Recent Developments in Approximation via Positive Definite Functions

Francis J. Narcowich

Abstract. Positive and conditionally positive definite functions, especially radial basis functions and similar functions for spheres, tori, and even Riemannian manifolds, are of interest because of the their well-known ability to *synthesize* a good surface fit from scattered data. More recently, positive definite basis functions have been employed to *analyze* scattered data. The methods used to do this involve constructing multiresolution analyses or multilevel approximations. This paper will discuss recent developments in the synthesis and analysis problems, point out new directions in their investigation, and remark on applications.

§1. Introduction

Positive definite and conditionally positive definite functions and kernels are used in areas that require fitting a surface to data taken at scattered points in Euclidean space or on some surface, a sphere or torus, say.

When the underlying space is Euclidean, radial basis functions (RBFs)—e.g., Gaussians, multiquadrics, and thin-plate splines—are employed. These are positive definite and conditionally positive definite kernels that are functions of only the distance between points. Because they retain their positivity no matter what the dimension of the space is, they are especially well suited for being activation functions in multi-input, feed-forward neural networks [5, 62]. They also produce pleasing, smooth surfaces [24] when used in fitting noisy data taken at two-dimensional sites. Some of the first RBF applications, which arose in geophysics [32, 33], were of this type. There are several good review articles on RBFs and related topics. Frequently cited ones are those written by Buhmann [9], Dyn [14], and Powell [63]. Cheney's article [11], although not specifically about RBFs, also includes a review of them. Finally, Schaback's article [71] contains a very nice review of many important properties of RBFs.

Positive definite functions on spheres, which Schoenberg [68] introduced, behave like RBFs, and have been called spherical (radial) basis functions (SBFs)

Approximation Theory IX, Volume 2: Computational Aspects
Charles K. Chui and Larry L. Schumaker (eds.), pp. 221–242.

ISBN 0-8265-1326-3.

or spherical splines; they have been applied to scattered data problems on the sphere by Freeden [25], Wahba [77, 78], and others [26]. Recent review articles dealing with positive definite functions on the sphere include Cheney's paper [11], a paper by Freeden *et al.* [26], and a comprehensive survey of approximation on the sphere by Fasshauer and Schumaker [21].

The general idea of positivity, including positive definiteness for kernels, is an old one. In [75], Stewart gives an historical review of it. The most recent review of positive definite kernels is in Cheney's article [11], which we mentioned above.

Fitting a surface to scattered data is the *synthesis* aspect of approximation with positive definite kernels. A second is the *analysis* of scattered data, both by wavelets [12, 13, 60, 61] and by a multi-level method introduced by Floater and Iske [22]; these methods make use of RBFs or similar positive definite kernels in constructing wavelets or fitting surfaces to residuals at various levels.

The paper is organized as follows. The notions of positive definite (PD) and strictly positive definite (SPD) functions and kernels are reviewed in Section 2. All of these are done from the point of view of distributions. A number of examples of such functions and kernels are also described there. In Section 3, we discuss how the SPD functions and kernels can be used in fitting surfaces to scattered data. The discussion includes variational principles, stability of approximations, "uncertainty relations," and rates of approximation.

The analysis of scattered data is discussed in Section 4. There we touch on wavelets developed for analyzing such data in Euclidean space and on spheres. In Section 5, we discuss applications to neural networks in radar, to fitting divergence-free vector fields such as those for magnetic fields and for incompressible fluids with neither sources nor sinks, to PDEs, and to positron emission tomography. In Section 6, we discuss possible avenues for new research.

The topics covered in this review reflect my own taste. For any gaps, omissions, or mistakes, I take responsibility—*mea culpa!*

§2. Positive Definite Functions and Kernels

2.1. Distributional Formulation

In this section, we want to define precisely the set of functions and kernels that we will be looking at. Throughout, $\mathbf{M}^m$ will be either $\mathbb{R}^m$ or a compact, C^∞, m-dimensional manifold, such as an m-sphere or a torus.

A continuous, complex-valued kernel $\kappa(\cdot,\cdot)$ is called *pointwise 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 semi-definite [16, 17, 52, 75].

In applications where fitting derivative information is needed, one needs a more general definition of positive definiteness. To avoid technicalities, we will assume that κ is C^∞. We will say that the kernel κ is positive definite on

$\mathbf{M}^m$ if and only if

$$Q(u) := (\bar{u} \otimes u, \kappa) \geq 0, \tag{1}$$

for every compactly supported distribution on $\mathbf{M}^m$. We remark that when $\mathbf{M}^m$ is a compact manifold, all distributions are compactly supported; the qualification only applies in the case of $\mathbb{R}^m$.

Let δ_p be the Dirac delta function located at p. If in (1) we let $u = \sum_{j=1}^n c_j \delta_{p_j}$, where the c_j's are arbitrary complex numbers, then $Q(u) \geq 0$ is equivalent to the matrix with entries $\kappa(p_j, p_k)$ being positive semi-definite. Thus being positive definite in the sense mentioned above implies that κ is pointwise positive definite. The converse is also true, as the following "extension theorem" shows.

Theorem 1. *Let κ be a C^∞ kernel that is pointwise positive definite. If u is any compactly supported distribution on $\mathbf{M}^m$, then $Q(u) = (\bar{u} \otimes u, \kappa) \geq 0$.*

Thus the two types of positive definiteness are equivalent. For Euclidean kernels, Madych and Nelson [47, §6] proved a version of this theorem. For compact manifolds and C^∞ kernels, Narcowich [52] obtained Theorem 1. A version this theorem that applies to the case when the kernels are in Sobolev spaces, and not necessarily C^∞, may be found in Dyn, Narcowich, and Ward [17, Theorem 2.1].

The theorem has a corollary that is useful for constructing positive definite kernels on manifolds embedded in $\mathbb{R}^m$.

Corollary 2. [16, Corollary 2.1]. *If $\kappa(p,q)$ is a smooth positive definite kernel on $\mathbb{R}^n$, and if $\mathbf{M}^m$ is embedded in $\mathbb{R}^n$, then the restriction of κ to $\mathbf{M}^m$ is positive definite.*

Proof: The matrix $[\kappa(p_j, p_k)]$ is positive semi-definite for any distinct set of points in $\mathbb{R}^n$, and it will therefore be positive definite when the points are restricted to $\mathbf{M}^m$. Since κ will also be smooth when restricted to $\mathbf{M}^m$, Theorem 1 applies and the result follows from it. $\square$

For the sphere, the utility in this result is that one can construct positive definite kernels by restricting an RBF to points on the sphere.

2.2. Strictly Positive Definite Functions and Kernels

Again assume that κ is a C^∞ positive definite kernel on $\mathbf{M}^m$. We will say that κ is strictly positive definite if the equation $Q(u) = 0$ implies that the distribution $u = 0$. This has two immediate consequences.

Proposition 3. *If κ is a smooth, strictly positive definite kernel, then*

$$[\![u, v]\!] := (\bar{v} \otimes u, \kappa), \tag{2}$$

defines an inner product on the space of distributions.

Corollary 4. *If $\{u_1 \ldots u_n\}$ is a linearly independent set of distributions, then the $n \times n$ matrix A, with entries*

$$A_{j,k} = (\bar{u}_k \otimes u_j, \kappa), \tag{3}$$

is positive definite and hence invertible.

Strictly pd vs. pd. There is no extension theorem analogous to Theorem 1. If in Theorem 1 we replace "positive definite" by "strictly positive definite," the result is false. A counter example was given by Ron and Sun [67]. To see why this is so, let us look at the case of the circle $\mathbf{T}^1$. If $\kappa(p,q) = P(p-q)$ is a positive definite function on $\mathbf{T}^1$, then it has the expansion

$$P(\theta) = \sum_{n\in\mathbb{Z}} a_n e^{in\theta}, \ a_n \geq 0.$$

The associated quadratic form $Q(u)$ is

$$Q(u) = \sum_{n\in\mathbb{Z}} a_n |u_n|^2, \ u_n := (u, e^{-in\theta}).$$

It is easy to see that P is *strictly* pd in the distributional sense if and only if $a_n > 0$ for all n. Simply use $u = e^{in\theta}$ to get $a_n = Q(e^{in\theta}) > 0$. Conversely, if all $a_n > 0$ and $Q(u) = 0$ for a distribution u, then each $u_n = 0$, forcing $u = 0$. The Ron/Sun example shows that P may define a strictly positive definite kernel in the *pointwise* sense, even if some of the a_n's actually vanish. This follows because for $u = \sum_{j=1}^N c_j \delta_{p_j}$, the Fourier coefficients are $u_n = \sum_{j=1}^N c_j e^{-inp_j}$; these have support spread out over infinitely many n. Thus, even if a few of the a_n vanish, $Q(u)$ will still be positive.

Examples. We will now list a number of examples of strictly (conditionally) positive definite functions and kernels.

- Radial basis functions (RBFs)
  - — Gaussians, Hardy multiquadric, and inverse Hardy multiquadric, and completely monotonic functions of r^2. Micchelli [51].
  - — Compactly supported RBFs. Narcowich and Ward [58], Wendland [79], Schaback and Wu [73], and Wu [82].
- Periodic basis functions (PBFs)
  - — "Wavians," $\exp(\sum_{j=1}^m \frac{\cos(\theta_j)-1}{\sigma_j^2})$.
  - — Periodic functions having positive Fourier coefficients.
  - — Bernoulli splines ($a_n = |n|^{-k}, n \neq 0, \ a_0 = 1$).
- Spherical basis functions (SBFs)
  - — $\exp(\cos(\theta))$. Here and below $\theta = \theta(p,q)$ is the geodesic distance on the m-sphere between points p and q.
  - — $\sum_{\ell=0}^\infty \sum_{m=1}^{d_\ell} a_{\ell,j} Y_{\ell,j}(p)\overline{Y_{\ell,j}(q)}$, where the $Y_{\ell,j}$'s are spherical harmonics on $\mathbf{S}^m$, $a_{\ell,m} > 0$, and d_ℓ is the dimension of the space of spherical harmonics of order ℓ.
  - — $\sum_{\ell=0}^\infty a_\ell P_\ell(m+1, \cos(\theta)), \ a_\ell > 0$. Xu and Cheney [83].
  - — $[1 - 2\sin(\theta/2)]_+^5 (1 + 10\sin(\theta/2))$ Dyn, *et al.* [16].

- SPD kernels on compact manifolds
  - — $\sum_{k=0}^{\infty} a_k F_k(p)\overline{F_k(q)}$, $a_k > 0$. The F_k's are the eigenfunctions of the Laplace-Beltrami operator on $\mathbf{M}^m$.
  - — $\sum_{k=0}^{\infty} e^{-\lambda_k t} F_k(p)\overline{F_k(q)}$, where $0 = \lambda_0 < \lambda_1 \le \cdots \le \lambda_k < \cdots$ are the eigenvalues of the Laplace Beltrami operator on $\mathbf{M}^m$, counted according to multiplicity.
- SPD functions on $SO(3)$
  - — $e^{\gamma \cos^2 \theta F(\varphi,\psi,\theta)}$, $\gamma > 0$. Gutzmer [31]. φ, θ, ψ are Euler angles, and $F(\varphi,\psi,\theta) := \cos\varphi\cos\psi - (1-\sin\varphi\sin\psi)\cos\theta - \sin\theta(\sin\varphi + \sin\psi)$.

§3. Fitting Surfaces to Scattered Data

3.1. Generalized Hermite Interpolation

Hermite interpolation problems involve fitting a curve or surface to data that contains information involving derivatives. For splines, the standard reference on Hermite interpolation is the book by Lorentz, Jetter, and Riemenschneider [45]. By Generalized Hermite interpolation we will mean interpolating data arising from integrating a (Schwartz) distribution against a smooth, unknown function f. For example, measuring the value of f at a point p can be thought of as obtaining (δ_p, f). More generally, take $\mathcal{U} := \mathrm{Span}\{u_j\}_{j=1}^N$ be a linearly independent (i.e., non-redundant) set of distributions on $\mathbf{M}^m$, and take the data to be

$$(\bar{u}_j, f) = \int_{\mathbf{M}^m} \bar{u}_j(p) f(p) d\mu(p) = d_j \in \mathbf{C}, \quad \text{for } j = 1, \ldots, N. \tag{4}$$

The form of the interpolant that we want here requires discussion. For the RBF case, with the distributions being delta functions and $\kappa(p,q) = F(\|p-q\|)$, the interpolant $\tilde{f}$ has the form

$$\tilde{f}(p) = \sum_j c_j F(\|p - p_j\|),\ p_j \in \mathcal{C}.$$

Rewriting this in terms of κ and $u = \sum_j c_j \delta_{p_j}$, we see that

$$\tilde{f}(p) = \underbrace{\int_{\mathbf{M}^m} \kappa(p,q) u(q) d\mu(q)}_{(\delta_p \otimes u, \kappa)}, \quad \text{where } u \in \mathcal{U}.$$

It is this expression that plays the rôle of interpolant in the general case. Let κ be a positive definite C^∞ kernel on $\mathbf{M}^m$, and let u be a distribution. Denote the function $p \mapsto (\delta_p \otimes u, \kappa)$

$$\kappa \star u(p) := (\delta_p \otimes u, \kappa) = \int_{\mathbf{M}^m} \kappa(p,q) u(q) d\mu(q). \tag{5}$$

We seek $u \in \mathcal{U}$ such that $\tilde{f} = \kappa \star u$ satisfies $(\bar{u}_j, \tilde{f}) = d_j$. When this problem can be solved uniquely for a specific set of linearly independent distributions $\mathcal{U}$, we will say that it is poised. If a problem is poised for every finite, linearly independent set of distributions $\mathcal{U}$, then we say the problem is well poised.

In [37], Jetter, Riemenschneider, and Shen considered the problem of cardinal Hermite interpolation on the lattice $\mathbb{Z}^d$. They employed a matrix-valued symbol to establish that certain classes of interpolation problems were well poised, and in certain cases they obtained approximation orders. Wu [81] was the first to show that Hermite interpolation of scattered data using RBFs was well poised; he used Kriging methods to do so. Independently and using methods different from Wu, although later, Sun [76] also discussed conditions under which Hermite scattered-data problems are well poised.

Narcowich and Ward [59] developed a distributional framework to investigate a class of generalized Hermite interpolation problems in which the interpolants are allowed to be divergence free, and the data set is generated by a collection of arbitrary, linearly independent distributions. We will discuss this in more detail below. We wish to point out that the introduction to [59] incorrectly states that in [81] the number of centers used must equal the number of linear functionals used. In fact, no such restriction is made by Wu in [81].

The well-poisedness of a generalized Hermite interpolation on a C^∞, compact, Riemannian manifold was treated by Narcowich in [52]. Both the m-sphere and m-torus are such manifolds.

3.2. Variational Formulation

We can rewrite the problem described above in terms of the inner product defined in Equation (2). Again, let $\mathcal{U} = \mathrm{Span}\left(\{u_j\}_{j=1}^N\right)$. Assume that there is a distribution v such that the unknown function f satisfies $f = \kappa \star v$. Of course, v is not known either. The point is that we can regard v as the object we wish to interpolate by a distribution in the finite dimension space $\mathcal{U}$. The set of data is then $d_j = (\bar{u}_j, f) = (\bar{u}_j \otimes v, \kappa) = [\![v, u_j]\!]$. The interpolant $\tilde{f} = \kappa \star u$, $u \in \mathcal{U}$, satisfies $d_j = [\![u, u_j]\!]$. If we put $u = \sum_{k=1}^N c_k u_k$, then this becomes

$$\begin{aligned} d_j &= \sum_{k=1}^N [\![u_k, u_j]\!] c_k \\ &= \sum_{k=1}^N (\bar{u}_j \otimes u_k, \kappa) c_k = \sum_{k=1}^N A_{k,j} c_j, \end{aligned}$$

where, by Corollary 4, $A = [A_{k,j}]$ is positive definite and hence invertible. Thus, there exists a unique $u \in \mathcal{U}$ that solves the interpolation problem; it also satisfies the following minimization principle.

Proposition 5. [17, Theorem 3.1]. *The distribution $v - u$ is orthogonal to $\mathcal{U}$ with respect to the inner product (2). In addition, if we let $[\![\cdot]\!]$ be the associated norm, then*

$$[\![v]\!]^2 = [\![v-u]\!]^2 + [\![u]\!]^2.$$

Finally, if $v \neq u$,

$$[\![u]\!] < [\![v]\!].$$

Proof: Standard inner product space methods. $\square$

While we do not need (or want!) to complete the set of distributions comprising the inner product space above—so we are not dealing with a Hilbert space—, the variational principles and results above are essentially those from the theory of reproducing kernel Hilbert spaces [30]. Such spaces have been used extensively in connection with RBFs, SBFs, and similar basis functions, albeit directly in terms of the functions rather than the distributions; they are called native spaces [71] or associated spaces [16] for κ. For RBFs, Madych and Nelson [46] first used them. In the case of the m-sphere, they have been used extensively by many researchers [17, 26, 29, [42]] to get estimates of rates of approximation. One useful tool in getting such rates is the hypercircle inequality [30], which is formulated below in our notation.

Corollary 6. [17, Proposition 3.2]. *If w is a distribution, then*

$$\big|(\bar{w}, f - \tilde{f})\big| = \big|[\![v - u, w]\!]\big| \leq \mathrm{dist}(w, \mathcal{U})\mathrm{dist}(v, \mathcal{U}), \tag{6}$$

where the distances on the right are computed using the norm $[\![\cdot]\!]$.

Proof: This follows from Schwarz's inequality together with the observation that $v - u$ being orthogonal to $\mathcal{U}$ implies that we may replace w by $w - u'$, where $u' \in \mathcal{U}$ is arbitrary. Minimizing over u' then yields the result. $\square$

Let us put this into a more familiar setting. Set $w = \delta_x$, and note that $\mathrm{dist}(v, \mathcal{U}) \leq [\![v]\!]$. Using these in (6) above yields $|f(x) - \tilde{f}(x)| \leq [\![v]\!]\mathrm{dist}(\delta_x, \mathcal{U})$. In the case where $\mathbf{M}^m$ is Euclidean and the u_j's are point evaluations, the function $P(x) = \mathrm{dist}(\delta_x, \mathcal{U})$ is precisely what Schaback calls the *power function* [70], which plays a role in his "uncertainty principle." (See Section 3.3 below.)

By analogy, in the general case we define the function

$$\wp(w) := \mathrm{dist}(w, \mathcal{U}), \tag{7}$$

which satisfies the following inequality.

Corollary 7. *If A_w is the interpolation matrix for κ with data generated by the distributions $\{u_1, \ldots, u_N, w\}$, then*

$$\wp(w)^2 \|A_w^{-1}\| \geq 1. \tag{8}$$

Proof: (The proof we give follows along lines of that in [70]). Because $\wp(w)$ is the distance from w to the finite dimension space $\mathcal{U}$, we can find a linear combination of the u_j's for which $\wp(w) = [\![w - \sum_{j=1}^N c_j u_j]\!]$. Letting $u_{N+1} := w$, and $b_j = -c_j$, $j = 1, \ldots, N$, and setting $b_{N+1} = 1$, we see that

$$\wp(w)^2 = [\![\sum_{j=1}^{N+1} b_j u_j]\!]^2 = \sum_{j,k=1}^{N+1} b_j \bar{b}_k [\![u_j, u_k]\!] = b^* A_w b.$$

Since the lowest eigenvalue of A_w is the reciprocal $\|A_w^{-1}\|$, we have that $b^* A_w b \geq \|A_w^{-1}\|^{-1} \|b\|^2$. In addition, because $b_{N+1} = -1$, $\|b\| \geq 1$. Combining these two inequalities and doing a little algebra yield (8). $\square$

3.3. Uncertainty Principles

A stability vs. goodness-of-fit trade-off occurs in many approximation schemes. For instance, those employing finitely many translates of a single function $g(x)$, whether g is an RBF, a spline, or a sigmoidal function exhibit it. Intuitively, this is because if $g(x)$ itself can be approximated well by a linear combination of $g(x-\xi_1), \ldots, g(x-\xi_n)$, then interpolation matrices associated with g will be nearly singular.

Corollary 7 quantifies this for generalized Hermite interpolation. $\|A_w^{-1}\|$ measures the stability of the interpolation method, whereas $\wp(w)$ is a universal control on the error.

From the hypercircle inequality (6), we see that the distance function $\wp(w)$ is one of two factors bounding the error on the left in (6). The other factor, $\mathrm{dist}(v, \mathcal{U})$, depends on the unknown function f, and is bounded above by $[\![v]\!]$. Although $\mathrm{dist}(v, \mathcal{U})$ does play an important role in controlling the error [16, 17] for a given function, that role is not a universal one. The distance function $\wp(w)$ does *not* depend on f. If it can be made small, the overall error can be made *uniformly* small. Thus (8) expresses quantitatively what we said above.

In the case of ordinary interpolation with RBFs, Schaback [70, 71] quantified this phenomenon, proving the inequality (8) for the Euclidean case and point evaluations. For obvious reasons, he called it the "uncertainty principle." We prefer the term stability-fit trade-off.

There is a second type of "uncertainty principle," one which is far closer to Heisenberg's original position-momentum uncertainty relation. In signal processing this is expressed as a trade-off between a signal's duration and its frequency content, and is used as a measure of how well various wavelets localize in time and frequency.

In connection with their investigation of wavelets on the circle and m-sphere, Narcowich and Ward [60, 61] constructed an uncertainty principle for the circle, which measures a trade-off between angular localization and frequency, and for $\mathbf{S}^2$, which measures the trade-off between localization on $\mathbf{S}^2$ and a "total" angular frequency, which is related to expansions in spherical harmonics.

3.4. Stability

We now turn to a discussion of how stable the interpolation matrix and its inverse are, as measured by norms of inverses and condition numbers.

Numerical results have long shown that the interpolation matrices associated with many RBFs were poorly conditioned. In [15], Dyn, Levin, and Rippa provided methods of preconditioning interpolation matrices for certain RBFs in special cases. However, RBFs and related basis functions are really quite varied in the way they behave, and getting a general theory of preconditioners is not practical. To really deal with this issue requires quantitative results on norms of inverses and condition numbers.

The first of these was a result by Ball [1], who studied the stability for the order 1 RBF $D(r) := r$.

Theorem 8. (Ball [1]). *If s is an odd integer, then for any finite set of distinct points $\{x_j\}_{j=1}^N \subset \mathbb{R}^s$, one has*

$$\|A^{-1}\| \leq \frac{3.55s}{2^s q} \binom{s-1}{\frac{1}{2}(s-1)}, \tag{9}$$

where $A_{j,k} = \|x_j - x_k\|$ and $q := \frac{1}{2} \min_{j \neq k} \|x_j - x_k\|$.

The estimate in (9) is for a special, simple case, and its proof relies on features of that case. Nevertheless, the fact that (9) depends *only* on the minimal separation distance for the data, and not on the number of data points nor on any other aspect of the data, holds in much more general cases. Using Fourier transform techniques, Narcowich and Ward [57, 58] developed a general approach for making such estimates for a large class of order m RBFs, with data in $\mathbb{R}^s$. To avoid technical detail, we will state only the $m = 0$ estimate. Define these:

$$\left.\begin{aligned} \varphi_s(x) :=& x^{-s/2} e^{-1/x}, \ x > 0 \\ \gamma_s :=& 2^{s+1} \Gamma(\frac{s+2}{2}) \\ \delta_s :=& (3^{s-1} \pi \gamma_s^2)^{1/(s+1)} \end{aligned}\right\}. \tag{10}$$

The most general $m = 0$ order RBF has the form $F(r) := \int_0^\infty e^{-tr^2} d\eta(t)$, where $d\eta$ is a nonnegative Borel measure on $[0, \infty)$ such that F is nonconstant and continuous on $[0, \infty)$. For the set of points $\{x_j\}_{j=1}^N \subset \mathbb{R}^s$, the interpolation matrix A associated with F is an $N \times N$ matrix with entries $A_{j,k} := F(\|x_j - x_k\|)$.

Theorem 9. [58, Corollary 2.5]. *With the notation of Theorem 8 and equation (10), we have*

$$\|A^{-1}\| \leq \left(\frac{q^s}{\gamma_s} \int_0^\infty \varphi_s(\frac{q^2 t}{\delta_s}) d\eta(t) \right)^{-1}. \tag{11}$$

The condition number of A is the product $c_A = \|A^{-1}\| \|A\|$. If $|A_{j,k}| \leq M$, then it is easy to show that $\|A\| \leq NM$, where N is the number of centers [57, §VII]. Combining this and (11) yields an N *dependent* bound on c_A. Are there cases for which c_A is *independent* of N? For 0-order RBFs of the form used in Theorem 9, Narcowich, Sivakumar, and Ward found that one has bounds that are *independent* of the number of centers if and only if $\int_0^1 t^{-s/2} d\eta(t) < \infty$ [54].

The estimates in Theorem 9 are independent of any characteristics of the set of centers, except for q, the separation radius, and s, the dimension of the underlying space. How good are they?

Ball, Sivakumar, and Ward [2] showed that in the case where $q > 0$ is small, one can find a finite set of centers with minimal separation $2q$ such that for many RBFs the estimates in (11) are best possible or nearly so; these centers lie on the regular lattice $2q\mathbb{Z}^s$.

The situation is different if the separation distance is fixed. Baxter [3, Theorem 3.10 & Corollary 3.11] found bounds for $\|A^{-1}\|$ when the centers are on the lattice $\mathbb{Z}^s$ (q=1/2), and there are mild restrictions on the RBF:

$$\|A^{-1}\| \leq \left(\sum_{k\in\mathbb{Z}^s} \left|\hat{\varphi}(\|\pi e + 2\pi k\|)\right| \right)^{-1}, \quad e = (1, 1, \cdots, 1) \in \mathbb{Z}^s, \tag{12}$$

where $\hat{\varphi}$ is the s-dimensional Fourier transformation of $F(\|x\|)$. If the symbol function $\xi \in \mathbb{R}^s \mapsto \sum_{k\in\mathbb{Z}^s} \hat{\varphi}(\|\xi + 2\pi k\|_2)$ is continuous, then the bound (12) is best possible. For example, for the Gaussian $F(r) = e^{-tr^2}$ in the $s = 1$ case, the right side of (12) is $\sqrt{\frac{t}{\pi}}\vartheta_2(0, \exp(-\frac{\pi^2}{4t}))^{-1}$, where ϑ_2 is a type-2 theta function [80, p. 464]. For t small, this is approximately $\sqrt{\frac{t}{4\pi}}\exp(\frac{\pi^2}{4t})$. For the same parameters ($q = 1/2$, $s = 1$), the bound in (11) is $\sqrt{2t}\exp(\frac{8\pi}{t})$, which is much larger.

This itself raises the question of whether one can *improve* stability if the placement of centers is at one's disposal. Schaback [69] gave a set of bounds from *below* for $\|A^{-1}\|$. Although bounds from below for $\|A^{-1}\|$ were given in the papers [2, 3], they still relied on gridded data; Schaback's do not. They apply to any finite set of points in $\mathbb{R}^s$, and to a wide class of RBFs. What they give are asymptotic estimates when the number of centers N is large, and simultaneously, the set of centers has a fixed diameter. Knowing these, Schaback showed that choosing gridded centers nearly optimizes stability.

Concerning questions of stability for RBFs with centers in $\mathbb{R}^s$, we wish to mention that preconditioning using a difference operator has been studied by Quak, Sivakumar, and Ward [66] and by Jetter and Stöckler [38]. Finally, the stability of interpolation matrices associated with certain Hermite interpolation problems was investigated in [59, §VII].

The stability of scattered-data interpolation using SBFs has recently become of interest. Levesley, Luo, and Sun [43], Menagatto [48], and Narcowich, Sivakumar, and Ward [55] independently have obtained upper bounds for $\|A^{-1}\|$ similar to those discussed above for RBFs. The methods used, the SBFs to which they apply, and even the formulation of the bounds are considerably different. Here is a nice example worked out by Levesley, Luo, and Sun.

Example 10. [43, Example 3.3]. *Let* $\Phi(\cos(\theta)) := 1 - \sqrt{(1-\cos(\theta))/2}$. *If* $q = \frac{1}{2}\min_{j\neq k}\theta(p_j\,, p_k)$ *is the (spherical) separation radius for the set of centers* $\{p_1\,, \ldots,\, p_N\} \subset \mathbf{S}^2$, *then*

$$\|A^{-1}\| \leq 2\csc^2(q/2).$$

3.5. Rates of Approximation

In $\mathbb{R}^s$, rates of approximation using translates of a single RBF have been obtained in several ways. Duchon [19] and Madych/Nelson [46, 47] used reproducing-kernel Hilbert spaces to obtain rates of approximation for scattered data RBF interpolants to a function f that belonged to the particular RBF's native space. These native spaces become smaller (and smoother) as the smoothness of the RBF increases.

De Boor, DeVore, and Ron [8] obtained L^2 rates of convergence for least-squares approximation from spaces generated by the translates of an RBF on a scaled lattice in $\mathbb{R}^s$ to functions in various Sobolev spaces. Here, the RBF may very well be smooth, and the function one wishes to approximate may be non-smooth. When the RBF has a Fourier transform that is singular at the origin, Dyn and Ron [18] obtained rates of convergence in the scattered-center case; the scattered-center case for RBFs with non-singular Fourier transform—Gaussians, for example—is still not well understood from this point of view. Finally, on the scaled lattice, Jetter, Riemenschneider, Shen [37] have obtained rates when the data set includes information about derivatives (Hermite data).

The question of how fast translates of RBFs can approximate a smooth function in the unit cube in $\mathbb{R}^s$, given the freedom to place N centers anywhere in that cube, was investigated by Mhaskar [49] for Gaussian RBFs and by Mhaskar and Prestin [50] for a broader class of functions similar to RBFs.

In non-Euclidean settings, most of the work has been directed towards obtaining rates for interpolation of data on the circle or the unit sphere $\mathbf{S}^m$. Freeden, Schreiner, and Franke [26] discuss scattered-data point-interpolation on the 2-sphere, and obtain linear rates. In [16, 17], Dyn, Narcowich, and Ward provided the variational approach described earlier to set up a framework for obtaining rates of approximation in non-traditional settings and data–i.e., compact manifolds and generalized Hermite data—, and used this framework to derive approximation rates for various classes of functions on the circle and the 2-sphere. For the circle, the rates in [17] were for a limited class Hermite problems, but included the case of quasi-uniform, scattered data. On the 2-sphere, they obtained rates substantially better than linear, but for point-interpolation of gridded data. Using this framework and a Markov-type inequality on $\mathbf{S}^m$, Jetter, Stöckler, and Ward [39] handled the scattered-data case, obtaining explicit rates that depend on the smoothness of the SBF and of the function to be interpolated, and the mesh size of the data. Taking a different approach, still variational, but closer in spirit to that of Madych and Nelson, von Golitschek and Light [29] obtained rates for conditionally positive definite kernels, with a fixed number of spherical harmonics being added to the interpolant. The rates obtained by the two groups are complementary. For a detailed discussion, we refer the reader to the survey article by Fasshauer and Schumaker [21].

§4. Analysis of Scattered Data

Originally, RBFs were used for surface reconstruction, either directly or via neural networks, in other words, for *synthesizing* scattered data. Two recently developed methods employ RBFs and other positive definite kernels to do *analysis* of such data as well.

4.1. Wavelets

The first method involves constructing wavelets out of non-lattice "translates" of positive definite functions. Without a dyadic lattice, one cannot expect to be able to directly employ Mallat's definition of a multiresolution analysis (MRA) to create such wavelets. The idea that is needed is a *nonstationary* MRA, where the approximation spaces $\mathcal{V}_j$, which are spans of translates of RBFs, satisfy the usual inclusion relations, but the strict scaling requirements are dropped.

These had been used earlier with non-dyadic lattices by Buhmann and Micchelli [10] and de Boor, De Vore, and Ron [7]. Chui, Jetter, Stöckler, and Ward [12, 13] used them to construct RBF wavelets capable of analyzing scattered data in Euclidean space.

The m-sphere, for $m \geq 2$, has no arbitrarily large gridded lattices. For example, for the 2-sphere, the largest lattice of regular points comprises the twenty vertices of the dodecahedron. Thus, to construct wavelets that are intrinsically defined, one needs to use some kind of nonstationary wavelet. There are a number of approaches to this problem. The two making use of SBFs are by Freeden and Windheuser [27], who discretize the continuous wavelet transform, and by Narcowich and Ward [60, 61], who construct classes of nonstationary wavelets by techniques similar to ones used in [12, 13].

4.2. Multilevel Methods

We want to briefly describe a multilevel method that uses strictly positive definite kernels to do pointwise interpolation of large data sets on $\mathbf{M}^m$. When the data sets are very large, on the order of several thousand data points, the interpolation (or least-squares) matrices generated by many SPDKs are full; solving interpolation equations involving them is especially time consuming. Because of well-known conditioning problems [2, 57, 58, 70] that occur when the distance between points becomes small, it may be very difficult to reliably solve interpolation equations in such cases. Even if a solution is available, evaluation of the SPDKs themselves can add to computation time.

The evaluation problem is currently being addressed by Baxter and Roussos [4], Beatson and Newsam [6], and Powell [64, 65] who are working toward fast evaluation and interpolation of large data sets using thin plate splines and other RBFs.

Multilevel methods address what really is a problem of scale. The one described below for RBFs was introduced by Floater and Iske [22].

We have available the data set $\{(\xi_1, f(\xi_1)) \ldots (\xi_N, f(\xi_N)\}$, where N is large. The data sites $\mathcal{C} = \{\xi_1, \ldots, \xi_N\}$ serve to provide the sets of "centers" to be used at the various levels. In general, $\mathcal{C}$ comprises distinct, scattered points. (For the important case of least-squares surface fitting, one could simply choose $\mathcal{C}$ to be a subset of the data sites, or some other suitable point set.)

The central idea is to extract from the original set $\mathcal{C}$ a finite sequence of m nested sets, $\mathcal{C}_1 \subset \cdots \subset \mathcal{C}_m := \mathcal{C}$, each $\mathcal{C}_j$ corresponding to a level j. The $\mathcal{C}_j$'s are chosen to reflect the natural scales occurring in the data sites. Thus, at any level j one has points that are neither too far apart, nor too close together. (In [22], a thinning algorithm was used for this purpose.) For the j^{th} level, one picks a basis of the form $\{\Phi_j(\cdot,\xi)\}_{\xi\in\mathcal{C}_j}$, where Φ_j is an SPDK appropriate for the level. The interpolant is then found as follows.

1) Find $f_1^\star \in \mathrm{Span}\{\Phi_1(\cdot,\xi)\}_{\xi\in\mathcal{C}_1}$ that interpolates f on $\mathcal{C}_1$.
2) Find $g_2^\star \in \mathrm{Span}\{\Phi_2(\cdot,\xi)\}_{\xi\in\mathcal{C}_2}$ that interpolates the residual $g_2 := f - f_1^\star$ on $\mathcal{C}_2$. The level 2 interpolant is then $f_2^\star = f_1^\star + g_2^\star$.
3) For $j = 3,\ldots,m$, find $g_j^\star \in \mathrm{Span}\{\Phi_j(\cdot,\xi)\}_{\xi\in\mathcal{C}_j}$ that interpolates the residual $g_j := g_{j-1} - g_{j-1}^\star$ on $\mathcal{C}_j$. The level j interpolant is $f_j^\star = f_1^\star + g_2^\star + \cdots + g_j^\star$.

We assume that f belongs to a space $\mathcal{W}_0$ of smooth functions, a Sobolev space for example. The native space $\mathcal{W}_j$ associated with each Φ_j (see Section 3.2) changes with level; smoother spaces capture trends, and "spikier" ones, details. These spaces are nested, with $\mathcal{W}_0$ being a subspace of them all.

$$\mathcal{W}_0 \subset \underbrace{\mathcal{W}_1}_{smooth} \subset \mathcal{W}_2 \subset \cdots \subset \mathcal{W}_{m-1} \subset \underbrace{\mathcal{W}_m}_{spikey}.$$

The Floater-Iske technique [22] is to pick the lowest (smoothest) level basis function Φ, and scale it: $\Phi_j(r) := \Phi(\alpha^{j-1}r)$, where $\alpha > 0$.

In Figure 1, we have used the well-known Franke function to generate data to illustrate the first few functions in a multilevel analysis. The Φ_j's were taken to be scaled versions of one of the compactly supported C^2 functions described in [79].

The multilevel method does not dictate the choice for the Φ_j's. Narcowich, Schaback, and Ward [53], who provided a framework with which to analyze rates, and in addition did so in specific cases, start with the highest ("spikey") level $\Phi_m = \Phi$, and then convolve (in the $\star$-product sense): $\Phi_j = \Phi\star\Phi_{j+1}$, $j = 1,\ldots,m-1$. It is both an interesting and practical question as to which ones should be selected for a given application. Fasshauer and Jerome [20] modify the method described above by adding in a smoothing step at each iteration, and obtain improved rates of convergence.

§5. Applications

5.1. Phased Array Antennas

A phased array radar antenna works by detecting the phase differences between the voltages of adjacent antenna elements. For a perfectly tuned array and a single source, all of the phase differences are the same angle (modulo

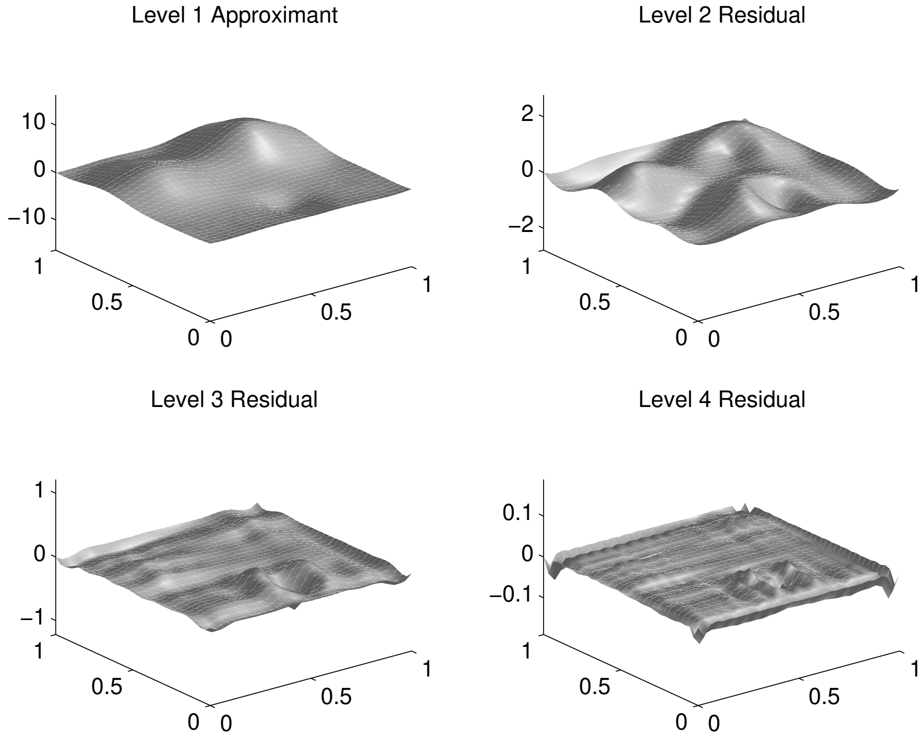

Fig. 1. Multilevel analysis for the Franke function.

2π), which is itself proportional to the sine of the angle of arrival (AOA) of an incoming electromagnetic plane wave. Unlike the familiar dish antenna, a phased array has no moving parts, and is thus less susceptible to breakdown. However, constructing perfectly matched, well-tuned antenna elements is costly. Minor damage to an element can change the relationship between its output and the direction of an incoming signal.

To overcome these drawbacks, Southall, Simmers, and O'Donnell [74] took a neural-network approach based on the premise that even when antennas are neither well-tuned nor perfectly matched, the information in the phase differences is still sufficient to determine the AOA of a single incoming source. The approach described in [74] uses a feed-forward artificial neural network with Gaussian RBFs in the processing nodes. In the language of approximation, this network creates a Gaussian RBF fit from scattered, naturally periodic input data. Because the input variables are periodic and Gaussians are not, the fit one gets can suffer from "edge effects." The remedy for this is to use the periodic basis functions mentioned earlier. See the paper by Southall and O'Donnell in this collection.

One unique feature of the network in [74] is the way the fit is arrived at. The network uses neither least squares nor simple interpolation. Instead, it employs a form of "local" Lagrange interpolation, which works well even though far too few centers are used for there to be a good fit under normal circumstances. This is not explicitly discussed in [74], so we now describe it, but in a general setting.

Suppose that we are given a region $\Omega \subset \mathbf{M}^m$ and a set of centers $\mathcal{C} = \{\xi_1, \ldots, \xi_N\} \subset \Omega$, and an SPD kernel $\kappa(p, q)$. The idea is that most of the points we are interested in are contained in Ω, and the centers $\mathcal{C}$, while scattered, are a quasi-uniform sample of Ω. If $A = [\kappa(\xi_k, \xi_j)]$ is the interpolation matrix for κ, then the ordinary Lagrange interpolants are

$$\Lambda_k(\cdot) = \sum_{j=1}^{N} \left(A^{-1}\right)_{j,k} \kappa(\cdot, \xi_j). \tag{13}$$

These satisfy $\Lambda_k(\xi_j) = \delta_{j,k}$. If at each ξ_k one had measured the value d_k, then the standard Lagrange interpolant would be

$$\tilde{f}_{global}(\xi) = \sum_{k=1} d_k \Lambda_k(\xi)\,. \tag{14}$$

In *local* Lagrange interpolation, to find the value that one should assign to a point $\xi \in \Omega$, one first finds the two indices corresponding to the two largest values in the set $\{\Lambda_k(\xi)\}_{k=1}^{N}$. Let these be k_1 and k_2. The interpolant in this case is then

$$\tilde{f}_{local} = d_{k_1} \frac{\Lambda_{k_1}(\xi)}{\Lambda_{k_1}(\xi) + \Lambda_{k_2}(\xi)} + d_{k_2} \frac{\Lambda_{k_2}(\xi)}{\Lambda_{k_1}(\xi) + \Lambda_{k_2}(\xi)}\,. \tag{15}$$

The two functions multiplying d_{k_1} and d_{k_2} may not be defined for all $\xi \in \Omega$. They *are*, however, Lagrange interpolants near ξ_{k_1} and ξ_{k_2}, whence the name local.

5.2. Incompressible Fluids and Magnetic Fields

In regions where there are neither sources nor sinks, an incompressible fluid, water for example, has a velocity field $\vec{v}(\vec{x})$ that is divergence free,

$$\nabla \cdot \vec{v} = 0\,. \tag{16}$$

Magnetic fields also satisfy (16), because there are no magnetic charges. Standard RBFs can be, and have been [34], used to fit a vector field to data taken from a fluid. Such fits do not, however, preserve the divergence-free nature of the field.

In [59], we introduced "matrix-valued" RBFs, with the expressed purpose of doing divergence-free fits to scattered data from the velocity field of an incompressible fluid or a magnetic field. Of course, one can have s-dimensional divergence free vector fields defined on $\mathbb{R}^s$, $s \geq 2$. An $s \times s$ matrix-valued RBF related to a Gaussian with spread parameter $t > 0$ is

$$h_t^{s\times s}(x) = \left\{(2(s-1)t - 4t^2\|x\|^2)I + 4t^2 x x^T\right\} e^{-t\|x\|^2}, \tag{17}$$

where I is the $s \times s$ identity, x is regarded as a column vector, and x^T is the transpose of x.

For example, if one wants to interpolate data $\{(\xi_j, \vec{d}_j)\}_{j=1}^N$, where ξ_j is a point (bound vector, in older terminology) in $\mathbb{R}^3$ and $\vec{d}_j$ is a (free) vector in $\mathbb{R}^3$, then we seek a vector field of the form

$$\vec{v}(x) = \sum_{j=1}^{N} h_t^{3\times 3}(x - \xi_j)\vec{c}_j \tag{18}$$

for which $\vec{v}(\xi_k) = \vec{d}_k$. (The multiplication in (18) involves matrices, and does not commute.) The interpolation matrix involved is always positive definite, and thus the equations can be inverted. The interpolant in (18) will also automatically satisfy (16).

5.3. Partial Differential Equations

RBFs have been used in solving PDEs in several ways: as a straight collocation method [40, 41], in connection with a variant of the boundary element method [28], and via a different RBF-based collocation method, one that uses RBF-Hermite interpolation [23]. Indeed in [23], Franke and Schaback showed that the Hermite-based method converges to solutions for a variety of PDEs. Recently, several researchers [20, 72] have employed multilevel methods either in direct application to solving PDEs or in related matters.

There are two other situations where RBFs techniques might be helpful. First, although the finite element method is a well-established method for solving PDE's in two and three dimensions, there are situations in which alternative methods would be welcome. For example, in the field of neutron transport, nodal methods are used in their place [35]. It might be possible to couple multilevel methods with nodal methods to obtain smoother solutions to transport equations. Second, the multilevel method has potential as a collocation tool for PDEs in higher dimensions. One such equation arises in finance, where Black-Scholes models [36, Chapter 10] might be analyzed using the multilevel method.

5.4. Tomography

Recently, RBFs have been successfully used by the Medical Image Processing Group at the University of Pennsylvania as a tool for reconstructing images from data obtained via *positron emission tomography* (PET) [44].

There are several interesting features that arise here. The data sets are huge, on the order of millions, and noisy. They are non-traditional, because they amount to line integrals of an unknown distribution function. The distribution function that one is trying to reconstruct is nonnegative, so the reconstruction task involves shape-preservation, and it is continuous, but has edges and corners and is not differentiable.

§6. New Directions

We have discussed many, but certainly not all, aspects of research involving positive definite functions and kernels. Along the way, we have pointed out a number of problems and areas for future research. There are two other new directions for future research. The first is in finding and dealing with non-traditional applications, such as the one involving positron emission tomography. The second is attacking a range of *nonlinear* problems, for instance the shape preservation problem inherent in simply keeping fits to positive densities positive. Another example is the nonlinear optimization problem encountered if one is given N possible centers along with the freedom to place them anywhere. Preliminary results may be found in [49] and [56].

Acknowledgments. I wish to thank Professors Chui and Schumaker for the opportunity to speak at the conference, and for their support. I also wish to thank Professor Ward for many helpful suggestions. This work is partially supported by the Air Force Office of Scientific Research, USAF, under grant number F49620-98-1-0204. The U.S. Government is authorized to reproduce and distribute reprints for Governmental, notwithstanding any copyright notation thereon.

References

1. Ball, K., Eigenvalues of Euclidean distance matrices, J. Approx. Theory **68** (1992), 74–82.
2. Ball, K., N. Sivakumar, and J. D. Ward, On the sensitivity of radial basis interpolation to minimal data separation distance, Constr. Approx. **8** (1992), 401–426.
3. Baxter, B. J. C., Norm estimates for inverses of Toeplitz distance matrices, J. Approx. Theory **79** (1994), 222–242.
4. Baxter, B. J. C. and G. Roussos, On the fast evaluation of conditionally positive definite functions via the fast Gauss transform, *IMACS RBF Conference*, Monterey, California, May 1997, pp. 27–30.
5. Beale, R. and T. Jackson, *Neural Computing*, Adam Hilger Press, Bristol, 1990.
6. Beatson, R. K. and G. N. Newsam, A far field expansion for generalised multiquadrics in $\mathbb{R}^n$, *Computational Techniques and Applications: CTAC97*, Adelaide, Australia, September 1997, p. 30.
7. de Boor, C., R. DeVore, and A. Ron, On the construction of multivariate (pre) wavelets, Constr. Approx. **9** (1993), 123–166.
8. de Boor, C., R. DeVore, and A. Ron, Approximation from shift-invariant subspaces of $L_2(R^d)$, Trans. Amer. Math. Soc. **341** (1994), 787–806.
9. Buhmann, M. D., New developments in the theory of radial basis function interpolation, in *Multivariate Approximation: From CAGD to Wavelets*, K. Jetter and F. I. Uteras (eds.), World Scientific, Singapore, 1993, pp. 35–75.

10. Buhmann, M. D. and C. A. Micchelli, Spline prewavelets for nonuniform knots, Numer. Math. **61** (1992), 455–474.
11. Cheney, W., Approximation using positive definite functions, in *Approximation Theory VIII, Vol. 1: Approximation and Interpolation*, C. K. Chui and L. L. Schumaker (eds.), World Scientific Publishing Co., Inc., Singapore, 1995, pp. 145–168.
12. Chui, C. K., K. Jetter, J. Stöckler, and J. D. Ward, Wavelets for analyzing scattered data: An unbounded operator approach, Appl. Comput. Harmonic Anal. **3** (1996), 254–267.
13. Chui, C. K., J. Stöckler, and J. D. Ward, Analytic wavelets generated by radial functions, Adv. Comp. Math. **5** (1996), 95–123.
14. Dyn, N., Interpolation and approximation by radial and related functions, in *Approximation Theory VI*, C. K. Chui, L. L. Schumaker, and J. Ward (eds.), Academic Press, New York, 1989, pp. 211–234.
15. Dyn, L., D. Rippa, and S. Rippa, Numerical procedures for surface fitting of scattered data by radial functions, SIAM J. Sci. Statist. Comput. **7** (1986), 639–659.
16. Dyn, N., F. J. Narcowich, and J. D. Ward, A framework for interpolation and approximation on Riemannian manifolds, in *Approximation Theory and Optimization*, M. D. Buhmann and A. Iserles (eds.), Cambridge University Press, 1997, pp. 133–144.
17. Dyn, N., F. J. Narcowich, and J. D. Ward, Variational principles and Sobolev-type estimates for generalized interpolation on a Riemannian manifold, Constr. Approx., to appear.
18. Dyn, N. and A. Ron, Radial basis function approximation: from gridded centers to scattered centers, J. London Math. Soc. **71**(3) (1996), 76–108.
19. Duchon, J., Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, Rairo Analyse Numerique **10** (1976), 5–12.
20. Fasshauer, G. E. and J. Jerome, Multistep approximation algorithms: improved convergence rates through postconditioning with smoothing kernels, 1997, preprint.
21. Fasshauer, G. E. and L. L. Schumaker, Scattered Data Fitting on the Sphere, in *Mathematical Methods for Curves and Surfaces II*, M. Dæhlen, T. Lyche and L. L. Schumaker (eds.), Vanderbilt Univ. Press, Nashville & London, 1998, pp. 117–166.
22. Floater, M. S. and A. Iske, Multistep scattered data interpolation using compactly supported radial basis functions, J. Comput. Appl. Math. **73**(1-2) (1996), 65–78.
23. Franke, C. and R. Schaback, Solving partial differential equations by collocation using radial basis functions, 1997, preprint.
24. Franke, R., Scattered data interpolation: tests of some methods, Math. Comp. **38** (1982), 181–199.

25. Freeden, W., On spherical spline interpolation and approximation, Math. Meth. Appl. Sci. **3** (1981), 551–575.

26. Freeden, W., M. Schreiner, and R. Franke, A survey on spherical spline approximation, Surveys Math. Indust. **7** (1997), 29–85.

27. Freeden, W. and U. Windheuser, *Spherical Wavelet Transform and its Discretization*, Berichte der Arbeitsgruppe Technomathematik von Universität Kaiserlautern # 125, 1995.

28. Goldberg, M. A. and C. S. Chen, The theory of radial basis functions applied to the BEM for inhomogeneous partial differential equations, Boundary Elements Communications, **5** (1994), 57–61.

29. Golitschek, M. von and W. A. Light, Interpolation by polynomials and radial basis functions, 1997, preprint.

30. Golomb, M. and H. F. Weinberg, Optimal approximation and error bounds, in *on Numerical Approximation*, by R. E. Langer (ed.), Madison, 1959, pp. 117–190.

31. Gutzmer, T., Interpolation by positive definite functions on locally compact groups with application to $SO(3)$, Results in Mathematics **29** (1996), 69–77.

32. Hardy, R. L., Multiquadric equations of topography and other irregular surfaces, J. Geophys. Res. **76** (1971), 1905–1915.

33. Hardy, R. L., Theory and applications of the multiquadric-biharmonic method, Comput. Math. Appl. **19** (1990), 163–208.

34. Hassan, Y., T. Blanchat, and C. Seeley, Jr., Simultaneous velocity measurements of both components of a two phase flow using particle image velocimetry, Int. J. Multiphase Flow **18** (1992), 1–23.

35. Hennart, J.-P. and E. del Valle, On nodal transport methods, in *highlights of MAFELAP 1996*, Brunel University, June 1996.

36. Hull, J. C., *Options, Futures, and Other Derivative Securities*, 2^{nd} edition, Prentice Hall, Englewood Cliffs, NJ, 1993.

37. Jetter, K, S. Riemenschneider, and Z. Shen, Hermite interpolation on the lattice $\mathbb{Z}^d$, SIAM J. Math. Anal. **25** (1994), 962–975.

38. Jetter, K. and J. Stöckler, A generalization of de Boor's stability result and symmetric preconditioning, Adv. Comput. Math. **3** (1995), 353–367.

39. Jetter, K., J. Stöckler, and J. D. Ward, Error estimates for scattered data interpolation on spheres, Math. Comp., to appear.

40. Kansa, E. J., Multiquadrics—a scattered data approximation scheme with applications to computational fluid dynamics—I. Surface approximations and partial derivative estimates, Computers Math. Applic. **19** (1990), 127–145.

41. Kansa, E. J., Multiquadrics—a scattered data approximation scheme with applications to computational fluid dynamics—II. Solutions to parabolic,

hyperbolic and elliptic partial differential equations, Computers Math. Applic. **19** (1990), 147–161.

42. Levesley, J., W. Light, D. Rogozin, and X. Sun, Variational theory for interpolation on spheres, 1997, preprint.
43. Levesley, J., Z. Luo, and X. Sun, Norm estimates of interpolation matrices and their inverses associated with strictly positive definite functions, 1997, preprint.
44. Lewitt, R. M., S. Matej, and G. T. Herman, Discretization and iterative solution of inverse problems in 3D tomography using bell-shaped radial basis functions having compact support, IMACS RBF Conference, Monterey, California, May 1997, pp. 27–30.
45. Lorentz, G. G., K. Jetter, and S. Riemenschneider, Birkhoff interpolation, Encyclopedia of Mathematics and its Applications, vol 19, Addison-Wesley, Reading, 1983.
46. Madych, W. R. and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions, Approx. Theory Appl. **4** (1988), 77–79.
47. Madych, W. R. and S. A. Nelson, Multivariate interpolation and conditionally positive definite functions II, Math. Comp. **54** (1990), 211–230.
48. Menegatto, V. A., Condition numbers associated with radial-function interpolation on spheres, 1997, preprint.
49. Mhaskar, H. N., Versitile Gaussian networks, in *Proc. of the IEEE Workshop on Nonlinear Image and Signal Processing*, I. Pitas (ed.), Halkidiki, Greece, June 1995, IEEE, pp. 70–73.
50. Mhaskar, H. N. and J. Prestin, On a choice of sampling nodes for optimal approximal of smooth functions by generalized translation networks, 1997, preprint.
51. Micchelli, C. A., Interpolation of scattered data: distances, matrices, and conditionally positive definite functions, Constr. Approx. **2** (1986), 11–22.
52. Narcowich, F. J., Generalized Hermite interpolation and positive definite kernels on a Riemannian manifold, J. Math. Anal. Applic. **190** (1995), 165–193.
53. Narcowich, F. J., R. Schaback, and J. D. Ward, *Multilevel Interpolation and Approximation*, Center for Approximation Theory Report # 379, Department of Mathematics, Texas A&M University, 1997.
54. Narcowich, F. J., N. Sivakumar, and J. D. Ward, On condition numbers associated with radial-function interpolation, J. Math. Anal. Appl. **186** (1994), 457–485.
55. Narcowich, F. J., N. Sivakumar, and J. D. Ward, Stability results for scattered-data interpolation on Euclidean spheres, Adv. Comp. Math., to appear.
56. Narcowich, F. J., P. W. Smith, and J. D. Ward, Density of translates of radial functions on compact sets, in *Approximation Theory VIII, Vol. 1:*

Approximation and Interpolation, C. K. Chui and L. L. Schumaker (eds.), World Scientific Publishing Co., Inc., Singapore, 1995, pp. 435–442.

57. Narcowich, F. J. and J. D. Ward, Norms of inverses and condition numbers for matrices associated with scattered data, J. Approx. Theory **64** (1991), 69–94.

58. Narcowich, F. J. and J. D. Ward, Norm estimates for the inverses of a general class of scattered-data radial-function interpolation matrices, J. Approx. Theory **69** (1992), 84–109.

59. Narcowich, F. J. and J. D. Ward, Generalized Hermite interpolation via matrix-valued conditionally positive definite functions, Math. Comp. **63** (1994), 661–688.

60. Narcowich, F. J. and J. D. Ward, Wavelets associated with periodic basis functions, Appl. Comput. Harmonic Anal. **3** (1996), 40–56.

61. Narcowich, F. J. and J. D. Ward, Nonstationary wavelets on the m-sphere for scattered data, Appl. Comp. Harm. Anal. **3** (1996), 324–336.

62. Poggio, T. and F. Girosi, A theory of networks for approximating and learning, MIT Artificial Intelligence Laboratory and Center for Biological Information Processing (Whitaker College), A.I. Memo No. 1140, C.B.I.P. Paper No. 31, 1989.

63. Powell, M. J. D., The theory of radial basis approximation in 1990, in *Advances in Numerical Analysis II: Wavelets, Subdivision and Radial Functions*, W. Light (ed.), Oxford University Press, 1990, pp. 105–210.

64. Powell, M. J. D., Truncated Laurent expansions for the fast evaluation of thin plate splines, Numer. Algorithms **5** (1993), 99–120.

65. Powell, M. J. D., Thin plate spline interpolation in two dimensions, *Computational Techniques and Applications: CTAC97*, Adelaide, Australia, September 1997, p. 29.

66. Quak, E., N. Sivakumar, and J. D. Ward, Least squares approximation by radial functions, SIAM J. Math Anal. **24** (1993), 1043–1066.

67. Ron A. and X. Sun, Strictly positive definite functions on spheres, Math. Comp. **65** (1996), 1513–1530.

68. Schoenberg, I. J., Positive definite functions on spheres, Duke Math. J. **9** (1942), 96–108.

69. Schaback, R., Lower bounds for norms of inverses of interpolation matrices for radial basis functions, J. Approx. Theory **79** (1994), 287–306.

70. Schaback, R., Error estimates and condition numbers for radial basis function interpolation, Adv. Comp. Math. **3** (1995), 251–264.

71. Schaback, R., Multivariate Interpolation and Approximation by Translates of a Basis Function, in *Approximation Theory VIII, Vol. 1: Approximation and Interpolation*, C. K. Chui and L. L. Schumaker (eds.), World Scientific Publishing Co., Inc., Singapore, 1995, pp. 491–514.

72. Schaback, R., Solving partial differential equations using radial basis functions, IMACS RBF Conference, Monterey, California, May 1997, pp. 27–30.

73. Schaback, R., and Z. Wu, Operators on radial functions, J. Comput. Appl. Math. **73**(1-2) (1996), 1–17.

74. Southall, H. L., J. A. Simmers, and T. H. O'Donnell, Direction finding in phased arrays with a neural network beamformer, IEEE Transactions on Antennas and Propagation **43** (1995), 1369–1374.

75. Stewart, J., Positive definite functions and generalizations, an historical survey, Rocky Mountain J. Math. **6** (1976), 409–434.

76. Sun, X., Scattered Hermite interpolation using radial basis functions, Linear Algebra Applic. **207** (1994), 135-146.

77. Wahba, G., Spline interpolation and smoothing on the sphere, SIAM J. Sci. Statist. Comput. **2** (1981), 5–16.

78. Wahba, G., Surface fitting with scattered noisy data on Euclidean d-space and on the sphere, Rocky Mountain J. Math. **14** (1984), 281–299.

79. Wendland, H., Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree, Adv. Comp. Math. **4** (1995), 389–396.

80. Whittaker, E. T. and G. N. Watson, *A Course in Modern Analysis*, Cambridge University Press, Cambridge, 1965.

81. Wu, Z., Hermite-Birkhoff interpolation of scattered data by radial basis functions, Approx. Theory & Its Appl. **8** (1992), 1–10.

82. Wu, Z., Multivariate compactly supported positive definite radial functions, Adv. Comp. Math. **4** (1995), 283–292

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

Francis J. Narcowich
Department of Mathematics
Texas A&M University
College Station, TX 77843
`fnarc@math.tamu.edu`
`http://www.math.tamu.edu/~francis.narcowich/`