Multilevel Interpolation and Approximation

F.J. Narcowich[1], R. Schaback[2] and J.D. Ward[1]

1. Center for Approximation Theory, Texas A&M University, College Station, TX, 77843-3368, U.S.A.

The work of F.J. Narcowich was supported by Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number 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, Air Force Material Command, USAF, under grant number F49620-95-1-0194. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

2. Universität Göttingen, Lotzestrasse 16-18, D-37083 Göttingen, Germany

Proposed running head: *Multilevel Interpolation and Approximation*

Send all correspondence to:

Francis J. Narcowich
Department of Mathematics
Texas A&M University
College station, TX 77843-3368
Phone: (409) 845-7369
FAX: (409) 845-6028
E-mail `fnarc@math.tamu.edu`

Abstract

Interpolation by translates of a given radial basis function (RBF) has become a well-recognized means of fitting functions sampled at scattered sites in R^d. A major drawback of these methods is their inability to interpolate very large data sets in a numerically stable way while maintaining a good fit. To circumvent this problem, a multilevel interpolation (ML) method for scattered data was presented by Floater and Iske. Their approach involves m levels of interpolation where at the j^{th} level, the residual of the previous level is interpolated. On each level, the RBF is scaled to match the data density. In this paper, we provide some theoretical underpinnings to the ML method by establishing rates of approximation for a technique that deviates somewhat from the Floater-Iske setting. The final goal of the ML method will be to provide a numerically stable method for interpolating several thousand points rapidly.

1 Introduction

1.1 Background

Interpolation by translates of a given radial basis function (RBF) has become a well-recognized means of fitting functions sampled at scattered sites in R^d. The initial theoretical groundwork was done by Duchon [2, 3] and later Micchelli [10] and Madych/Nelson [8, 9]. In these papers, the basic interpolation method by RBF's was introduced, rates of approximation were given for a class of smooth functions and the invertibility of the interpolation matrices was established.

Nevertheless, there were primarily two drawbacks to these methods. First, the RBF's used were globally supported, so the interpolation matrices were full and inversion was slow. For the most part, this problem has been overcome with the advent of a class of "designer" RBF's by Wu and Wendland [19, 20]. This class consists of functions that are positive definite on a given Euclidean space R^d and that are compactly supported, easy to evaluate, and belong to a prescribed smoothness class. Of course, the interpolation matrices corresponding to this class of functions are sparse.

The second drawback was the inability of RBF's to interpolate large data sets (several thousand points) in a numerically stable way while maintaining good reproduction of the original function. This problem is most easily understood in terms of the "Uncertainty Principle" described in [16], which establishes a trade-off between the smoothness of an RBF, which implies better reproduction of a function and faster rates of approximation, and the numerical stability of the interpolation matrices. In short, one has to sacrifice good reproduction quality for good stability or vice versa.

In an effort to circumvent this problem, a multilevel interpolation method for scattered data was presented by Floater and Iske in [5, 6]. A full description of this method will be presented in §2, but briefly this approach involves m levels of interpolation where at the j^{th} level, the residual of the previous level is interpolated. On each level, the radial basis function is scaled to match the data density. Thus, at the lowest level, interpolation of coarse data by a function of relatively large support is done to capture the broad details of the given function. At the highest level, one interpolates very dense data with a narrow peak function to insure numerical stability. Floater and Iske [5] showed numerically that one could interpolate several thousand points in this fashion. However, no theoretical estimates for rates of approximation using this method were given, and an analysis of their method poses severe problems. This may be due to the fact that they used radial basis functions with the same smoothness at each level.

In this paper, we provide some theoretical underpinnings to the multilevel interpolation method (ML) by establishing rates of approximation for a technique that deviates somewhat from the original Floater-Iske setting in the sense that it uses smoother functions on coarse data than on fine data. The final goal of the ML method (using compactly supported RBF's) will be to provide a numerically stable method for interpolating several thousand points with rapid approximation.

In the next section, the ML method is fully described and a motivating example from spline theory is given. Then §3 provides a framework for obtaining recursive Jackson bounds, while in §4 error bounds for multilevel interpolation are obtained for the case of the torus, the circle, the sphere, and compact sets in Euclidean space.

2 Multilevel Approximation

In this section, we will discuss the notion of multilevel approximation. The method may be described as a process where one works on several levels, each time treating the residuals of the previous level. Since the technique is not confined to interpolation by radial basis functions, we describe it here in full generality. At the same time, we present an overview of our proof technique.

2.1 Formulation of the Multilevel Technique

For a normed linear space $\mathcal{W}$ with a closed linear subspace W, the determination of a best approximation $f^* \in W$ to a given $f \in \mathcal{W}$ consists of solving the minimization problem

$$\inf_{g \in W} \|f - g\|_{\mathcal{W}} = \|f - f^*\|_{\mathcal{W}}$$

for the norm $\|f - g\|_{\mathcal{W}}$ of the residual function $f - g$. To assess the possible quality of the result, one is interested in *a-priori* bounds for the optimal error $\|f - f^*\|_{\mathcal{W}}$. Since $g = 0$ is admissible, there always is the primitive error bound

$$\|f - f^*\|_{\mathcal{W}} \leq \|f\|_{\mathcal{W}}. \tag{2.1}$$

Unfortunately, this bound is useless in practice, because it in no way involves the space of approximants W. Moreover in most cases, without additional information on f itself, it cannot be improved.

To obtain a useful error bound, one must restrict f to a proper subspace $\mathcal{W}_0$ of $\mathcal{W}$, where $\mathcal{W}_0$ carries a norm that provides one with information that can be used to get improved error bounds. Such bounds have the general form

$$\|f - f^*\|_{\mathcal{W}} \leq K(\mathcal{W}, \mathcal{W}_0, W)\|f\|_{\mathcal{W}_0} \tag{2.2}$$

with a constant $K(\mathcal{W}, \mathcal{W}_0, W)$ that is significantly smaller than 1 and depends in a highly nontrivial way on the subspace W that furnishes the admissible approximations to f. For historical reasons we shall call such an error estimate a *Jackson bound.*

The multilevel approximation method can be motivated as follows. Suppose that we want to work in the setting above, but are not satisfied with the error predicted by (2.2), and futhermore, suppose that we do not want to change $\mathcal{W}, \mathcal{W}_0, f$ or W. There are of course many reasons for not wishing to change these things; here we just list a few possibilities:

1. The subspace W cannot be enlarged without exceeding computational limitations or because it would require new and unavailable data.
2. $\mathcal{W}$ cannot be made larger (or $\|\cdot\|_{\mathcal{W}}$ made less restrictive) due to the inherent assumptions on the application problem (e.g. the required smoothness of approximants or the required type of norm).
3. $\mathcal{W}_0$ cannot be made smaller (or $\|\cdot\|_{\mathcal{W}_0}$ made more restrictive) because there is saturation (no gain by picking smaller spaces $\mathcal{W}_0$, as is the case, for example, for univariate splines of fixed degree), or because there is no suitable subspace of $\mathcal{W}_0$ that contains f.

To escape this dilemma, one can employ multilevel approximation processes. These use spaces intermediate between $\mathcal{W}$ and $\mathcal{W}_0$, if all of the conditions mentioned above apply. To describe the technique, let us assume that there is a sequence of nested spaces

$$\mathcal{W}_0 \subset \mathcal{W}_1 \subset \mathcal{W}_2 \cdots \subset \mathcal{W}_m = \mathcal{W}, \tag{2.3}$$

that connects the $\mathcal{W}$ and $\mathcal{W}_0$ of (2.2). In each of the spaces $\mathcal{W}_k$ we pose an approximation problem

$$\inf_{g \in W_k} \|f_k - g\|_{\mathcal{W}_k} = \|f_k - f_k^*\|_{\mathcal{W}_k}$$

to approximate an element $f_k \in \mathcal{W}_k$ by elements from a closed subspace $W_k \subset \mathcal{W}_k$. The function f_k, however, will always be the *residual* of the previous step, i.e.,

$$f_k := f_{k-1} - f_{k-1}^* \in \mathcal{W}_{k-1}, \qquad 2 \leq k \leq m, f_1 := f.$$

The final approximation after step m to the first input $f = f_1$ will be

$$g_m^* := \sum_{k=1}^{m} f_k^*,$$

and it will be from the space

$$V_m := \sum_{k=1}^{m} W_k,$$

where the sum need not be direct or orthogonal. Note that the spaces V_m are nested as in a multiresolution analysis, but the spaces W_m may not (and probably won't) be orthogonal.

Furthermore, some situations do not require intermediate spaces, but rather an extension of the inclusion $\mathcal{W}_0 \subset \mathcal{W}$ to the left or right. Extension to the left, for instance, will always occur if the given function f can be assumed to lie in a much smaller subspace that $\mathcal{W}_0$.

2.2 Error Bounds

The intermediate spaces of (2.3) should allow a sequence of recursive Jackson bounds

$$\|f_k - f_k^*\|_{\mathcal{W}_k} \leq K_k \|f_k\|_{\mathcal{W}_{k-1}}, \qquad 1 \leq k \leq m \tag{2.4}$$

where we use the abbreviated notation

$$K_k := K(\mathcal{W}_k, \mathcal{W}_{k-1}, W_k).$$

If the subspaces W_k are sufficiently large, each of these constants

will be significantly less than one. The error bounds (2.4) can be applied recursively with the final result taking the form

$$\|f - g_m^*\|_{\mathcal{W}} = \|f_1 - g_m^*\|_{\mathcal{W}_m} \leq \left(\prod_{k=1}^{m} K_k\right) \|f_1\|_{\mathcal{W}_0} \tag{2.5}$$

which replaces (2.2), and where all the single Jackson constants of (2.4) are *multiplied.*

The multilevel error bound (2.5) usually involves a lot more information than any single-level bound (2.4), and the method itself uses more degrees of freedom for approximation. A fair comparison would thus contrast (2.5) with the Jackson bound

$$\|f_1 - f_m^{**}\|_{\mathcal{W}_m} \leq K(\mathcal{W}_m, \mathcal{W}_0, V_m)\|f_1\|_{\mathcal{W}_0}, \tag{2.6}$$

where f_m^{**} is a best single-level approximation to f_1 from the space V_m in the norm of $\mathcal{W}_m$. Note that this problem will have a solution that differs in general from the solution of the multilevel method, since the intermediate norms differ. Furthermore, the above problem may not be solvable at all because of computational limitations. In such cases the multilevel method would still be preferable, since it allows one to break an unwieldy large-scale approximation problem into manageable chunks. As mentioned earlier, this work originated from promising computational results of Floater and Iske [5, 6] concerning reconstruction of multivariate functions from very large sets of scattered data.

If we ignore the above motivation based on numerical stability, and concentrate on theoretical error bounds, the multistep method would only be an improvement over the single step method if the relevant Jackson constants satisfy

$$K(\mathcal{W}_m, \mathcal{W}_0, V_m) > \prod_{k=1}^{m} K(\mathcal{W}_k, \mathcal{W}_{k-1}, V_k).$$

In section §4.4 we shall discuss the validity of this inequality for the multilevel method on the circle. But at least in situations where the single-step bound (2.6) is weak due to saturation, while the bounds (2.4) avoid saturation, the above inequality will hold and make the multilevel method superior. A simple example will be given in the next subsection. On the downside, there will be no improvement except for a possible savings in computational complexity, if all norms stem from a single inner product and if the spaces W_k form an orthogonal decomposition of V_m. But this case, which is typical of standard wavelet or Fourier series expansions, is not covered by this paper, because the norms involved in a Jackson bound will usually differ substantially.

2.3 Example

We end this section with an example showing how the multilevel procedure works and illustrating its potential advantages over the single-step method. The example will model a case where we assume the final level (the level utilizing all the interpolation points) is forced to use a simple and efficient method due to computational restrictions. However the resulting error estimates will then be weak. Thus we take, for illustration purposes, interpolation of functions f on [0,1] by splines $f^*_{1,h}$ of degree one on a mesh of width h. The standard Jackson bound then is [1, p. 40]

$$\|f - f^*_{1,h}\|_\infty \leq \frac{h^2}{8}\|f''\|_\infty, \tag{2.7}$$

and this is saturated in the sense that even for much smoother functions f there will be no improvement over the $\mathcal{O}(h^2)$ accuracy. A very similar situation occurs if one interpolates a very large set of multivariate scattered data. The sheer size of the problem will require either a solution without solving a system of equations, or at least a massively sparse system of equations. In both cases one must be satisfied with a weak error bound, if sticking to a single-step method.

Suppose, in the above example, we assume that f is much smoother than C^2. The Jackson bound (2.7) then suggests to prefix a first step that has a Jackson bound on second derivatives by higher derivatives. If, for instance, the function f is in $C^4[0,1]$ with $f''(0) = f''(1) = 0$, then cubic spline interpolation will serve the purpose, the required bound being

$$\|f'' - (f^*_{3,h})''\|_\infty \leq ch^2\|f^{(4)}\|_\infty.$$

If one interpolates with cubic splines at the coarse level and then interpolates the residual by piecewise linear splines at the fine level, the multistep approach will now yield an explicit $\mathcal{O}(h^4)$ bound

$$\|f - f^*_{3,Kh} - (f - f^*_{3,Kh})^*_{1,h}\|_\infty \leq \frac{cK^2}{8}h^4\|f^{(4)}\|_\infty,$$

even if the cubic spline is calculated on a very coarse mesh of width Kh. This yields an improvement over (2.7) as soon as

$$cK^2h^2\|f^{(4)}\|_\infty \leq \|f''\|_\infty$$

holds, e.g.for h sufficiently small.

For examples of the good performance of a variation of the multistep method using compactly supported radial basis functions applied to large multivariate problems we refer the reader to the papers by Floater and Iske [5, 6]. This should suffice as a practical motivation for our work. Since a theoretical basis for multilevel approximation is missing, we thus concentrate on theoretical questions.

3 Radial and Related Basis Functions

We now discuss the spaces and classes of functions that we will deal with in this paper. We confine ourselves to methods using (not necessarily radial) basis functions, and we develop a notation for transforms that allows us to deal with functions on Euclidean space, torus, and sphere in a unified way.

3.1 Transforms and Positive Definite Functions

Let $\mathcal{M}$ be a metric space. We say that a conjugate symmetric kernel $\kappa \in C(\mathcal{M}\times\mathcal{M})$ is positive definite [18], if the $N\times N$ selfadjoint matrix with entries $\kappa(p_j, p_k)$ is positive semidefinite for any arbitrary finite subset $\{p_1, \dots, p_N\}$ of $\mathcal{M}$ with distinct points.

Three important choices for $\mathcal{M}$ are n-dimensional Euclidean space, $\mathbf{R}^n$, the n-torus, $\mathbf{T}^n$, and the n-sphere, $\mathbf{S}^n$. On each of these we can define positive definite kernels via *positive definite functions*; such functions were introduced long ago by Bochner and Schoenberg (cf. [18]).

The first two choices are similar, because both allow a group of translations to act on $\mathcal{M}$, making harmonic analysis by Fourier transforms possible. Suppose that $\Phi : \mathbf{R}^n \to \mathbf{C}$ is in $C(\mathbf{R}^n) \cap L^1(\mathbf{R}^n)$. From Bochner's Theorem, we know that Φ is positive definite if and only if its Fourier transform $\widehat{\Phi}(\xi)$ is nonnegative. Similarly, if $\Phi : \mathbf{T}^n \to \mathbf{C}$ is in $C(\mathbf{T}^n) \cap L^1(\mathbf{T}^n)$, and if the Fourier coefficients of Φ, which we denote by $\widehat{\Phi}(\alpha)$, with $\alpha \in \mathbf{Z}^n$ being a multi-index, are nonnegative, then Φ will be positive definite on $\mathbf{T}^n$. In both cases, the positive semidefinite kernel associated with Φ is $\kappa(p, q) := \Phi(p-q)$.

The case of $\mathbf{S}^n$ is somewhat different, because now the orthogonal group acts on $\mathcal{M}$, replacing standard harmonic analysis by expansions into spherical harmonics. Using ultraspherical polynomials, Schoenberg [17] gave a representation for the continuous positive definite functions on $\mathbf{S}^n$. Because the Legendre polynomials $P_\ell(n+1;x)$ are somewhat easier to use than the ultraspherical polynomials, and are, up to constant multiples, the same set, we will use them to give Schoenberg's representation theorem [17], eq. (1.5):

$$\Phi(\cos(\theta)) = \sum_{\ell=0}^{\infty} \widehat{\Phi}(\ell) P_\ell(n+1;\cos(\theta)), \tag{3.1}$$

where the $\widehat{\Phi}(\ell)$'s are nonnegative and decay sufficiently fast for $\Phi(\cos(\theta))$ to be continuous. In this case, the positive semidefinite kernel associated with Φ is $\kappa(p,q) = \Phi(p\cdot q)$, where $p\cdot q$ is the usual Euclidean dot product in $\mathbf{R}^{n+1}$. Note that points in $\mathbf{S}^n$ may be regarded as vectors of unit length in $\mathbf{R}^{n+1}$, such that the angle θ between two points p and q satisfies $p\cdot q = \cos\theta$. The resulting functions of the form $\Phi(\cos(\theta)) = \kappa(p,q)$ are called *zonal*.

It is useful to expand $\Phi(p\cdot q)$ in terms of the spherical harmonics Y_j^ℓ on $\mathbf{S}^n$ (cf. [11, 12]). This we can do by employing the famous Addition Theorem for spherical harmonics [11]:

$$P_\ell(n+1;p\cdot q) = \frac{\omega_n}{d_n(\ell)} \sum_{j=1}^{d_n(\ell)} Y_j^\ell(p)\bar{Y}_j^\ell(q), \tag{3.2}$$

Here, $d_n(\ell)$ is the dimension of the space of $n+1$ dimensional harmonic polynomials homogeneous of degree ℓ and ω_n is the volume of $\mathbf{S}^n$. This results in the expansion

$$\Phi(p\cdot q) = \sum_{\ell=0}^{\infty} \sum_{j=1}^{d_n(\ell)} \widehat{\Phi}(\ell,j) Y_j^\ell(p)\bar{Y}_j^\ell(q), \text{ where } \widehat{\Phi}(\ell,j) := \widehat{\Phi}(\ell)\frac{\omega_n}{d_n(\ell)}. \tag{3.3}$$

3.2 Dual Pairs of Sobolev and Native Spaces

Positive definite functions may be used to define positive semi-definite Hermitian forms on distributions. We first describe this in the case of $\mathcal{M} = \mathbf{R}^n$. Recall that if u is a compactly supported distribution on $\mathbf{R}^n$, then it has a well-defined analytic Fourier transform $\hat{u}(\xi)$ such that the action of u on a sufficiently smooth and absolutely integrable function f can be written as

$$u(f) = \int_{\mathbf{R}^n} \hat{u}(\omega)\hat{f}(\omega)d\omega.$$

The same identity holds for pairs (u,f) of tempered distributions and functions in the sense of L. Schwartz. In general, the above identity serves to define dual pairs of spaces depending on the behavior of the respective Fourier transforms of u and f. This is done by introducing a nonnegative weight function $\rho(\omega)$ such that $\rho(\omega)$ penalizes $\hat{u}(\omega)$ while $(\rho(\omega))^{-1}$ penalizes $\hat{f}(\omega)$. The most popular case of this technique takes the weight function $\rho(\omega) = (1 + |\omega|^2)^s$ and leads to Sobolev spaces $H_s(\mathbf{R}^n)$, where the inner product

$$\llbracket u, v \rrbracket_s := \int_{\mathbf{R}^n} (1 + |\omega|^2)^s \hat{u}(\omega)\overline{\hat{v}}(\omega) d\omega \tag{3.4}$$

is used to define the dual pair of spaces

$$\begin{aligned} H_s(\mathbf{R}^n) &:= \{ f \ : \ \llbracket f \rrbracket_s < \infty \} \\ H_{-s}(\mathbf{R}^n) &:= \{ u \ : \ \llbracket u \rrbracket_{-s} < \infty \} \\ &= H_s(\mathbf{R}^n)^*. \end{aligned}$$

Upon completion, these form a pair of mutually dual Hilbert spaces.

We now bring positive definite functions Φ into play and define an associated dual pair of "native" spaces. If Φ is a continuous, even, and absolutely integrable function such that $\hat{\Phi}$ is nonnegative on $\mathbf{R}^n$, then the function Φ and the associated kernel $\kappa(p, q) := \Phi(p - q)$ are positive definite and one can use $\hat{\Phi}$ as a weight function to define a semidefinite inner product of two finitely supported distributions u and v by

$$\llbracket u, v \rrbracket_\Phi := u^p v^q \Phi(p - q) = u^p v^q \kappa(p, q) = \int_{\mathbf{R}^n} \hat{u}(\omega)\overline{\hat{v}}(\omega)\hat{\Phi}(\omega) d\omega, \tag{3.5}$$

where u^p means the action of u with respect to the variable p. The final form of this expression and the correspondence to the Sobolev case lead us to introduce a similar semidefinite inner product of two sufficiently smooth and absolutely integrable functions f and g by

$$\llbracket f, g \rrbracket_\Phi := \int_{\mathbf{R}^n} \hat{f}(\omega)\overline{\hat{g}}(\omega) \left(\hat{\Phi}(\omega)\right)^{-1} d\omega. \tag{3.6}$$

Since we carefully distinguish between functions $f, g, \ldots$ and distributions $u, v, w, \ldots$ (as functionals acting on functions), there will be no confusion between these two inner products. Note that the above integral requires the Fourier transforms of f and g to vanish wherever the Fourier transform of Φ vanishes, but a quick look at the classical Whittaker–Shannon sampling theorem corresponding to the choice $\Phi(x - y) = \operatorname{sinc}(x - y)$ shows that this requirement is quite natural for recovery questions, restricting the recoverable functions to those with the proper bandwidth. If the function $\hat{\Phi}$ is positive almost everywhere in $\mathbf{R}^n$, the above inner products are positive definite. Upon Hilbert space completion, we then get the dual pair of "native" Hilbert spaces

$$\begin{aligned} H_\Phi &:= \{ f \ : [\![f]\!]_\Phi < \infty \} \\ H_\Phi^* &:= \{ u \ : [\![u]\!]_\Phi < \infty \}. \end{aligned}$$

As in the Sobolev case, it does not matter much from which (sufficiently rich) starting point the completion is done; starting from tempered functions and distributions is sufficient.

We now want to relate Sobolev spaces $H_s(\mathbf{R}^n)$ to the somewhat unwieldy native spaces H_Φ. To this end, we define

$$\|\Phi\|_{s,\infty} := \sup_{\omega \in \mathbf{R}^n} (1 + |\omega|^2))^s |\hat{\Phi}(\omega)| \tag{3.7}$$

and easily calculate the following bound for the bilinear form (3.6)

Proposition 3.1 *If Φ is a continuous, positive definite function on $\mathbf{R}^n$ for which $\|\Phi\|_{s,\infty}$ is finite, then the bilinear form (3.5) is continuous on $H_{-s}(\mathbf{R}^n)$ and satisfies the bound*

$$|[\![u, v]\!]_\Phi| \le \|u\|_{-s} \|v\|_{-s} \|\Phi\|_{s,\infty} \tag{3.8}$$

for all u and v in $H_{-s}(\mathbf{R}^n)$. The associated spaces have continuous embeddings

$$H_{-s}(\mathbf{R}^n) \subseteq H_\Phi^*, \quad H_\Phi \subseteq H_s(\mathbf{R}^n). \tag{3.9}$$

Note that for all of the well–known positive definite RBF's there is some s such that Proposition 3.1 holds. Gaussians and inverse multiquadrics will allow any nonnegative s.

3.3 Generalization to the Torus and the Sphere

If u is a distribution on $\mathbf{S}^n$ or $\mathbf{T}^n$, then u belongs to the Sobolev space $H_s(\mathbf{S}^n)$ if

$$\|u\|_s^2 := \sum_{\ell=0}^{\infty} \sum_{j=1}^{d_n(\ell)} (1 + \ell(\ell + n - 1))^s |\hat{u}(\ell, j)|^2$$

and to $H_s(\mathbf{T}^n)$ if

$$\|u\|_s^2 := \sum_{\alpha\in\mathbf{Z}^n} (1+|\alpha|^2)^s |\hat{u}(\alpha)|^2/(2\pi)^n < \infty.$$

Here, $\hat{u}(\ell, j)$ is the Fourier coefficient of u relative to the orthonormal basis of spherical harmonics, and $\hat{u}(\alpha)$ is the corresponding quantity relative to $\{e^{i\alpha\cdot\theta}\}$, the usual orthogonal basis for $L^2(\mathbf{T}^n)$.

The formulas above make use of eigenfunction transforms for the Laplace-Beltrami operator associated with the underlying manifold [12]. Many of the results that we state can be conveniently described by using common notation for these eigenfunction transforms. To that end, we will denote the appropriate index set by $\mathcal{O}$, the index by ω, the measure appropriate to $\mathcal{O}$ (discrete or continuous, as needed) by $d\nu(\omega)$, and we will let

$$\lambda(\omega) := \left\{ \begin{array}{ll} |\xi|^2 & \omega = \xi,\ (\mathbf{R}^n) \\ |\alpha|^2 & \omega = \alpha,\ (\mathbf{T}^n) \\ \ell(\ell+n-1) & \omega = (\ell, j),\ (\mathbf{S}^n) \end{array} \right\}. \tag{3.10}$$

This allows us to write the Sobolev norms as

$$\|u\|_s^2 = \int_{\mathcal{O}} (1+\lambda(\omega))^s |\hat{u}(\omega)|^2 d\nu(\omega). \tag{3.11}$$

In addition to the Sobolev Hilbert space norms above, we need the corresponding seminorms introduced by positive definite functions Φ. These are defined similarly to (3.5) and (3.6) with just a generalized notion of transform:

$$\begin{array}{rcl} [\![u, v]\!]_\Phi & := & \displaystyle\int_{\mathcal{O}} \hat{u}(\omega)\overline{\hat{v}(\omega)}\widehat{\Phi}(\omega) d\nu(\omega) \\ [\![f, g]\!]_\Phi & := & \displaystyle\int_{\mathcal{O}} \hat{f}(\omega)\overline{\hat{g}(\omega)}\left(\widehat{\Phi}(\omega)\right)^{-1} d\nu(\omega). \end{array} \tag{3.12}$$

We can then generalize (3.7) to

$$\|\Phi\|_{s,\infty} := \sup_{\omega\in\mathcal{O}} (1+\lambda(\omega))^s |\widehat{\Phi}(\omega)| \tag{3.13}$$

and get a straightforward generalization of Proposition 3.1 to the torus and the sphere. We include it into the formulation of Proposition 3.2 below.

3.4 Strictly Positive Definite Functions

The condition that $\widehat{\Phi}$ be positive almost everywhere is very important for solving recovery problems. It not only guarantees that the first of the Hermitian forms in (3.12) is actually an inner product on H_Φ^* and its subspaces $H_{-t} \subseteq H_{-s} \subseteq H_\Phi^*$ for all $t \le s$ provided that (3.13) is finite, but it also allows for the solution of generalized Hermite interpolation problems [4, 12, 13], the interpolants being of the form

$$K_\Phi[v](p) := v^q \kappa_\Phi(p, q), \tag{3.14}$$

with v being a distribution acting on the second argument of Φ. The following result, which is proved for compact manifolds in [4], and which is well known for $\mathbf{R}^n$ kernels, tells us which spaces are involved.

Proposition 3.2 *Let Φ be a positive definite function satisfying $\|\Phi\|_{s,\infty} < \infty$. Then for all $t \le s$ the operator K_Φ maps $H_{-t}(\mathcal{M})$ boundedly into $H_{2s-t}(\mathcal{M})$, with the norm being $\|\Phi\|_{s,\infty}$. Furthermore, we have*

$$|[\![u, v]\!]_\Phi| \le \|u\|_{-t} \|v\|_{-t} \|\Phi\|_{s,\infty} \tag{3.15}$$

for all u and v in $H_{-t}(\mathcal{M})$, and the associated spaces have continuous embeddings

$$H_{-t}(\mathcal{M}) \subseteq H_{-s}(\mathcal{M}) \subseteq H_\Phi^*, \quad H_\Phi \subseteq H_s(\mathcal{M}) \subseteq H_t(\mathcal{M}). \tag{3.16}$$

Let $\widehat{\Phi}$ be strictly positive and let $s \ge t$. For data generated by applying distributions from $\mathcal{U} = \mathrm{Span}\{u_1, \ldots, u_N\} \subset H_{-t}(\mathcal{M})$ to some unknown function f, we can always find $v^* \in \mathcal{U}$ for which $f^* = K_\Phi[v^*]$ such that the generalized interpolation conditions

$$u_j(f) = u_j(f^*), \quad 1 \le j \le N \tag{3.17}$$

hold, because the interpolation matrix is the $N \times N$ Gramian $[\![u_j\, u_k]\!]$, which is of course positive definite and therefore invertible.

Positive definite functions Φ for which $\widehat{\Phi}$ is strictly positive will be termed *strictly positive definite*. In the case of $\mathbf{R}^n$, if Φ is strictly positive definite and depends only on $|x|$, Φ is called a *radial basis function* (RBF). Strictly positive definite functions for $\mathbf{T}^n$ or $\mathbf{S}^n$ will be called *periodic basis functions* (PBFs) or *spherical basis functions* (SBFs), respectively.

The kernels induced by the Φ in $\mathbf{R}^n$ and $\mathbf{T}^n$ are of convolution type. This is also true for the case of the n-sphere, although there the convolution involved is less familiar and requires group theory to adequately describe. For our purposes, we need only note that if we formally define

$$f * g := \sum_{\ell=0}^{\infty} \sum_{j=1}^{d_n(\ell)} \hat{f}(\ell, j)\hat{g}(\ell, j) Y_j^\ell$$

for two arbitrary functions (or distributions) f and g, then fixing p and using equations (3.2) and (3.3), one gets

$$v^q \Phi(p \cdot q) = (\Phi * v)(p). \tag{3.18}$$

For the rest of the paper, we shall use the convolution form of the kernel for all three cases. In addition, we shall assume that $\widehat{\Phi} > 0$ and that $\|\Phi\|_{s,\infty}$ is finite.

Remark 3.3 There are two observations that we want to make. *First, if Φ_1 and Φ_2 are strictly positive definite functions on $\mathcal{M}$, then so is $\Phi_1 * \Phi_2$. Second, if $\widehat{\Phi}_1 \le c\widehat{\Phi}_2$, with c some positive constant, then $H_{\Phi_1} \subseteq H_{\Phi_2}$.* Note that the function Ψ_s with $\hat{\Psi}_s(\omega) = (1 + \lambda(\omega))^{-s}$ is positive definite, and thus (3.9) is a special case of the second observation, where $\Phi_1 = \Phi,\ \Phi_2 = \Psi_s$.

3.5 Standard Error Bounds

Let us return to the discussion of interpolation using functions constructed by applying functionals from a finite dimensional subspace $\mathcal{U} \subset H_{-s}(\mathcal{M}) \subseteq H_\Phi^*$ to Φ. If the function generating the data via application of the distributions in $\mathcal{U}$ is in the space H_Φ, so that it has the form $f = f_v = K_\Phi[v] = \Phi * v$ for some $v \in H_\Phi^*$, then we can adopt the terminology of (3.17) to write the generalized interpolant as $f^* = K_\Phi[v^*] = \Phi * v^*$ for some $v^* \in H_\Phi^*$. Standard arguments then show that $v - v^*$ must be orthogonal to $\mathcal{U}$ in the inner product $[\![\cdot,\cdot]\!]_\Phi$. Thus v^* is the orthogonal projection of v onto $\mathcal{U}$ and is automatically a best approximation to v from $\mathcal{U}$. Moreover, if w is in H_Φ^* and u in $\mathcal{U}$, then this orthogonality implies that $[\![v - v^*, w]\!] = [\![v - v^*, w - u]\!]$. These observations lead to a simple (and standard) bound for the action of $\bar{w}$ on $f_v - f_{v^*}$:

$$\begin{aligned} |\overline{w}(f_v - f_{v^*})| &= |[\![v - v^*, w]\!]_\Phi| \\ &= |[\![v - v^*, w - u]\!]_\Phi| \\ &\le [\![w - u]\!]_\Phi [\![v - v^*]\!]_\Phi\,. \end{aligned} \tag{3.19}$$

Since $\|v - v^*\|_\Phi = \|f_v - f_{v^*}\|_\Phi$ and v^* is the orthogonal projection of v onto $\mathcal{U}$, we have

$$\|v - v^*\|_\Phi = \|f_v - f_{v^*}\|_\Phi = \mathrm{dist}_\Phi(v, \mathcal{U}).$$

In addition, observe that the left side of (3.19) is independent of u, so that we may choose u to be the orthogonal projection of w onto $\mathcal{U}$. Taking this into account, we arrive at the estimate

$$|\overline{w}(f_v - f_{v^*})| \le \mathrm{dist}_\Phi(w, \mathcal{U})\mathrm{dist}_\Phi(v, \mathcal{U}), \tag{3.20}$$

which is in terms of a product of two bounds for optimal approximation problems, and which can be viewed as a generalization of the hypercircle inequality [7, 4].

We remark that there is a useful connection between this bound and more familiar bounds involving the notion of "power function". If δ_q is the Dirac functional at $q \in \mathcal{M}$, the power function $\mathrm{dist}_\Phi(\delta_q, \mathcal{U})$ for $q \in \mathcal{M}$ arises in many publications and there are various papers proving upper bounds for it; see [15, 16]. If we take $s > n/2$ in our setting, then δ_q is in $H_{-s}(\mathcal{M}) \subseteq H_\Phi^*$. Consequently, we may set $w = \delta_q$ in (3.20). Using this in connection with (3.20) then gives us the standard pointwise error bound

$$|(f_v - f_{v^*})(q)| \le \mathrm{dist}_\Phi(\delta_q, \mathcal{U})\|f_v - f_{v^*}\|_\Phi \tag{3.21}$$

for all $q \in \mathcal{M}$.

3.6 Jackson Bounds

We know from §2 that we need to construct recursive Jackson bounds of type (2.4). Their construction is one of the major goals of this paper. In the $\mathcal{M} = \mathbf{R}^n$ case, the usual bounds are pointwise and in the form of (3.21) without using Sobolev spaces. On the other hand, in the $\mathcal{M} = \mathbf{S}^n$ or $\mathcal{M} = \mathbf{T}^n$ cases the available bounds involve Sobolev spaces and are not in the form of (3.21). In both cases there are general techniques to arrive at recursive Jackson bounds, and we choose the spaces $\mathcal{W}_k$ in (2.4) to be H_{Φ_k} normed by $\|\,.\,\|_{\Phi_k}$. Generalized interpolation at level k involves Φ_k, while the error is measured in terms of Φ_{k-1}. These two functions will be related by convolution

$$\Phi_{k-1} = \Phi_k * \Phi_k \tag{3.22}$$

in this paper, but future developments may allow for different choices.

Let us start with the cases $\mathcal{M} = \mathbf{S}^n$ or $\mathcal{M} = \mathbf{T}^n$. There the literature [4] (also see §4.1 and §4.2 below) provides bounds of the form

$$\text{dist}_\Phi(v, \mathcal{U}) = \|f - f^*\|_\Phi \leq c_s \|v\|_{-s} \tag{3.23}$$

for $f = v * \Phi$ with $v \in H_{-s} \subseteq H_\Phi^*$, where c_s is a small constant depending on $\mathcal{M}$, $\mathcal{U}$, s, and Φ. To be recursive, such a bound must be extended to the right by something containing $\|f\|_\Psi$ for another strictly positive definite function Ψ. In particular, we have this result.

Proposition 3.4 *Let f belong to $H_\Psi \subseteq H_\Phi$, and suppose that*

$$C_s^2 := \sup_{\omega \in \mathcal{O}} \frac{\hat{\Psi}(\omega)}{(1 + \lambda(\omega))^s \hat{\Phi}(\omega)^2} \tag{3.24}$$

is finite. If $f = \Phi * v$, where $v \in H_{-s}(\mathcal{M})$, then

$$\|v\|_{-s} \leq C_s \|f\|_\Psi \,. \tag{3.25}$$

Proof: We have that $\hat{v}(\omega) = \hat{f}(\omega)/\hat{\Phi}(\omega)$, and so

$$\begin{aligned}
\|v\|_{-s}^2 = & \int_{\mathcal{O}} |\hat{v}(\omega)|^2 (1 + \lambda(\omega))^{-s} d\nu(\omega) \\
= & \int_{\mathcal{O}} \left(|\hat{f}(\omega)|^2 / \hat{\Phi}(\omega)^2\right) (1 + \lambda(\omega))^{-s} d\nu(\omega) \\
= & \int_{\mathcal{O}} \left(|\hat{f}(\omega)|^2 / \hat{\Psi}(\omega)\right) \frac{\hat{\Psi}(\omega)}{(1 + \lambda(\omega))^s \hat{\Phi}(\omega)^2} d\nu(\omega) \\
\leq & \; C_s^2 \|f\|_\Psi^2 \,.
\end{aligned}$$

Taking square roots above yields (3.25). ■

Using the notation of (3.19) and putting both ingredients together implies the required Jackson bound (2.4) in the form

$$\|f - f^*\|_\Phi \leq c_s C_s \|f\|_\Psi \tag{3.26}$$

for all $f \in H_\Psi$. Applications can now use any choice of Φ, s, and Ψ that satisfies (3.24) and $\|\Phi\|_{s,\infty} < \infty$, for instance $\Psi := \Phi * \Phi$. We will deal with specific cases in §4.

In case $\mathcal{M} = \mathbf{R}^n$ we convert the pointwise error estimate (3.21) into an L_2 bound by simply summing up over the bounded and measurable domain $\Omega \subset \mathbf{R}^n$ in which we do the interpolation. This yields

$$\|f_v - f_{v^*}\|_{L_2(\Omega)} \leq \|\operatorname{dist}_\Phi(\delta_\cdot, \mathcal{U})\|_{L_2(\Omega)} \|f_v - f_{v^*}\|_\Phi.$$

To get a recursive Jackson bound, we proceed along the lines of [14] and impose an additional "boundary condition"

$$v \in L_2(\Omega), \quad \operatorname{supp} v \subseteq \Omega \tag{3.27}$$

for the function $f = f_v = \Phi * v$ that generates the data. But since $\hat{f}_v = \hat{\Phi} \cdot \hat{v}$, we can conclude that (3.27) implies

$$\|v\|_{L_2(\mathbf{R}^n)} = \|v\|_{L_2(\Omega)} = \|f_v\|_{\Phi * \Phi}. \tag{3.28}$$

Then we apply the standard orthogonality argument to prove

$$\begin{aligned}
\|f_v - f_{v^*}\|_\Phi^2 &= \int_{\mathbf{R}^n} (\widehat{f_v} - \widehat{f_{v^*}}) \frac{1}{\hat{\Phi}} (\overline{\widehat{f_v}} - \overline{\widehat{f_{v^*}}}) \\
&= \int_{\mathbf{R}^n} (\widehat{f_v} - \widehat{f_{v^*}}) \frac{\overline{\widehat{f_v}}}{\hat{\Phi}} d\lambda \\
&= \int_{\mathbf{R}^n} (\widehat{f_v} - \widehat{f_{v^*}}) \overline{\hat{v}} d\lambda \\
&= \int_{\mathbf{R}^n} (f_v - f_{v^*}) \overline{v} d\lambda \\
&= \int_\Omega (f_v - f_{v^*}) \overline{v} d\lambda \\
&\leq \|f_v - f_{v^*}\|_{L_2(\Omega)} \|v\|_{L_2(\Omega)} \\
&= \|f_v - f_{v^*}\|_{L_2(\Omega)} \|f_v\|_{\Phi * \Phi}. \\
&\leq \|\operatorname{dist}_\Phi(\delta_\cdot, \mathcal{U})\|_{L_2(\Omega)} \|f_v - f_{v^*}\|_\Phi \|f_v\|_{\Phi * \Phi}.
\end{aligned}$$

Here, the final line used our previous bound, and cancelling a factor yields the required recursive Jackson bound

$$\|f_v - f_{v^*}\|_\Phi \leq \|\operatorname{dist}_\Phi(\delta_\cdot, \mathcal{U})\|_{L_2(\Omega)} \|f_v\|_{\Phi * \Phi} \tag{3.29}$$

which is adapted to convolution (3.22) between different levels.

4 Specific Examples

For $\mathcal{M} = \mathbf{T}^n$ or $\mathcal{M} = \mathbf{S}^n$ there are bounds of the form (3.23), provided some additional restrictions on $\mathcal{U}$ hold [4]. We now make these restrictions precise, first in the case of the circle and second for the 2-sphere.

4.1 The Circle

We wish to discuss a multilevel interpolation problem for the case of the circle. To do this, we will first need to give specific estimates on the constants c_s and C_s in equation (3.26). In particular, to get c_s we need the following result alluded to in §3.6.

Proposition 4.1 *Let $\mathcal{U}$ be the span of the set $\{\delta_{\phi_j}\}_{j=0}^{N-1}$, where the angles ϕ_j are given by*

$$\phi_j = \frac{2\pi}{N}(j + \varepsilon_j),$$

where the ϵ_j's are real numbers that satisfy

$$\sup_j |\epsilon_j| = L, \quad 0 \le L < 1/4,$$

and let $s > \frac{1}{4}$. If v is in $H_{-s}(\mathbf{T}^1)$ and Φ is in $H_{2s}(\mathbf{T}^1)$, and if $\sum_k (1+k^2)^s \widehat{\Phi}(k)$ converges, then

$$\operatorname{dist}_\Phi(v, \mathcal{U}) \le \|v\|_{-s} \sigma(L,N) \left(\sum_{k \notin \mathcal{I}_N} (1+k^2)^s \widehat{\Phi}(k) \right)^{1/2},$$

where $\mathcal{I}_N := [-[N/2], [(N-1)/2]\,] \cap \mathbf{Z}$ and

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

Proof: See Theorem 5.3 in [4]. ■

This result combined with equation (3.23) provides this bound on c_s:

$$c_s \le \sigma(L,N) \left(\sum_{k \notin \mathcal{I}_N} (1+k^2)^s \widehat{\Phi}(k) \right)^{1/2}$$

The choice of Ψ in Proposition 3.4 will of course determine C_s. As we mentioned at the beginning of §3.6, we will choose $\Psi = \Phi * \Phi$. For the case of the circle, this gives $\widehat{\Psi}(k) = \widehat{\Phi}(k)^2$. Inserting this in equation (3.24) gives $C_s = 1$, since s is greater than 0 and $\lambda(k) = k^2$ can take on the value 0. When this result and the previous estimate on c_s are inserted in equation (3.26), we arrive at the Jackson bound that we will require here:

$$\|f - f^*\|_\Phi \leq \sigma(L, N) \left(\sum_{k \notin \mathcal{I}_N} (1 + k^2)^s \widehat{\Phi}(k) \right)^{1/2} \|f\|_\Psi, \tag{4.1}$$

with $\Psi = \Phi * \Phi$ and the other quantities as defined above in Proposition 4.1.

4.2 The 2-sphere

We now deal with the 2-sphere. The space $\mathcal{U}$ will be taken to be the span of point evaluations at points $p_{j,k}$. To describe these points, we will adopt the convention for spherical coordinates that is customary in physics and was employed in [4]: the angle $\theta \in [0, \pi]$ is measured off the positive z-axis and the angle $\phi \in [0, 2\pi)$ is measured off the x-axis in the x-y plane. In addition, take Λ to be a fixed positive integer, then let

$$\begin{aligned} \theta_j &= \tfrac{\pi j}{2\Lambda}, \ j = 0, \ldots, 2\Lambda - 1, \\ \phi_k &= \tfrac{\pi k}{\Lambda}, \ k = 0, \ldots, 2\Lambda - 1, \\ (\theta_j, \phi_k) &= \text{coordinates of } p_{j,k}. \end{aligned}$$

As in the case of the circle, we are faced with the problem of determining the constants c_s and C_s in equation in (3.26).

Proposition 4.2 *Take $s > 1/2$. Let v be in H_{-s}, $\mathcal{U}$ be the span of $\{\delta_{p_{j,k}}\}$ and Φ be a spherical basis function in $H_{2s}(\mathbf{S}^2)$. If Λ is a power of 2, and if the series $\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} (1 + \ell(\ell+1))^s \widehat{\Phi}(\ell, m)$ is convergent, then*

$$\begin{aligned} \mathrm{dist}_\Phi(v, \mathcal{U}) \leq & \left(1 + \tfrac{1}{\sqrt{\pi}} \Lambda^{\frac{3}{2}} \log_2(16\Lambda)\right) \times \\ & \left(\sum_{\ell=\Lambda}^{\infty} \sum_{m=-\ell}^{\ell} (1 + \ell(\ell+1))^s \widehat{\Phi}(\ell, m) \right)^{1/2} \|v\|_{-s}. \end{aligned}$$

Proof: See Theorem 6.6 in [4]. ■

By this result it is clear that we have

$$c_s = \left(1 + \frac{1}{\sqrt{\pi}}\Lambda^{\frac{3}{2}} \log_2(16\Lambda)\right) \left(\sum_{\ell=\Lambda}^{\infty} \sum_{m=-\ell}^{\ell} (1+\ell(\ell+1))^s \widehat{\Phi}(\ell, m)\right)^{1/2} .$$

If we again take $\Psi = \Phi * \Phi$, we have $C_s = 1$ as in the case of the circle.

We want to use an SBF of the form (3.1), where from equation (3.3), with $d_2(\ell) = 2\ell + 1$ and $\omega_2 = 4\pi$, we see that

$$\widehat{\Phi}(\ell, m) = \frac{4\pi}{2\ell+1}\widehat{\Phi}(\ell),\ m = -\ell, \ldots, \ell\,.$$

Using this equation in the expression for c_s and summing over m, we obtain

$$c_s = 2\left(\sqrt{\pi} + \Lambda^{\frac{3}{2}} \log_2(16\Lambda)\right) \left(\sum_{\ell=\Lambda}^{\infty} (1+\ell(\ell+1))^s \widehat{\Phi}(\ell)\right)^{1/2} .$$

When this result and $C_s = 1$ are inserted in equation (3.26), we arrive at the Jackson bound that we will require for the spherical case:

$$\|f - f^*\|_\Phi \le 2\left(\sqrt{\pi} + \Lambda^{\frac{3}{2}} \log_2(16\Lambda)\right) \left(\sum_{\ell=\Lambda}^{\infty} (1+\ell(\ell+1))^s \widehat{\Phi}(\ell)\right)^{1/2} \|f\|_\Psi, \tag{4.2}$$

with $\Psi = \Phi * \Phi$ and the other quantities as defined above in Proposition 4.2.

4.3 Radial Basis Functions in $\mathbf{R}^n$

The recursive evaluation of the Jackson bound (3.29) poses no further problems in case of standard Lagrange interpolation data on scattered locations, because there are many sources (e.g. the discussion in [16]) for pointwise bounds on the power function, thus yielding L_2 bounds on compact domains $\Omega \subset \mathbf{R}^n$. In general, these bounds improve with the smoothness on Φ_k at the current level k and with decreasing sparsity

$$h_k := \sup_{p \in \Omega} \min_{q \in P_k} |p - q|$$

of the k-th level data set $P_k \subset \Omega$. Refined bounds using the boundary conditions (3.27) are in [14].

However, the boundary conditions (3.27) still pose a serious research problem, since they are (so far) required to hold on each level. Using residuals between levels implies the recursions

$$
\begin{array}{rcl}
f_{k+1} & = & f_k - f_k^* = \\
v_{k+1} * \Phi_{k+1} & = & (v_k - v_k^*) * \Phi_k = \\
v_{k+1} * \Phi_{k+1} & = & (v_k - v_k^*) * \Phi_{k+1} * \Phi_{k+1} \\
v_{k+1} & = & (v_k - v_k^*) * \Phi_{k+1},
\end{array}
\tag{4.3}
$$

and if v_k and v_k^* are supported in Ω, the support of v_{k+1} will in general not be confined to Ω. Before we discuss a special case in which we partially overcome the problem, it should be pointed out that the boundary conditions may possibly be unnecessary if other proof techniques for the recursive Jackson bounds are provided by future work. In fact, using the $\|.\|_{\Phi_k}$ norms is quite restrictive, and weaker norms may not require boundary conditions. This is motivated by looking at the cubic spline case, where boundary conditions usually are necessary for convergence of second or higher derivatives, but not for convergence in weaker norms.

In case of compactly supported functions Φ_k, the recursive convolution preserves compact supports and eliminates problems with L_2 integrability, if we start with $f_0 = v_0 * \Phi_0$ satisfying the boundary conditions with respect to some compact subdomain Ω_0 of Ω. The latter is an awkwark hypothesis, but our current technique cannot get away with a less restrictive assumption. If the distributions v_k and v_k^* are supported in some compact subdomain Ω_k whose points have a distance of at least r_k from the boundary of Ω, and if the functions Φ_k have a support radius of $\rho_k = \rho_0 2^{-k}$, then (4.3) implies $r_{k+1} \geq r_k - 2\rho_k$. If we start with $r_0 > 4\rho_0$, induction yields

$$r_{k+1} \geq r_0 - 2\rho_0(1 + 2^{-1} + \ldots + 2^{-k}) \geq r_0 - 4\rho_0 > 0$$

for all k. Thus the boundary conditions are satisfied at all levels.

We now apply the multilevel technique and start at the finest level m with the bound (3.21)

$$|(f_m - f_m^*)(q)| \leq P_m(q)\|f_m - f_m^*\|_{\Phi_m}$$

using the power function $P_m(q) := \text{dist}_\Phi(\delta_q, \mathcal{U})_m$ on points $q \in \Omega_m$ and involving Φ_m. We can assume the L_∞ norm of P_m on Ω_m to be expressible as a function of sparsity

$$h_m := \sup_{p\in\Omega_m} \min_{q\in P_m} |p - q|$$

of the m-th level data set $P_m \subset \Omega_m \subset \Omega$. Thus

$$\|f_m - f_m^*\|_{L_\infty(\Omega_m)} \leq \|P_m\|_{L_\infty(\Omega_m)} \|f_m - f_m^*\|_{\Phi_m}$$

serves as a starting point for the recursion using

$$\|f_{j+1}\|_{\Phi_j} \leq \|P_j\|_{L_2(\Omega_j)} \|f_j\|_{\Phi_{j-1}}$$

for $j = k, k-1, \ldots, 1$. Again, the crucial factor $\|P_j\|_{L_2(\Omega_j)}$ usually is boundable by a function of h_j defined as above. Note that on the levels with coarser data and smoother RBF's with larger supports, the data ara allowed to stay further away from the boundary. Now these bounds multiply neatly and end up with

$$\|f_m - f_m^*\|_{L_\infty(\Omega_m)} \leq \|f_0\|_{\Phi_0} \|P_m\|_{L_\infty(\Omega_m)} \Pi_{j=1}^m \|P_j\|_{L_2(\Omega_j)},$$

as required.

4.4 Comparison

We finally contrast the multilevel method with the standard interpolation error on a single level. Our general multilevel estimate starts with

$$\left| \left(f - \sum_{j=1}^m f_j^* \right)(x) \right| \leq \mathrm{dist}_{\Phi_m}(\delta_x, \mathcal{U}_m)\ \mathrm{dist}_{\Phi_m}(v_m, \mathcal{U}_m) \tag{4.4}$$

where we can evaluate the last factor recursively. Since the use of smoother functions than Φ_m at the finest level m usually is numerically unfeasible, we contrast (4.4) with the estimates obtained for the standard case using the function Φ_m at the maximal number of interpolation points (distribution space $\mathcal{U}_m$) directly to the given data. Again appealing to [4] this has the bound

$$|(f - f_m^{**})(x)| \leq \mathrm{dist}_{\Phi_m}(\delta_x, \mathcal{U}_m)\ \mathrm{dist}_{\Phi_m}(\tilde{v}_m, \mathcal{U}_m), \tag{4.5}$$

where we are using the convolution representation of f as $f = \Phi_m * \tilde{v}_m$. Note that v_m and $\tilde{v}_m$ are different, because v_m is subject to the recursion (4.3).

We wish to contrast the predicted errors in (4.4) and (4.5). A reasonable method is to compare the ratio of the two error bounds, which, after cancelling the common factor $\mathrm{dist}_{\Phi_m}(\delta_x, \mathcal{U}_m)$, leads to the ratio

$$\frac{\mathrm{dist}_{\Phi_m}(v_m, \mathcal{U}_m)}{\mathrm{dist}_{\Phi_m}(\tilde{v}_m, \mathcal{U}_m)}. \tag{4.6}$$

We first discuss the quantity $\text{dist}_{\Phi_m}(v_m, \mathcal{U}_m)$. Upon recursion, it can be bounded above by

$$\|v_1\|_{\Phi_0} \prod_{j=1}^{m} K_j,$$

where the constants K_j are norms of residual operators on nested native spaces, i.e.

$$K_j := \sup_{f_j \in H_{\Phi_j}} \frac{\|f_j - f_j^*\|_{\Phi_j}}{\|f_j\|_{\Phi_{j-1}}},$$

and where best approximations f_j^* to f_j are taken with respect to data functionals in $\mathcal{U}_j$. On the other hand, the denominator has the bound $K_m \|\tilde{v}_m\|_{\Phi_{m-1}}$ by the same reason. Again, we compare ratios of bounds rather than the ratios themselves, and thus we cancel K_m. This leaves the other factors K_j, which we can make small by using sufficiently many data functionals in all of our scenarios. The finally remaining ratio $\|v_1\|_{\Phi_0}/\|\tilde{v}_m\|_{\Phi_{m-1}}$ can be large, but is independent of the data. This shows that the bounds for the multilevel method are superior to the bounds for the single-level method on the last level, provided that there are sufficiently many data.

Acknowledgement

The second author thanks Texas A&M University for its hospitality during two visits in January of 1996 and 1997, and gratefully acknowledges partial financial support provided through the auspices of Professor Charkes K. Chui and by the Mathematics Department of Texas A&M University

References

[1] C. de Boor. *A Practical Guide to Splines.* Springer-Verlag, New York, 1978. Applied mathematical Series 27.

[2] J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *Rev. Française Automat. Informat. Rech. Opér. Anal. Numer.*, 10:5–12, 1976.

[3] J. Duchon. Splines minimizing rotation–invariate semi–norms in sobolev spaces. In W. Schempp and K. Zeller, editors, *Constructive Theory of Functions of Several Variables*, pages 85–100. Springer, Berlin–Heidelberg, 1979.

[4] N. Dyn, F.J. Narcowich, and J.D. Ward. Variational principles and Sobolev-type estimates for generalized interpolation on a Riemannian manifold. Technical Report 371, Department of Mathematics, Texas A & M University, 1996.

[5] M.S. Floater and A. Iske. Multistep scattered data interpolation using compactly supported radial basis functions. To appear in *Journal of Computational and Applied Mathematics*, 1995.

[6] M.S. Floater and A. Iske. Scattered data analysis and industrial applications. In F. Fontanella, K. Jetter, and P.-J. Laurent, editors, *Advanced Topics in Multivariate Approximation*, pages 1–10. World Scientific Publishing, 1996.

[7] M. Golomb and H. F. Weinberger. Optimal approximation and error bounds. In *On Numerical Approximation*, pages 117–190. R. E. Langer, editor; Madison, 1959, 1959.

[8] W.R. Madych and S.A. Nelson. Multivariate interpolation: a variational theory. Manuscript, 1983.

[9] W.R. Madych and S.A. Nelson. Multivariate interpolation and conditionally positive definite functions. *Approximation Theory and its Applications*, 4:77–89, 1988.

[10] C.A. Micchelli. Interpolation of scattered data: distance matrices and conditionally positive definite functions. *Constructive Approximation*, 2:11–22, 1986.

[11] C. Müller. *Spherical Harmonics.* Springer-Verlag, Berlin, 1966.

[12] F.J. Narcowich. Generalized Hermite interpolation positive definite kernels on a Riemannian manifold. *Journal of Mathematical Analysis and Applications*, 190:165–193, 1995.

[13] F.J. Narcowich and J.D. Ward. Generalized Hermite interpolation via matrix-valued conditionally positive definite functions. *Mathematics of Computation*, 63:661–687, 1994.

[14] R. Schaback. Improved error bounds for scattered data interpolation by radial basis functions. Submitted to *Mathematics of Computation*, 1996.

[15] R. Schaback. Comparison of radial basis function interpolants. In *Multivariate Approximation. From CAGD to Wavelets*, pages 293–305. K. Jetter and F. Utreras, editors; World Scientific, London, 1993.

[16] R. Schaback. Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264, 1995.

[17] I. J. Schoenberg. Positive definite functions on spheres. *Duke Math. J.*, 9:96–108, 1942.

[18] J. Stewart. Positive definite functions and generalizations, an historical survey. *Rocky Mountain J. Math.*, 6:409–434, 1976.

[19] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4:389–396, 1995.

[20] Z. Wu. Multivariate compactly supported positive definite radial functions. *Advances in Computational Mathematics*, 4:2833–292, 1995.