Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Approximation Theory 133 (2005) 269–283
Journal of Approximation Theory
www.elsevier.com/locate/jat

Approximation in rough native spaces by shifts of smooth kernels on spheres

J. Levesley[a], X. Sun[b,*,1]

[a] *Department of Mathematics, University of Leicester, University Road, Leicester LE1 7RH, UK*
[b] *Department of Mathematics, Southwest Missouri State University, Springfield, MO 65804, USA*
Received 20 February 2004; received in revised form 22 September 2004; accepted in revised form 8 December 2004

Dedicated to the memory of Will Light

Abstract

Within the conventional framework of a native space structure, a smooth kernel generates a small native space, and "radial basis functions" stemming from the smooth kernel are intended to approximate only functions from this small native space. Therefore their approximation power is quite limited. Recently, Narcowich et al. (J. Approx. Theory 114 (2002) 70), and Narcowich and Ward (SIAM J. Math. Anal., to appear), respectively, have studied two approaches that have led to the empowerment of smooth radial basis functions in a larger native space. In the approach of [NW], the radial basis function interpolates the target function at some scattered (prescribed) points. In both approaches, approximation power of the smooth radial basis functions is achieved by utilizing spherical polynomials of a (possibly) large degree to form an intermediate approximation between the radial basis approximation and the target function. In this paper, we take a new approach. We embed the smooth radial basis functions in a larger native space generated by a less smooth kernel, and use them to approximate functions from the larger native space. Among other results, we characterize the best approximant with respect to the metric of the larger native space to be the radial basis function that interpolates the target function on a set of finite scattered points after the action of a certain multiplier operator. We also establish the error bounds between the best approximant and the target function.

Keywords: Native space; Radial basis function; Reproducing kernels

* Corresponding author.
E-mail address: xis280f@smsu.edu (X. Sun).
[1] Partially supported by the Engineering and Physical Science Research Council (EPSRC) of Great Britain.

0021-9045/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jat.2004.12.005

1. Introduction

Reconstructing an unknown function from scattered data is both theoretically interesting and practically important. The tools utilized for the reconstruction have included traditionally polynomials and splines, and lately radial basis functions (RBFs). If the domains are the d-sphere S^d embedded in the $(d+1)$-dimensional Euclidean space R^{d+1}, then the term "radial basis function" often refers to a linear combination of spherical shifts of a fixed strictly positive definite zonal kernel ϕ. The kernel ϕ generates a reproducing kernel Hilbert space (RKHS) $\mathcal{N}_\phi$, called the native space of ϕ in the community of approximation theory. Let Ξ be a finite subset of S^d. We always assume in this paper that Ξ consists of distinct points, and we use $|\Xi|$ to denote the cardinality of Ξ. For each $f \in \mathcal{N}_\phi$, let $s_\phi[f]$ be the best approximant of f from the $|\Xi|$-dimensional subspace $\Phi_\Xi :=\text{span}\{\phi(\langle\cdot,\xi\rangle) : \xi \in \Xi\}$, where $\langle\cdot,\cdot\rangle$ denotes the usual inner product on R^{d+1}, and the best approximant is with respect to the metric of $\mathcal{N}_\phi$. Then $s_\phi[f]$ is uniquely determined by the interpolation condition:

$$s_\phi[f](\xi) = f(\xi), \quad \xi \in \Xi.$$

An optimal error estimate for $\|s_\phi[f] - f\|$ is also available, where $\|\cdot\|$ denotes the uniform norm. We refer the readers to [DNW;GL;JSW;LLRS;MN;Sc] for details on these standard results. The characterization of the best approximant and consequently the error estimate therein only apply to functions in the native space $\mathcal{N}_\phi$. If the kernel ϕ is smooth, the associated native space $\mathcal{N}_\phi$ is small in the sense that it is composed of smooth functions. Narcowich et al. [NSW] made this observation and proposed the following remedy: If f belongs to a larger Sobolev-type space W not contained in the native space $\mathcal{N}_\phi$, then they approximate f first by $s_L(f)$, its L-th partial sum of its Fourier series. They then approximate $s_L(f)$ by a radial basis function from Φ_Ξ. They obtained the same approximation order as if f is being approximated by radial basis functions stemming from a less smooth kernel whose native space contains W. Narcowich and Ward [NW] recently took a more elaborate approach in this direction. They first succeeded (remarkably) in finding a spherical harmonics $P_L \in \mathcal{H}_L$, the space of all spherical harmonics of degree L or less, such that

(i) P_L interpolates the target function f on Ξ,
(ii) $\|P_L - f\| \leqslant \text{const} \cdot \text{dist}(f, \mathcal{H}_L)$.

They then transfer the approximation power of P_L to a smooth radial basis function from Φ_Ξ. The overall approximation power is comparable to the optimal one. In both approaches, approximation power is achieved by utilizing spherical polynomials of a (possibly) large degree as an intermediate approximation between the radial basis approximant and the target function.

In this paper, we take a new approach. Suppose that ϕ and ψ are two zonal kernels, ϕ is smoother than ψ. Therefore $\mathcal{N}_\phi \subset \mathcal{N}_\psi$. We embed Φ_Ξ in $\mathcal{N}_\psi$ and approximate $f \in \mathcal{N}_\psi$ by Φ_Ξ in $\mathcal{N}_\psi$. We show (Theorem 3.3) that the best approximant $s_\phi[f]$ of f, with respect to the metric of $\mathcal{N}_\psi$, is characterized by the interpolation condition

$$T(s_\phi[f])|_\Xi = T(f)|_\Xi,$$

where T is a multiplier operator explicitly determined by ϕ and ψ. We also show that the error estimate for $\|s_\phi[f]-f\|$ is of the same order as $\|s_\psi[f]-f\|$, where $s_\psi[f]$ denotes the best approximant of f from the $|\Xi|$-dimensional space: $\text{span}\{\psi(\langle\cdot,\xi\rangle):\xi\in\Xi\}$. In obtaining the error estimate we will use the "norming set" method developed by Jetter et al. [JSW], and a general Bernstein-type inequality established by Ditzian [D]. We will also use a native space duality argument advocated by Morton and Neamtu [MN].

This approach can also be applied to the Euclidean space setting in which we approximate functions from a larger native space by smoother radial basis functions, such as the multiquadrics. We will discuss this issue in a forthcoming publication.

The current paper is arranged as follows. In Section 2, we discuss strictly positive definite kernels and strictly positive definite functions on S^d. In Section 3, we study the structure of a native space generated by a strictly positive definite zonal kernel, which culminates in the characterization of the best approximant $s_\phi[f]$ from Φ_Ξ for a function $f\in\mathcal{N}_\psi$. In Section 4, we obtain various estimates for $\|s_\phi[f]-f\|$.

2. Strictly Positive definite kernels and functions on spheres

For each $k=0,1,\ldots$, let $\mathcal{H}_k^{(0)}$ be the linear space of homogeneous harmonic polynomials of degree k, and

$$\mathcal{H}_L := \bigoplus_{k\leqslant L}\mathcal{H}_k^{(0)}.$$

The dimension of $\mathcal{H}_k^{(0)}$ is denoted by d_k. It is well known that

$$d_0=1,\quad d_k=\frac{(2k+d-1)\Gamma(k+d-1)}{\Gamma(k+1)\Gamma(d)}$$

and that $\mathcal{H}_k^{(0)}=\Pi_k\cap\Pi_{k-1}^{\perp}$, where Π_k denotes the linear space of homogeneous polynomials of degree k. For each $k=0,1,\ldots$, let $\{Y_{k,j}:j=1,\ldots,d_k\}$ be an orthonormal basis of $\mathcal{H}_k^{(0)}$ with respect to the standard inner product

$$\langle p,q\rangle=\int_{S^d}p(x)q(x)\,d\mu(x),\quad p,q\in\mathcal{H}_k^{(0)},$$

where $d\mu$ is the restriction on S^d of the Lebesgue measure in R^{d+1} whose total mass is denoted by $w\omega_d$, i.e., $\omega_d=\int_{S^d}d\mu(x)$.

The set

$$\{Y_{k,\mu}:\mu=1,\ldots,d_k,\ k=0,1,\ldots,\}$$

forms an orthonormal basis for $L^2(S^d)$. Thus, the collection

$$\{Y_{k,\mu}\cdot Y_{l,v}:\mu,v=1,\ldots,d_k,\ k,l=0,1,\ldots,\}$$

is an orthonormal basis for $L^2(S^d\times S^d)$. As usual, we will call a function defined on $S^d\times S^d$ a kernel on S^d, or just a kernel. A function $\mathcal{K}\in C(S^d\times S^d)$ is called a positive definite, if for any finite subset of S^d, and arbitrary real numbers c_ξ, $\xi \in \Xi$, we have

$$\sum_{\xi\in\Xi}\sum_{\zeta\in\Xi} c_\xi c_\zeta \mathcal{K}(\xi,\zeta) \geqslant 0. \tag{2.1}$$

A rather general result of Bochner [B] implies that the totality of all positive definite kernels $\mathcal{K}$ may be represented by the expression

$$\mathcal{K}(x,y) = \sum_{k=0}^{\infty}\sum_{\mu=1}^{d_k}\sum_{\nu=1}^{d_k} a_{k,\mu,\nu} Y_{k,\mu}(x) Y_{k,\nu}(y), \tag{2.2}$$

in which, for each fixed k, the $d_k \times d_k$ matrix $(a_{k,\mu,\nu})$, is nonnegative definite, and

$$\sum_{k=0}^{\infty} d_k \sum_{\mu=1}^{d_k}\sum_{\nu=1}^{d_k} |a_{k,\mu,\nu}| < \infty.$$

It is not difficult to see that a kernel $\mathcal{K}$ as expressed in Eq. (2.2) is positive definite. A straightforward calculation shows that

$$\sum_{\xi\in\Xi}\sum_{\zeta\in\Xi} c_\xi c_\zeta \mathcal{K}(\xi,\zeta) = \sum_{k=0}^{\infty}\sum_{\mu=1}^{d_k}\sum_{\nu=1}^{d_k} a_{k,\mu,\nu}\left(\sum_{\xi\in\Xi} c_\xi Y_{k,\mu}(\xi)\right)\left(\sum_{\xi\in\Xi} c_\xi Y_{k,\nu}(\xi)\right).$$

For each $k = 0, 1, \ldots$, the quadratic form

$$\sum_{\mu=1}^{d_k}\sum_{\nu=1}^{d_k} a_{k,\mu,\nu}\left(\sum_{\xi\in\Xi} c_\xi Y_{k,\mu}(\xi)\right)\left(\sum_{\xi\in\Xi} c_\xi Y_{k,\nu}(\xi)\right)$$

is nonnegative because of the nonnegative definiteness of the matrix $(a_{k,\mu,\nu})$.

We will use $\langle x, y\rangle$ or xy, depending on the mathematical context, to denote the usual inner product of x and y in R^d. Let x and y be two points on S^d. The geodesic distance between x and y is denoted by $g(x,y)$, and we have $xy = \cos(g(x,y))$. A kernel $\mathcal{K}$ is called rotational invariant, if

$$\mathcal{K}(\rho x, \rho y) = \mathcal{K}(x,y) \quad \text{for all } x, y \in S^d \quad \text{and for all rotations } \rho.$$

It can be shown that a continuous rotational invariant kernel depends only on the distance between x and y, that is, there is a function $\phi : [-1,1] \to R$, such that $\phi(xy) = \mathcal{K}(x,y)$ for all $x, y \in S^d$; see [SW, Chapter IV]. Therefore a rotational invariant kernel is also called a zonal kernel in the literature. The study of the zonal kernels is also facilitated by the famous summation formula:

$$C_k^{(\lambda)}(xy) = \frac{\omega_d}{d_k}\sum_{\mu=1}^{d_k} Y_{k\mu}(x) Y_{k\mu}(y),$$

where $\lambda = (d-2)/2$, and $C_k^{(\lambda)}$ is the order λ Gegenbauer polynomials of degree k, normalized so that $C_k^{(\lambda)}(1) = 1$; see [Sz]. Hence, a continuous zonal kernel enjoys the simpler expansion:

$$\mathcal{K}(x, y) = \phi(xy) \sim \sum_{k=0}^{\infty} A_k \sum_{\mu=1}^{d_k} C_k^{(\lambda)}(xy).$$

The summation formula also yields the following useful relation:

$$\sum_{\mu=1}^{d_k} |Y_{k\mu}(x) Y_{k\mu}(y)| \leqslant \sum_{\mu=1}^{d_k} Y_{k\mu}^2(x) = \frac{d_k}{\omega_d}, \quad x, y \in S^d. \tag{2.3}$$

Schoenberg [S] defined a continuous function $\phi : [-1, 1] \to R$ to be "positive definite" on S^d if the kernel $\mathcal{K}(x, y) := \phi(xy)$ is positive definite. In the same paper, Schoenberg established the following remarkable result: in order that ϕ be positive definite on S^d , it is necessary and sufficient that

$$\phi(t) = \sum_{k=0}^{\infty} A_k C_k^{(\lambda)}(t), \quad t \in [-1, 1], \tag{2.4}$$

in which $A_k \geqslant 0$, for all $k = 0, 1, \ldots$, and $\sum_{k=0}^{\infty} A_k < \infty$.

We note that the positive definiteness of a function ϕ amounts to the positive definiteness of the kernel $\phi(xy)$. Therefore, Schoenberg's result is a special case of the afore-mentioned result of Bochner. However, Schoenberg's result came first even though his paper appeared one year later than Bochner's; see [B].

If the quadratic form in Eq. (2.1) is positive whenever the c_ξ are not all zero, then the kernel $\mathcal{K}(x, y)$ is called strictly positive definite. In the zonal kernel case, the univariate function ϕ is called strictly positive definite on S^d. This notation of strictly positive definiteness on spheres was introduced by Xu and Cheney [XC] with the motivation of interpolating arbitrary data on a finite subset Ξ of S^d by a unique function in the span$\{\phi(\langle \cdot, \xi \rangle) : \xi \in \Xi\}$. It is important to characterize all the strictly positive definite functions on spheres. Such an endeavor has been taken by authors in, among others, [XC,M1,M2,RS,CMS,Su]. A characterization is recently given by Chen et al. [CMS] for strictly positive definite functions on S^d $(d \geqslant 2)$. Their result asserts that in order that the zonal kernels as expressed in Eq. (2.4) be strictly positive definite on S^d $(d \geqslant 2)$, it is necessary and sufficient that A_k be positive for infinitely many odd k's and infinitely many even k's. The problem is still open for the case $d = 1$. Some substantial recent progress is reported in [Su].

In applications, however, the most useful strictly positive kernels appear to be the ones given by Xu and Cheney [XC] which require that all expansion coefficients A_k in Eq. (2.4) be positive. We will refer them as Xu–Cheney kernels in this paper. The importance of the Xu–Cheney kernels are reinforced by the following two results:

1. A Xu–Cheney kernel $\phi(xy)$ enjoys the stronger strict positive definiteness advocated by Narcowich [N]. Namely, for any function $f \in C(S^d)$, not identically zero, we have

$$\int_{S^d} \int_{S^d} \phi(xy) f(x) f(y) \, d\mu(x) \, d\mu(y) > 0.$$

We caution here that there exist functions that are strictly positive definite in the conventional sense but fail to satisfy this stronger requirement; see [RS].

2. Span$\{\phi(\xi\cdot), \xi \in S^d\}$ is dense in $C(S^d)$; see [SC,RL].

In the next two sections, we will use Xu–Cheney kernels to set up native spaces and study approximations in these spaces.

3. Best approximation in native spaces

Despite the simpler form as in Eq. (2.4), we choose to write a Xu–Cheney kernel $\phi(xy)$ in the following form for the convenience of our presentation:

$$\phi(xy) = \sum_{k=0}^{\infty} a_k \sum_{\mu=1}^{d_k} Y_{k,\mu}(x) Y_{k,\mu}(y), \tag{3.1}$$

where $a_k > 0$, for all $k = 0, 1, 2, \ldots$, and $\sum_{k=0}^{\infty} a_k d_k < \infty$. Of course, the coefficient a_k in Eq. (3.1) and the coefficient A_k in Eq. (2.4) have the relation $A_k = a_k \frac{d_k}{\omega_d}$. We consider the linear space consisting of all the finite linear combinations of zonal shifts of ϕ. Denote this linear space by PH_ϕ, and define the following bilinear form on PH_ϕ:

$$\langle f, g \rangle_{PH_\phi} := \sum_{\xi \in \Xi} \sum_{\zeta \in \Theta} c_\xi d_\zeta \phi(\xi\zeta),$$

where $f = \sum_{\xi \in \Xi} c_\xi \phi(\xi\cdot)$, and $g = \sum_{\zeta \in \Theta} d_\zeta \phi(\zeta\cdot)$, and both Ξ and Θ are finite subsets of S^d. It is easy to verify that this is an inner product on PH_ϕ thanks to the strict positive definiteness of ϕ.

The native space $\mathcal{N}_\phi$, associated with ϕ, is the subspace of $L_2(S^d)$ defined by

$$\mathcal{N}_\phi := \left\{ f = \sum_{k=0}^{\infty} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu} Y_{k,\mu}(\cdot) : \sum_{k=0}^{\infty} a_k^{-1} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu}^2 < \infty \right\},$$

where $\hat{f}_{k,\mu}$ is defined by

$$\hat{f}_{k,\mu} := \int_{S^d} f(x) Y_{k,\mu}(x) \, d\mu(x).$$

The native space $\mathcal{N}_\phi$ is a Hilbert space with the inner product:

$$\langle f, g \rangle_{\mathcal{N}_\phi} = \sum_{k=0}^{\infty} a_k^{-1} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu} \hat{g}_{k,\mu}.$$

Proposition 3.1. *The native space $\mathcal{N}_\phi$ is a reproducing kernel Hilbert space, and the reproducing kernel is $\phi(xy)$.*

Proof. We already know that $\mathcal{N}_\phi$ is a Hilbert space. So we just need to check that $\phi(xy)$ is the reproducing kernel. Take $f \in \mathcal{N}_\phi$, and write

$$f = \sum_{k=0}^{\infty} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu} Y_{k,\mu} \quad \text{with} \quad \sum_{k=0}^{\infty} a_k^{-1} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu}^2 < \infty.$$

For a fixed $x \in S^d$, we have

$$\langle f(\cdot), \phi(x\cdot)\rangle_{\mathcal{N}_\phi} = \sum_{k=0}^{\infty} a_k^{-1} \sum_{\mu=1}^{d_k} a_k \hat{f}_{k,\mu} Y_{k,\mu}(x) = f(x). \qquad \square$$

We can now reveal the relationship between $\mathcal{N}_\phi$ and PH_ϕ.

Proposition 3.2. *The native space $\mathcal{N}_\phi$ is the completion of PH_ϕ.*

Proof. It is obvious that $PH_\phi \subset \mathcal{N}_\phi$. For $f \in PH_\phi$, we write

$$f(\eta) = \sum_{\xi \in \Xi} c_\xi \phi(\xi\eta) = \sum_{k=0}^{\infty} a_k \sum_{\mu=1}^{d_k} \left(\sum_{\xi \in \Xi} c_\xi Y_{k,\mu}(\xi) \right) Y_{k,\mu}(\eta).$$

We then have

$$\| f \|_{PH_\phi}^2 = \sum_{k=0}^{\infty} a_k \sum_{\mu=1}^{d_k} \left(\sum_{\xi \in \Xi} c_\xi Y_{k,\mu}(x) \right)^2 = \| f \|_{\mathcal{N}_\phi}^2.$$

To complete the proof, we need to show that PH_ϕ is dense in $\mathcal{N}_\phi$. Suppose that $g \in \mathcal{N}_\phi$, and $\langle f, g\rangle_{\mathcal{N}_\phi} = 0$ for all $f \in PH_\phi$. In particular, $\langle \phi(\cdot x), g(\cdot)\rangle_{\mathcal{N}_\phi} = 0$ for each fixed $x \in S^d$. But

$$\langle \phi(\cdot x), g(\cdot)\rangle_{\mathcal{N}_\phi} = g(x),$$

because ϕ is the reproducing kernel of $\mathcal{N}_\phi$. Thus $g(x) = 0$ for all $x \in S^d$. The desired density result thus follows from the Hilbert space version of the Riesz representation theorem. $\square$

Let $\phi(xy)$, $\psi(xy)$ be two Xu–Cheney kernels. Write

$$\phi(xy) = \sum_{k=0}^{\infty} a_k \sum_{\mu=1}^{d_k} Y_{k,\mu}(x) Y_{k,\mu}(y),$$

$$\psi(xy) = \sum_{k=0}^{\infty} b_k \sum_{\mu=1}^{d_k} Y_{k,\mu}(x) Y_{k,\mu}(y).$$

Assuming that $a_k \leqslant b_k$ for all $k = 0, 1, \ldots$, we then see that $\mathcal{N}_\phi$ is a subset of $\mathcal{N}_\psi$. However, the embedding operator $f \mapsto f$ from $\mathcal{N}_\phi$ to $\mathcal{N}_\psi$ is generally not isometric. We define a multiplier operator T from $\mathcal{N}_\psi$ to $\mathcal{N}_\phi$ by the following rule:

$$Tf = \sum_{k=0}^{\infty} \frac{a_k}{b_k} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu} Y_{k,\mu}.$$

Clearly, T turns into the identity operator if and only if $a_k = b_k$ for all k. We intend to approximate an arbitrary function $f \in \mathcal{N}_\psi$ by the $|\Xi|$-dimensional subspace $\Phi_\Xi := \operatorname{span}\{\phi(\cdot\xi), \xi \in \Xi\}$. We use $s_\phi[f]$ to denote the best approximant of f from Φ_Ξ with respect to the metric of $\mathcal{N}_\psi$, and we use P_ϕ to denote the projection operator that maps $f \in \mathcal{N}_\psi$ to $s_\phi[f]$. We give a characterization for $s_\phi[f]$ in terms of a target function f and the multiplier operator T.

Theorem 3.3. *Let $f \in \mathcal{N}_\psi$, and let $s_\phi[f]$ be the unique best approximant of f from Φ_Ξ with respect to the metric of $\mathcal{N}_\psi$. Write*

$$s_\phi[f] = \sum_{\xi\in\Xi}^{N} c_\xi \phi(\cdot\xi).$$

Then the coefficients c_ξ, $\xi \in \Xi$, are determined by the interpolation conditions

$$T(s_\phi[f])|_\Xi = T(f)|_\Xi.$$

Proof. Since both the multiplier operator T and the projection operator P_ϕ are linear and bounded from $\mathcal{N}_\psi$ to $\mathcal{N}_\phi$, in view of Proposition 3.2, we may assume that f is of the form: $f(\cdot) = \psi(x\cdot)$ for a fixed x. Note that $s_\phi[f]$ is the best approximant to f if and only if $s_\phi[f] = P_\phi f$, which is equivalent to

$$(f - s_\phi[f]) \perp \phi(\cdot\xi) \quad \text{for all } \xi \in \Xi.$$

That is

$$\langle s_\phi[f], \phi(\cdot\xi)\rangle_{\mathcal{N}_\psi} = \langle f, \phi(\cdot\xi)\rangle_{\mathcal{N}_\psi} \quad \text{for all } \xi \in \Xi. \tag{3.2}$$

We calculate

$$\begin{aligned}\langle s_\phi[f], \phi(\cdot\xi)\rangle_{\mathcal{N}_\psi} &= \sum_{\zeta\in\Xi} c_\zeta \sum_{k=0}^{\infty} a_k \frac{a_k}{b_k} \sum_{\mu=1}^{d_k} Y_{k,\mu}(\zeta) Y_{k,\mu}(\xi)\\ &= \left.\left(\sum_{\zeta\in\Xi} c_\zeta T(\phi(\cdot\zeta)\right)\right|_\xi\\ &= T(s_\phi[f])|_\xi) \end{aligned} \tag{3.3}$$

and

$$\langle f, \phi(\cdot\xi)\rangle_{\mathcal{N}_\psi} = \sum_{k=0}^{\infty} b_k \frac{a_k}{b_k} \sum_{\mu=1}^{d_k} Y_{k,\mu}(x) Y_{k,\mu}(\xi) = T(f)|_\xi. \tag{3.4}$$

Substituting Eqs. (3.3) and (3.4) into Eq. (3.2), we have $T(s_\phi[f])|_\Xi = T(f)|_\Xi$. $\square$

4. Error estimates

The reproducing kernel $\psi(xy)$ plays an important role in the approximation of functions in $\mathcal{N}_\psi$ by functions in Φ_Ξ. We first fix an $x \in S^d$, and investigate how well the function $\psi_x : y \mapsto \psi(xy)$ can be approximated by functions in Φ_Ξ in the metric of the native space $\mathcal{N}_\psi$. The mesh norm h of Ξ defined by

$$h := \sup_{x\in S^d} \min_{\xi\in\Xi} d(x, \xi),$$

where $d(x, \xi)$ denotes the geodesic distance between x and ξ, is an important gauge for the approximation. In what follows, we take the liberty of using C to denote a constant whose exact value may be different from line to line.

To establish our error estimates, we will need the following two important results.

Lemma 4.1. *Let M be a multiplier operator defined on $\mathcal{H}_L$ (embedded in $C(S^d)$) by*

$$M(p) = \sum_{k=0}^{L} \sum_{\mu=1}^{d_k} m_{k,\mu} a_{k,\mu} Y_{k,\mu},$$

where

$$p = \sum_{k=0}^{L} \sum_{\mu=1}^{d_k} a_{k,\mu} Y_{k,\mu}.$$

Then

$$\|M(p)\| \leqslant C \left(\max_{k,\mu} m_{k,\mu} \right) \|p\|,$$

where C is a constant independent of f and L.

Lemma 4.1 is a special case of a rather general Bernstein-type inequality established by Ditzian [D, Theorem 3.2].

Lemma 4.2. *Let $\Xi \subset S^d$ be a finite knot set with mesh norm $h(\Xi) \leqslant 1/(2L)$. Then for any linear functional σ on $\mathcal{H}_L$ (embedded in $C(S^d)$), $\|\sigma\|_* = 1$, there exist $|\Xi|$ real numbers α_ξ, $\xi \in \Xi$ with $\sum_{\xi\in\Xi} |\alpha_\xi| \leqslant 2$, so that*

$$\sigma(f) = \sum_{\xi\in\Xi} \alpha_\xi \delta_\xi(f),$$

for all $f \in \mathcal{H}_L$, where δ_ξ denotes the point evaluation functional at the point ξ in Ξ.

Lemma 4.2 is due to Jetter et al. [JSW]. The result is the corner stone for the so-called "norming set" approach that has quickly become a standard tool for authors in native space approximation.

Theorem 4.3. *Let $\Xi \subset S^d$ be a finite set with mesh norm $h(\Xi) \leqslant 1/(2L)$. Assume that the sequence b_k/a_k is monotone increasing. Then for each fixed $x \in S^d$, we have*

$$\|\psi_x - s_\phi[\psi_x]\|_{\mathcal{N}_\psi} \leqslant C \left(\sum_{k>L}^{\infty} b_k d_k \right)^{1/2}.$$

Here the constant C is independent of f and x.

Before we embark on the proof of the theorem, we make a comment on the assumption that the sequence b_k/a_k be monotone increasing. This assumption is not necessary, and can be relaxed considerably. However, imposing the condition reduces some inessential details of the proof which would otherwise be too long. After all, the primary goal in this paper is to expand the approximation power of the Ξ-dimensional space Φ_Ξ, which can be achieved by conveniently choosing $b_k = k^m a_k$ for a proper natural number m. With this choice, the sequence b_k/a_k is monotone increasing.

Proof of Theorem 4.3. It suffices to show that there exists an $s \in \Phi_\Xi$, $s = \sum_{\xi \in \Xi} \beta_\xi \phi(\xi \cdot)$, such that

$$\|s - \psi_x\|_{\mathcal{N}_\psi} \leqslant C \left(\sum_{k>L}^{\infty} b_k d_k \right)^{1/2}.$$

We have

$$\begin{aligned} \|s - \psi_x\|_{\mathcal{N}_\psi} &= \sup_{\substack{v \in \mathcal{N}_\psi \\ \|v\|_{\mathcal{N}_\psi}=1}} \langle s - \psi_x, v \rangle_{\mathcal{N}_\psi} \\ &= \sup_{\substack{v \in \mathcal{N}_\psi \\ \|v\|_{\mathcal{N}_\psi}=1}} \sum_{k=0}^{\infty} b_k^{-1} \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} \left(a_k \sum_{\xi \in \Xi} \beta_\xi Y_{k,\mu}(\xi) - b_k Y_{k,\mu}(x) \right) \\ &= \sup_{\substack{v \in \mathcal{N}_\psi \\ \|v\|_{\mathcal{N}_\psi}=1}} \sum_{k=0}^{\infty} a_k b_k^{-1} \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} \left(\sum_{\xi \in \Xi} \beta_\xi Y_{k,\mu}(\xi) - b_k a_k^{-1} Y_{k,\mu}(x) \right). \end{aligned}$$

Let T_L be the multiplier operator defined on $\mathcal{H}_L$ (embedded in $C(S^d)$) by

$$T_L(Y_{k,\mu}) = \frac{b_k}{a_k} Y_{k,\mu}$$

for each $k = 0, 1, \ldots, L$, and all $\mu = 1, 2, \ldots, d_k$, and extended linearly throughout $\mathcal{H}_L$. Let σ be the linear functional on $\mathcal{H}_L$ defined by: $\sigma = \delta_x \circ T_L$. That is $\sigma(p) = (T_L(p))\,(x)$

for each $p \in \mathcal{H}_L$. By Lemma 4.1 and the assumption that the sequence b_k/a_k is monotone increasing, we have

$$|\sigma(p)| = |\,(T_L(p))\,(x)| \leqslant \|T_L(p)\| \leqslant C \max_{0 \leqslant k \leqslant L} \frac{b_k}{a_k} \|p\| = C\,\frac{b_L}{a_L} \|p\|,$$

in which C is a constant independent of p and L. By Lemma 4.2, there exist $|\Xi|$ real numbers β_ξ, $\xi \in \Xi$ with $\sum_{\xi\in\Xi} |\beta_\xi| \leqslant 2C\frac{b_L}{a_L}$ such that

$$\sum_{\xi\in\Xi} \beta_\xi Y_{k,\mu}(\xi) = \frac{b_k}{a_k}\, Y_{k,\mu}(x), \qquad Y_{k,\mu} \in \mathcal{H}_L.$$

With β_ξ thus chosen, we have

$$\begin{aligned}
\|s - \psi_x\|_{\mathcal{N}_\psi} &= \sup_{\substack{v\in\mathcal{N}_\psi\\ \|v\|_{\mathcal{N}_\psi}=1}} \sum_{k>L} a_k b_k^{-1} \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} \left(\sum_{\xi\in\Xi} \beta_\xi Y_{k,\mu}(\xi) - b_k a_k^{-1} Y_{k,\mu}(x) \right)\\
&= \sup_{\substack{v\in\mathcal{N}_\psi\\ \|v\|_{\mathcal{N}_\psi}=1}} \left[\sum_{\xi\in\Xi} \beta_\xi \sum_{k>L} a_k b_k^{-1} \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} Y_{k,\mu}(\xi) - \sum_{k>L} \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} Y_{k,\mu}(x) \right]\\
&\leqslant \sup_{\substack{v\in\mathcal{N}_\psi\\ \|v\|_{\mathcal{N}_\psi}=1}} \left[\sum_{\xi\in\Xi} |\beta_\xi| \sum_{k>L} a_k b_k^{-1} \left| \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} Y_{k,\mu}(\xi) \right| \right.\\
&\qquad \left. + \left| \sum_{k>L} \sum_{\mu=1}^{d_k} \hat{v}_{k,\mu} Y_{k,\mu}(x) \right| \right].
\end{aligned}$$

We use the Cauchy–Schwartz Inequality and the relation in (2.3) to estimate the sums to get

$$\begin{aligned}
\|s - \psi_x\|_{\mathcal{N}_\psi} \leqslant{}& \left(\sum_{\xi\in\Xi} |\beta_\xi| \right) \cdot \max_{\xi\in\Xi} \left(\sum_{k>L} \frac{a_k^2}{b_k} \sum_{\mu=1}^{d_k} Y_{k,\mu}^2(\xi) \right)^{1/2}\\
&\cdot \sup_{\substack{v\in\mathcal{N}_\psi\\ \|v\|_{\mathcal{N}_\psi}=1}} \left(\sum_{k>L} \sum_{\mu=1}^{d_k} \frac{\hat{v}_{k,\mu}^2}{b_k} \right)^{1/2} + \left(\sum_{k>L} b_k \sum_{\mu=1}^{d_k} Y_{k,\mu}^2(x) \right)^{1/2}\\
&\cdot \sup_{\substack{v\in\mathcal{N}_\psi\\ \|v\|_{\mathcal{N}_\psi}=1}} \left(\sum_{k>L} \sum_{\mu=1}^{d_k} \frac{\hat{v}_{k,\mu}^2}{b_k} \right)^{1/2}
\end{aligned}$$

$$\leqslant 2C\,\frac{b_L}{a_L}\left(\sum_{k>L}\frac{a_k^2}{b_k}\,\frac{d_k}{\omega_d}\right)^{1/2}+\left(\sum_{k>L}b_k\,\frac{d_k}{\omega_d}\right)^{1/2}$$
$$\leqslant (2C+1)\left(\sum_{k>L}b_k\,\frac{d_k}{\omega_d}\right)^{1/2}.$$

The last inequality is true because b_k/a_k is a monotone increasing sequence. □

Remark 4.4. The duality argument we used here is adapted from that of Proposition 10 by Morton and Neamtu [MN]. If $b_k = a_k$ for all $k = 0, 1, 2, \ldots$, then $C = 1$, and Theorem 2.6 reduces itself to Proposition 10 in [MN].

Corollary 4.5. *Let $\Xi \subset S^d$ be a finite set with mesh norm $h(\Xi)\leqslant 1/(2L)$. Then for every $f \in \mathcal{N}_\psi$, we have*

$$\|f - s_\phi[f]\|\leqslant C\|f - s_\phi[f]\|_{\mathcal{N}_\psi}\left(\sum_{k\geqslant L}^{\infty} b_k d_k\right)^{1/2}.$$

Here the constant C is independent of f.

Proof. For each fixed $x \in S^d$, we have

$$\begin{aligned}
&|f(x) - s_\phi[f](x)|\\
&\quad = |\langle f - s_\phi[f], \psi_x\rangle_{\mathcal{N}_\psi}|\\
&\quad = |\langle f - s_\phi[f], \psi_x - s_\phi[\psi_x]\rangle_{\mathcal{N}_\psi}|\\
&\quad \leqslant \|f - s_\phi[f]\|_{\mathcal{N}_\psi}\|\psi_x - s_\phi[\psi_x]\|_{\mathcal{N}_\psi}\\
&\quad \leqslant C\|f - s_\phi[f]\|_{\mathcal{N}_\psi}\left(\sum_{k\geqslant L}^{\infty} b_k d_k\right)^{1/2}.
\end{aligned}$$

In the second line of the above argument, we used the fact that $\psi(xy)$ is the reproducing kernel of the native space $\mathcal{N}_\psi$. In the third line, we used the orthogonality $(f - s_\phi[f]) \perp \Phi_\Xi$. In the fourth line, we used the Cauchy–Schwartz Inequality. In the fifth line, we used Theorem 4.3. □

Remark 4.6. The factor $\|f - s_\phi[f]\|_{\mathcal{N}_\psi}$ in Corollary 4.5 is hard to estimate, and therefore the obvious inequality

$$\|f - s_\phi[f]\|_{\mathcal{N}_\psi}\leqslant\|f\|_{\mathcal{N}_\psi}$$

is often applied to yield the following "cleaner" error estimate in Corollary 4.5:

$$\|f - s_\phi[f]\|\leqslant C\|f\|_{\mathcal{N}_\psi}\sum_{k\geqslant L}^{\infty} b_k d_k.$$

In the special case $b_k^2 = a_k$ and $f \in \mathcal{N}_\phi$, a so-called "error doubling" technique has been used to derive an estimate for $\|f - s_\phi[f]\|_{\mathcal{N}_\psi}$. The technique was first explored by Schaback [Sc] in the Euclidean space setting, and has been subsequently adapted in spherical domains in [LLRS,GL,MN], among other publications. Note that $b_k^2 = a_k$ is equivalent to $\phi = \psi * \psi$, where

$$(\psi * \psi)(x, y) = \int_{S^d} \psi(xz)\psi(zy)\, d\mu(z).$$

In Proposition 4.7 below, we modify the proof of Proposition 12 by Morton and Neamtu [MN] to obtain a slightly more general result.

Proposition 4.7. *Let $\Xi \subset S^d$ be a finite knot set with mesh norm $h(\Xi) \leqslant 1/(2L)$. Suppose that there is a $C > 0$ such that $a_k/b_k^2 \leqslant C$. Then for each $f \in \mathcal{N}_\phi$, we have*

$$\|f - s_\phi[f]\|_{\mathcal{N}_\psi} \leqslant C \left(\sum_{k>L}^{\infty} b_k d_k \right)^{1/2} \|f\|_{\mathcal{N}_\phi}.$$

Proof. To avoid awkward writing, we use the abbreviation $\hat{s}_{k,\mu}$ to denote the (k, μ)-coefficient of the Fourier series expansion of $s_\phi[f]$, that is,

$$s_\phi[f] = \sum_{k=0}^{\infty} \sum_{\mu=1}^{d_k} \hat{s}_{k,\mu} Y_{k,\mu}.$$

We use the Cauchy–Schwartz inequality to deduce that

$$\begin{aligned} \|f - s_\phi[f]\|^2_{\mathcal{N}_\psi} &= \langle f, f - s_\phi[f] \rangle_{\mathcal{N}_\psi} \\ &= \sum_{k=0}^{\infty} b_k^{-1} \sum_{\mu=1}^{d_k} \hat{f}_{k,\mu} \, (\hat{f}_{k,\mu} - \hat{s}_{k,\mu}) \\ &\leqslant \left(\sum_{k=0}^{\infty} a_k^{-1} \sum_{\mu=1}^{d_k} (\hat{f}_{k,\mu})^2 \right)^{1/2} \left(\sum_{k=0}^{\infty} \frac{a_k}{b_k^2} \sum_{\mu=1}^{d_k} (\hat{f}_{k,\mu} - \hat{s}_{k,\mu})^2 \right)^{1/2} \\ &= C \|f\|_{\mathcal{N}_\phi} \|f - s_\phi[f]\|_2. \end{aligned} \tag{4.1}$$

Here $\|f - s_\phi[f]\|_2$ denotes the regular L_2 norm of the function $f - s_\phi[f]$. Hölder's inequality yields

$$\|f - s_\phi[f]\|_2 \leqslant \omega_d^{1/2} \|f - s_\phi[f]\| \tag{4.2}$$

and Corollary 4.5 yields

$$\|f - s_\phi[f]\| \leqslant C \omega_d^{1/2} \|f - s_\phi[f]\|_{\mathcal{N}_\psi} \left(\sum_{k>L}^{\infty} b_k d_k \right)^{1/2}. \tag{4.3}$$

Connecting (4.1)–(4.3), we obtain

$$\|f - s_\phi[f]\|^2_{\mathcal{N}_\psi} \leqslant C\|f - s_\phi[f]\|_{\mathcal{N}_\psi} \left(\sum_{k>L}^{\infty} b_k d_k\right)^{1/2} \|f\|_{\mathcal{N}_\phi}. \tag{4.4}$$

If $\|f - s_\phi[f]\|_{\mathcal{N}_\psi} = 0$, then the result is automatically true. Otherwise we divide by $\|f - s_\phi[f]\|_{\mathcal{N}_\psi}$ on both sides of Inequality (4.4) to get the desired result. □

On some occasions, it is desirable to express the error estimate in terms of h. This can be done by making h and $1/L$ compatible in the sense that $h = \mathcal{O}(1/L)$. In Corollary 4.8 that follows, we deal with an important special case.

Proposition 4.8. *Let $f \in C^{2l}(S^d)$, and let $s_\phi[f]$ be the best approximant of f from Φ_Ξ as characterized by Theorem* 3.3. *Assume that $a_k \leqslant Ck^{-2\tau}$ for some $\tau \geqslant 2l > d/2$, where C is a constant independent of k. Then we have the estimate*

$$\|f - s_\phi[f]\| \leqslant Ch^{2l-d/2}\|f - s_\phi[f]\|_{\mathcal{N}_\phi},$$

where C is a constant independent of f and Ξ.

Proof. Let $f \in C^{2l}(S^d)$. Then by Inequality (2.7) in [NW], we have $f \in \mathcal{N}_\psi$ with

$$\psi(xy) = \sum_{k=0}^{\infty} k^{-4l} \sum_{\mu=1}^{d_k} Y_{k,\mu}(x)Y_{k,\mu}(y).$$

If $a_k \leqslant ck^{-2\tau}$, then

$$\Phi_\Xi \subset \mathcal{N}_\psi.$$

By Corollary 4.5, we have

$$\|f - s_\phi[f]\| \leqslant C\|f - s_\phi[f]\|_{\mathcal{N}_\psi} \left(\sum_{k>L}^{\infty} k^{-4l} d_k\right)^{1/2}.$$

Using the relation $d_k \sim k^{d-1}$, we get the estimate

$$\left(\sum_{k>L}^{\infty} k^{-4l} d_k\right)^{1/2} \sim (1/L)^{2l-d/2} \sim h^{2l-d/2}.$$

The desired result then follows. □

Remark 4.9. The order of approximation given in Proposition 4.8 is the same as in Proposition 3.1 in [NSW] and Corollary 3.3 in [NW]. We remind readers that the approximating functions used in the three occasions are all different. In particular, ours are characterized by Theorem 3.3 in this paper.

Acknowledgments

The second author is partially supported by a grant from The Engineering and Physical Sciences Research Council (EPSRC) of Great Britain. He thanks EPSRC for their support and the hospitality of University of Leicester during his 2003 summer visit. The referee offered several suggestions that has improved the exposition of the paper for which both authors are grateful.

References

[B] S. Bochner, Hilbert distances and positive definite functions, Ann. of Math. 42 (1941) 647–656.

[CMS] D. Chen, V.A. Menegatto, X. Sun, A necessary and sufficient condition for strictly positive definite functions on spheres, Proc. Amer. Math. Soc. 131 (2003) 2733–2740.

[D] Z. Ditzian, Fractional derivatives and best approximation, Acta. Math. Hungar. 81 (4) (1998) 323–348.

[DNW] N. Dyn, F. Narcowich, J. Ward, Variational principles and Sobolev-type estimates for generalized interpolation on a Riemannian manifold, Constr. Approx. 15 (1999) 175–208.

[GL] M. von Golitschek, W. Light, Interpolation by polynomials and radial basis functions on spheres, Constr. Approx. 17 (2001) 1–18.

[JSW] K. Jetter, J. Stöckler, J.D. Ward, Error estimates for scattered data interpolations on spheres, Math. Comp. 68 (1999) 733–747.

[LLRS] J. Levesley, W. A. Light, D. Rogozin, X. Sun, A simple approach to variational theory for interpolation on spheres, International Theory of Numerical Analysis, vol. 132, Birkhauser, Verlag Base/Switerzland, 1999.

[M1] V.A. Menegatto, Strictly positive definite kernels on the Hilbert spheres, Appl. Anal. 55 (1994) 114–119.

[M2] V.A. Menegatto, Interpolation on the complex Hilbert sphere, Approx. Theory Appl. 11 (1995) 1–9.

[N] F.J. Narcowich, Generalized Hermite interpolation and positive definite kernels on a Riemannian manifold, J. Math. Anal. Appl. 190 (1995) 165–193.

[NSW] F.J. Narcowich, R. Schaback, J.D. Ward, Approximation in Sobolev spaces by kernel expansions, J. Approx. Theory 114 (2002) 70–83.

[NW] F.J. Narcowich, J.D. Ward, Scattered data interpolation on sphere: error estimates and locally supported basis functions, SIAM J. Math. Anal. 36 (2004) 284–300.

[RL] D. Ragozin, J. Levesley, The fundamentality of translates of spherical functions on compact homogeneous spaces, J. Approx. Theory 103 (2000) 252–268.

[RS] A. Ron, X. Sun, Strictly positive definite functions on spheres, Math. Comp. 65 (1996) 1513–1530.

[S] I.J. Schoenberg, Positive definite functions on spheres, Duke Math. J. 9 (1942) 96–108.

[Sc] R. Schaback, Improved error bounds for scattered interpolation by radial basis functions, Math. Comp. 68 (1999) 201–216.

[Su] X. Sun, Strictly positive definite functions on the unit circle, Math. Comp. 74 (2005) 709–721.

[SC] X. Sun, E.W. Cheney, Fundamental sets of continuous functions on spheres, Constr. Approx. 13 (1997) 245–250.

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

[Sz] G. Szegő, Orthogonal Polynomials, fourth ed., American Mathematical Society Colloquium Publication, vol. 23, American Mathematical Society, Providence, RI, 1959.

[MN] Tanya M. Morton, M. Neamtu, Error bounds for solving pseudo-differential equations on spheres by collocation with zonal kernels, J. Approx. Theory 114 (2002) 242–268.

[XC] Y. Xu, E.W. Cheney, Strictly positive definite functions on spheres, Proc. Amer. Math. Soc. 116 (1992) 977–981.