\blacksquare$

We are now ready to apply Proposition 3.6. In what follows, we assume that the SPD kernel $\kappa \in H_{2s}(\mathbf{S}^2 \times \mathbf{S}^2)$ has the form

$$\kappa(p,q) := \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell,m}Y_{\ell,m}(p)\overline{Y}_{\ell,m}(q), \tag{6.16}$$

where the coefficients satisfy $a_{\ell,m} > 0$ for all ℓ and m. We will now prove the estimate on the distance from a distribution w to the subspace $\mathcal{U}$.

Theorem 6.6. *Let w be in H_{-s}, let $\mathcal{U}$ be as in* (6.1), *and let κ be as in* (6.16). *If Λ is a power of* 2, *and if the series $\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}(1+\ell(\ell+1))^s a_{\ell,m}$ is convergent, then*

(6.17)
$$\operatorname{dist}(w,\mathcal{U}) \le \left(1+\frac{1}{\sqrt{\pi}}\Lambda^{3/2}\log_2(16\Lambda)\right)\|w\|_{-s}\left(\sum_{\ell=\Lambda}^{\infty}\sum_{m=-\ell}^{\ell}(1+\ell(\ell+1))^s a_{\ell,m}\right)^{1/2}.$$

Proof. Combine Lemmas 6.2 and 6.5 with Proposition 3.6. ■

We can now use this result to obtain rate estimates for the case of the 2-sphere, $\mathbf{S}^2$.

Corollary 6.7. *With the notation and assumptions of Theorem* 6.6, *if $u \in H_{-s}(\mathbf{S}^2)$ is the unique distribution in $\mathcal{U}$ for which $\kappa * u$ solves Problem* 2.2′, *and if $v \in H_{-s}(\mathbf{S}^2)$ is any other distribution for which $\kappa * v$ also interpolates the data in Problem* 2.2′, *then for all $w \in H_{-s}$*

$$|(\bar{w},\ f-\tilde{f})| \le \left(1+\frac{1}{2\sqrt{\pi}}\Lambda^{3/2}\log_2(16\Lambda)\right)^2\|v\|_{-s}\|w\|_{-s} \times\left(\sum_{\ell=\Lambda}^{\infty}\sum_{m=-\ell}^{\ell}(1+\ell(\ell+1))^s a_{\ell,m}\right). \tag{6.18}$$

Proof. By Proposition 3.2,

$$|(\bar{w},\ f-\tilde{f})| \le \operatorname{dist}(v,\mathcal{U})\operatorname{dist}(w,\mathcal{U}).$$

Both w and v are in H_{-s}. Thus, (6.18) follows on applying the distance estimates in Theorem 6.6. ■

We now wish to obtain a result for the case in which f is in $f \in H_t(\mathbf{S}^2)$, where $t > s$. To do this, we will assume that (3.12) holds. If we also assume that the series $\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}(1+\ell(\ell+1))^s a_{\ell,m}$ converges, then we must have $\sum_{\ell=0}^{\infty}(2\ell+1)(1+\ell(\ell+1))^{-(t-s)/2}$ convergent, which implies the additional constraint that t has to be larger than $s+2$.

Corollary 6.8. *Let $t > s+2$, and adopt the notation of Corollary* 6.7. *If $f \in H_t(\mathbf{S}^2)$ and if* (3.12) *holds, then*

$$|(w,\ f-\tilde{f})| \le c\|f\|_t\|w\|_{-s}\left(1+\frac{1}{2\sqrt{\pi}}\Lambda^{3/2}\log_2(16\Lambda)\right)^2 \times\left(\sum_{\ell=\Lambda}^{\infty}\sum_{m=-\ell}^{\ell}(1+\ell(\ell+1))^s a_{\ell,m}\right), \tag{6.19}$$

where c is as in (3.12).

Proof. Since (3.12) holds and since $t > s+2 > s$, Proposition 3.5 applies. Thus, we see that $f = \kappa * v$, with $\|v\|_{-s} \le c\|f\|_t$. Using this in (6.18) yields (6.19). ■

Remark 6.9. As in the the case of the circle (see Remark 5.6), we would like to have the $a_{\ell,m}$'s decay as quickly as possible, but again (3.12) limits the rate at which these coefficients can decay. This roughly translates into limiting the smoothness class to which κ belongs. If $s' \geq s$, then we can show that if $\kappa \in H_{s+s'}(\mathbf{S}^2 \times \mathbf{S}^2)$, then $s \leq s' < t-1$. ■

Appendix

Consider angles of the form

$$\text{(A.1)} \qquad \varphi_j = \frac{2\pi}{N}(j+\varepsilon_j), \qquad i = 0, \ldots, N-1, \quad |\varepsilon_j| \leq L < \tfrac{1}{4}.$$

Recall that in (5.8), we defined

$$\text{(A.2)} \qquad \tilde{c}_k := \sum_{j=0}^{N-1} c_j e^{-ik\varphi_j}$$

for all k. Let $\mathcal{J}_N := [-[N/2], [(N-1)/2]] \cap \mathbf{Z}$. From the proof of Lemma 5.1, we can solve the equations above to get the c_j's in terms of the $\tilde{c}_k$'s, with $k \in \mathcal{J}_N$. We wish to get estimates on the 2-norm of the column matrix $c := (\, c_0 \quad \cdots \quad c_{N-1}\,)^T$ in terms of the 2-norm of $\tilde{c} := (\, \tilde{c}_{-[N/2]} \quad \cdots \quad \tilde{c}_{[(N-1)/2]}\,)^T$. We remark that, if $\hat{c} := (\, \hat{c}_{-[N/2]} \quad \cdots \quad \hat{c}_{[(N-1)/2]}\,)^T$, where

$$\text{(A.3)} \qquad \hat{c}_k = \sum_{j=0}^{N-1} c_j e^{2\pi ikj/N}$$

is the discrete Fourier transform of the c_j's, then $\tilde{c} = \hat{c}$.

The technique that we use in this appendix is a modification of the one employed by Young to prove a version of a famous theorem of Kadec [32, 1.10]. We begin with the observation that for $k \in \mathcal{J}_N$,

$$\text{(A.4)} \qquad \begin{aligned} e^{-ik\varphi_j} &= e^{2\pi ijk/N}(1-(1-e^{2\pi i\varepsilon_j k/N})) \\ &= e^{2\pi ijk/N} - e^{2\pi ijk/N}(1-e^{2\pi i\varepsilon_j k/N}). \end{aligned}$$

From this equation and (A.3), we can write $\tilde{c}_k$ in the form

$$\text{(A.5)} \qquad \tilde{c}_k = \hat{c}_k - \sum_{j=0}^{N-1} c_j e^{-2\pi ijk/N}(1-e^{-2\pi i\varepsilon_j k/N}).$$

For $k \in \mathcal{J}_N$, set

$$t = -\frac{2\pi k}{N} \qquad \text{and} \qquad \delta = \varepsilon_j$$

in the expansion [32, p. 43]

$$\begin{aligned} 1 - e^{i\delta t} = \left(1 - \frac{\sin \pi\delta}{\pi\delta}\right) &+ \sum_{m=1}^{\infty} \frac{2(-1)^m \delta \sin(\delta\pi)}{\pi(m^2-\delta^2)} \cos(mt) \\ &+ i \sum_{m=1}^{\infty} \frac{(-1)^m 2\delta \cos(\pi\delta)}{\pi((m-\frac{1}{2})^2-\delta^2)} \sin((m-\tfrac{1}{2})t), \end{aligned}$$

and insert the result in (A.5):

$$
\begin{aligned}
\tilde{c}_k = \hat{c}_k &- \sum_{j=0}^{N-1} c_j e^{-2\pi ijk/N}\left(1-\frac{\sin(\pi\varepsilon_j)}{\pi\varepsilon_j}\right)\\
&- \sum_{m=1}^{\infty}\frac{2(-1)^m\cos(2\pi km/N)}{\pi}\sum_{j=0}^{N-1} c_j e^{-2\pi ijk/N}\frac{\varepsilon_j\sin(\varepsilon_j\pi)}{m^2-\varepsilon_j^2}\\
&- \sum_{m=1}^{\infty}\frac{2i(-1)^{m+1}\sin\left(2\pi k\left(m-\frac{1}{2}\right)/\pi\right)}{N}\sum_{j=0}^{N-1} c_j e^{-2\pi ijk/N}\frac{\varepsilon_j\cos(\varepsilon_j\pi)}{\left(m-\frac{1}{2}\right)^2-\varepsilon_j^2}.
\end{aligned}
$$

Define the quantities

$$
\text{(A.6)}\qquad \begin{cases}
A_j = 1-\dfrac{\sin(\pi\varepsilon_j)}{\pi\varepsilon_j},\\
B_j^m = \dfrac{\varepsilon_j\sin(\varepsilon_j\pi)}{m^2-\varepsilon_j^2},\\
C_j^m = \dfrac{\varepsilon_j\cos(\varepsilon_j\pi)}{(m-\frac{1}{2})^2-\varepsilon_j^2}.
\end{cases}
$$

This puts $\tilde{c}_k$ into the form

$$
\begin{aligned}
\tilde{c}_k = \hat{c}_k &- \sum_{j=0}^{N-1} c_j A_j e^{-2\pi ijk/N}\\
&- \sum_{m=1}^{\infty}\frac{2(-1)^m\cos\left(\frac{2\pi km}{N}\right)}{\pi}\left(\sum_{j=0}^{N-1} c_j B_j^m e^{-2\pi ijk/N}\right)\\
&- \sum_{m=1}^{\infty}\frac{2(-1)^{m+1}i\sin\left(\frac{2\pi k}{N}(m-\frac{1}{2})\right)}{\pi}\left(\sum_{j=0}^{N-1} c_j C_j^m e^{-2\pi ijk/N}\right)
\end{aligned}
$$

By the convolution theorem, we have

$$
\begin{aligned}
\tilde{c}_k = \hat{c}_k &- \frac{1}{N}(\hat{c}*\hat{A})_k\\
&- \sum_{m=1}^{\infty}\frac{2(-1)^m}{\pi}\cos\left(\frac{2\pi km}{N}\right)\frac{(\hat{c}*\hat{B}^m)_k}{N}\\
&- \sum_{m=1}^{\infty}\frac{2(-1)^{m+1}i}{\pi}\sin\left(\frac{2\pi k}{N}(m-\frac{1}{2})\right)\frac{(\hat{c}*\hat{C}^m)_k}{N}.
\end{aligned}
$$

Setting

$$
\begin{aligned}
\hat{\mathcal{B}}_k^m &= (-1)^m\cos\left(\frac{2\pi km}{N}\right)\frac{(\hat{c}*\hat{B}^m)_k}{N},\\
\hat{\mathcal{C}}_k^m &= (-1)^{m+1}i\sin\left(\frac{2\pi k}{N}(m-\frac{1}{2})\right)\frac{(\hat{c}*\hat{C}^m)_k}{N},
\end{aligned}
$$

we may rewrite the previous equation as

$$\tilde{c} = \hat{c} - T\hat{c} = (I - T)\hat{c}, \tag{A.7}$$

where the linear transformation $T\colon \mathbf{C}^N \to \mathbf{C}^N$ is defined via

$$T\hat{c} = \frac{1}{N}(\hat{c} * \hat{A}) + \sum_{m=1}^{\infty} \frac{2}{\pi}(\hat{\mathcal{B}}^m + \hat{\mathcal{C}}^m). \tag{A.8}$$

We wish to estimate $\|T\|_{2,2}$. By the triangle inequality, we have

$$\|T\hat{c}\| \le \frac{1}{N}\|\hat{c} * \hat{A}\| + \sum_{m=1}^{\infty} \frac{2}{\pi}(\|\hat{\mathcal{B}}^m\| + \|\hat{\mathcal{C}}^m\|).$$

Recall that $\|\hat{X}\|_2 = \sqrt{N}\|X\|$, so

$$\begin{aligned}
\left\| \frac{\hat{c} * \hat{A}}{N} \right\| &= \|\mathcal{F}[\cdots c_j A_j \cdots]\|_2 \\
&= \sqrt{N}\|\cdots c_j A_j \cdots\|_2 \\
&\le \sqrt{N}\|c\|_2\|A\|_\infty = \|\hat{c}\|_2\|A\|_\infty.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|\hat{\mathcal{B}}^m\| &\le \left\| \frac{\hat{c} * \hat{B}^m}{N} \right\| \le \|\hat{c}\|_2\|B^m\|_\infty, \\
\|\hat{\mathcal{C}}^m\| &\le \left\| \frac{\hat{c} * \hat{C}^m}{N} \right\| \le \|\hat{c}\|_2\|C^m\|_\infty.
\end{aligned}$$

We thus have the following estimate on $T\hat{c}$:

$$\frac{\|T\hat{c}\|_2}{\|\hat{c}\|_2} \le \|A\|_\infty + \sum_{m=1}^{\infty} \frac{2}{\pi}\left(\|B^m\|_\infty + \|C^m\|_\infty\right). \tag{A.9}$$

From the definitions of A_j, B_j^m, C_j^m in (A.6), we have

$$\begin{cases}
\|A\|_\infty \le 1 - \dfrac{\sin(\pi L)}{\pi L}, \\
\|B^m\|_\infty \le \dfrac{L\sin(\pi L)}{m^2 - L^2}, \\
\|C^m\|_\infty \le \dfrac{L\cos(\pi L)}{(m-\frac{1}{2})^2 - L^2}.
\end{cases} \tag{A.10}$$

Combining (A.9) and (A.10), we get

$$\begin{aligned}
\frac{\|T\hat{c}\|_2}{\|\hat{c}\|_2} \le 1 - \frac{\sin(\pi L)}{\pi L} &+ \sum_{m=1}^{\infty} \frac{2L\sin(\pi L)}{\pi(m^2 - L^2)} \\
&+ \sum_{m=1}^{\infty} \frac{2L\cos(\pi L)}{\pi\left((m-\frac{1}{2})^2 - L^2\right)}.
\end{aligned}$$

The series on the right are known expansions:

$$\begin{cases} \displaystyle\sum_{m=1}^{\infty} \frac{2L}{\pi(m^2 - L^2)} = \frac{1}{\pi L} - \cot(L\pi), \\ \displaystyle\sum_{m=1}^{\infty} \frac{2L}{\pi\left((m-\frac{1}{2})^2 - L^2\right)} = \tan(L\pi). \end{cases}$$

Using these above gets us

$$\frac{\|T\hat{c}\|_2}{\|\hat{c}\|_2} \le 1 - \cos(\pi L) + \sin(\pi L) = 1 - \sqrt{2}\sin\left(\frac{\pi}{4} - L\pi\right). \tag{A.11}$$

We state this result as a lemma:

Lemma A.1. *If $L < \frac{1}{4}$, then*

$$\|T\|_{2,2} \le 1 - \sqrt{2}\sin\left(\frac{\pi}{4} - \pi L\right) < 1.$$

The point is that we can now invert the linear transformation in (A.7) and estimate the norm of $I - T$. When $L < \frac{1}{4}$, we have $\hat{c} = (I - T)^{-1}\tilde{c}$. Since $\|T\| < 1$, the standard Neumann expansion gives

$$\begin{aligned} \|\hat{c}\|_2 &\le \frac{\|\tilde{c}\|_2}{1 - \|T\|_{2,2}} \\ &\le (\sqrt{2}/2)\csc\left(\frac{\pi}{4} - \pi L\right)\|\tilde{c}\|_2, \end{aligned} \tag{A.12}$$

which leads to the following result:

Theorem A.2. *If $L < \frac{1}{4}$ and $\tilde{c} = (\,\tilde{c}_{-[N/2]} \quad \cdots \quad \tilde{c}_{[(N-1)/2]}\,)^T$, then for all $k \notin \mathcal{J}_N$*

$$|\tilde{c}_k| \le \sqrt{2}/2\csc\left(\frac{\pi}{4} - \pi L\right)\|\tilde{c}\|_2 \le \sqrt{N/2}\csc\left(\frac{\pi}{4} - \pi L\right)\sup_{k\in\mathcal{J}_N}|\tilde{c}_k|. \tag{A.13}$$

Proof. From (A.2), Schwarz's inequality, and $\|\hat{X}\|_2 = \sqrt{N}\|X\|$, we have that for all $k \notin \mathcal{J}_N$

$$|\tilde{c}_k| \le \sqrt{N}\|c\|_2 = \|\hat{c}\|_2.$$

Combining this inequality with (A.12) then yields the left inequality in (A.13). The right inequality follows from the observation that $\|\tilde{c}\|_2 \le \sqrt{N}\|\tilde{c}\|_\infty$. ■

Acknowledgments. The work of F. J. Narcowich was supported by Air Force Office of Scientific Research Grant F49620-95-1-0194, and that of J. D. Ward was supported by

National Science Foundation Grant DMS-9505460 and by Air Force Office of Scientific Research Grant F49620-95-1-0194.

References

1. P. H. BÉRARD (1980): Spectral Geometry: Direct and Inverse Problems. Berlin: Springer-Verlag.
2. YU. M. BEREZANSKII (1961): *A generalization of a multidimensional theorem of Bochner* (English). Soviet Math., **2**:143–147.
3. P. J. DAVIS (1975): Interpolation and Approximation. New York: Dover.
4. J. R. DRISCOLL, D. M. HEALY (1994): *Computing fourier transforms and convolutions for the* 2*-sphere*. Adv. in Appl. Math., **15**:202–250.
5. J. DUCHON (1976): *Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces*. RAIRO Anal. Numér., **10**:5–12.
6. J. DUCHON (1977): *Splines minimizing rotation invariant semi-norms in Sobolev spaces*. In: Constructive Theory of Functions of Several Variables. Oberwolfach 1976 (W. Schempp and K. Zeller, eds.). Berlin: Springer-Verlag, pp. 85–100.
7. W. FREEDEN (1981): *On spherical spline interpolation and approximation*. Math. Methods Appl. Sci., **3**:551–575.
8. F. G. FRIEDLANDER (1975): The Wave Equation on a Curved Space–Time. Cambridge: Cambridge University Press.
9. P. B. GILKEY (1974): The Index Theorem and the Heat Equation. Boston, MA: Publish or Perish.
10. M. GOLOMB, H. F. WEINBERG (1959): *Optimal approximation and error bounds*. In: On Numerical Approximation (R. E. Langer, ed.). Madison, pp. 117–190.
11. N. J. HICKS (1965): Notes on Differential Geometry. Princeton, NJ: D. van Nostrand.
12. M. G. KREIN (1946): *On a general method for decomposing positive-definite kernels into elementary products* (in Russian). Dokl. Akad. Nauk SSSR, **53**:3–6.
13. J. L. LIONS, E. MAGENES (1972): Non-Homogeneous Boundary Value Problems and Applications, Vol. I. New York: Springer-Verlag.
14. K. MAURIN (1967): Methods of Hilbert Spaces. New York: Hafner.
15. W. R. MADYCH, S. A. NELSON (1988): *Multivariate interpolation and conditionally positive definite functions*. Approx. Theory Appl., **4**:77–79.
16. W. R. MADYCH, S. A. NELSON (1990): *Multivariate interpolation and conditionally positive definite functions II*. Math. Comp., **54**: 211–230.
17. V. A. MENEGATTO (to appear): *Strictly positive definite kernels on the Hilbert sphere*. Appl. Anal.
18. C. MÜLLER (1966): Spherical Harmonics. Berlin: Springer-Verlag.
19. F. J. NARCOWICH (1995): *Generalized Hermite interpolation and positive definite kernels on a Riemannian manifold*. J. Math. Anal. Appl., **190**:165–193.
20. F. J. NARCOWICH, J. D. WARD (1994): *Generalized Hermite interpolation via matrix-valued conditionally positive definite kernels*. Math. Comp., **63**:661–688.
21. M. J. D. POWELL (1990): *The theory of radial basis approximation in* 1990. In: Wavelets, Subdivision and Radial Functions (W. Light, ed.). Oxford, UK: Oxford University Press.
22. A. RON, X. SUN (1996): *Strictly positive definite functions on spheres*. Math. Comp., **65**:1513–1530.
23. R. SCHABACK (1995): *Error estimates and condition numbers for radial basis function interpolation*. Adv. in Comput. Math., **3**:251–264.
24. I. J. SCHOENBERG (1942): *Positive definite functions on spheres*. Duke Math. J., **9**:96–108.
25. J. STEWART (1976): *Positive definite functions and generalizations, an historical survey*. Rocky Mountain J. Math., **6**:409–434.
26. G. WAHBA (1981): *Spline interpolation and smoothing on the sphere*. SIAM J. Sci. Statist. Comput., **2**:5–16.
27. G. WAHBA (1984): *Surface fitting with scattered noisy data on Euclidean d-space and on the sphere*. Rocky Mountain J. Math., **14**:281–299.
28. F. W. WARNER (1971): Foundations of Differentiable Manifolds and Lie Groups. Glenview, IL: Scott, Foresman.

29. Z. WU (1992): *Hermite–Birkhoff interpolation of scattered data by radial basis functions*. Approx. Theory Appl., **24**:201–215.
30. Z. WU, R. SCHABACK (1993): *Local error estimates for radial basis function interpolation of scattered data*. IMA J. Numer. Anal., **13**:13–27.
31. Y. XU, E. W. CHENEY (1992): *Strictly positive definite functions on spheres*. Proc. Amer. Math. Soc., **116**:977–981.
32. R. M. YOUNG (1980): An Introduction to Nonharmonic Fourier Series. New York: Academic Press.

N. Dyn
School of Mathematical Sciences
Tel-Aviv University
Tel-Aviv 69978
Israel

F. J. Narcowich
Department of Mathematics
Texas A&M University
College Station
Texas 77843
USA

J. D. Ward
Department of Mathematics
Texas A&M University
College Station
Texas 77843
USA