Acta Numerica (2000), pp. 1–38

Radial basis functions

M. D. Buhmann
Mathematical Institute,
Justus Liebig University,
35392 Giessen, Germany
E-mail: Martin.Buhmann@math.uni-giessen.de

Radial basis function methods are modern ways to approximate multivariate functions, especially in the absence of grid data. They have been known, tested and analysed for several years now and many positive properties have been identified. This paper gives a selective but up-to-date survey of several recent developments that explains their usefulness from the theoretical point of view and contributes useful new classes of radial basis function. We consider particularly the new results on convergence rates of interpolation with radial basis functions, as well as some of the various achievements on approximation on spheres, and the efficient numerical computation of interpolants for very large sets of data. Several examples of useful applications are stated at the end of the paper.

CONTENTS

1. Introduction

There is a multitude of ways to approximate a function of many variables: multivariate polynomials, splines, tensor product methods, local methods and global methods. All of these approaches have many advantages and some disadvantages, but if the dimensionality of the problem (the number of variables) is large, which is often the case in many applications from statistics to neural networks, our choice of methods is greatly reduced, unless

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

we resort solely to tensor product methods. In fact, sometimes we are given scattered data, immediately excluding the use of tensor product methods. Also, tensor product methods in high dimensions always require many, often too many data. In this situation, the method of choice is often a radial basis function approach which is, incidentally, also highly useful in lower-dimensional problems and as an alternative to (piecewise) polynomials, because of its excellent approximation properties; at any rate it is universally applicable independent of dimension.

In order to formulate the problem as we will see it in this review, let Ξ be a finite set of distinct points in $\mathbb{R}^n$, which are traditionally called *centres* in radial basis function jargon, because our basis functions will be radially symmetric about these points. The goal of our work is to approximate an unknown function that is only given at those centres via a set of real numbers f_ξ, $\xi \in \Xi$. These are almost always interpreted as function evaluations of some smooth function $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$, so that $f_\xi = f(\xi)$. Here, Ω is a domain in $\mathbb{R}^n$. This point of view will allow us to measure conveniently the uniform approximation error between f and its approximant s. This error depends on the choice of the approximant, on Ξ, on f, and in particular on its smoothness.

In order to approximate with s, which is usually by interpolation, we take a univariate continuous function ϕ that is radialized by composition with the Euclidean norm on $\mathbb{R}^n$, or a suitable replacement thereof when we are working on a sphere in n-dimensional Euclidean space, for instance. This $\phi : \mathbb{R}_+ \to \mathbb{R}$ is the *radial basis function*. Additionally, we take the given centres ξ from the given finite set Ξ of distinct points and use them simultaneously for shifting the radial basis function and as interpolation (collocation) points. Therefore, our standard radial function approximants now have the form

$$s(x) = \sum_{\xi \in \Xi} \lambda_\xi \phi(\|x - \xi\|), \qquad x \in \mathbb{R}^n, \tag{1.1}$$

suitable adjustments being made when x is not from the whole space, and the coefficient vector $\lambda = (\lambda_\xi)_{\xi \in \Xi}$ is an element of $\mathbb{R}^\Xi$. In many instances, particularly those that will interest us in Section 3, the interpolation requirements

$$s\,|_\Xi = f\,|_\Xi \tag{1.2}$$

for given data $f\,|_\Xi$ lead to a *positive definite* interpolation matrix $A = \{\phi(\|\xi - \zeta\|)\}_{\xi,\zeta \in \Xi}$. In that case, we call the radial basis function 'positive definite' as well. If it is, the linear system of equations that comes from (1.1) and (1.2) and uses precisely that matrix A yields a unique coefficient vector $\lambda \in \mathbb{R}^\Xi$ for the interpolant (1.1).

All radial basis functions of Section 3 have this property of positive defin-

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

iteness, as does for instance the Gaussian radial basis function $\phi(r) = \mathrm{e}^{-c^2r^2}$ for all positive parameters c and the inverse multiquadric function $\phi(r) = 1/\sqrt{r^2+c^2}$.

However, in some instances such as the so-called thin-plate spline radial basis function, the radial function ϕ is only *conditionally positive definite* of some order k on $\mathbb{R}^n$, say, a notion that we shall explain and use in the subsequent section. Now, in this event, polynomials $p(x) \in \mathbb{P}_n^{k-1}(x)$ of degree $k-1$ in n unknowns are augmented to the right-hand side of (1.1) so as to render the interpolation problem again uniquely solvable. Consequently we have as approximant

$$s(x) = \sum_{\xi\in\Xi} \lambda_\xi \phi(\|x-\xi\|) + p(x), \qquad x \in \mathbb{R}^n. \tag{1.3}$$

Then the extra degrees of freedom are taken up by requiring that the coefficient vector $\lambda \in \mathbb{R}^\Xi$ is orthogonal to the polynomial space $\mathbb{P}_n^{k-1}(\Xi)$, that is, all polynomials of total degree less than k in n variables restricted to Ξ:

$$\mathbb{R}^\Xi \ni \lambda \perp \mathbb{P}_n^{k-1}(\Xi) \quad \Longleftrightarrow \quad \sum_{\xi\in\Xi} \lambda_\xi q(\xi) = 0, \quad \forall\, q \in \mathbb{P}_n^{k-1}. \tag{1.4}$$

In order to retain uniqueness, Ξ has to contain a $\mathbb{P}_n^{k-1}$-unisolvent subset in this case. When $k = 2$, for instance, this means that the centre-set must not be a subset of a straight line. This is a requirement that can easily be met in most cases.

The two probably best-known and most often applied radial basis functions are called multiquadrics and thin-plate splines, respectively. The former is, for a positive parameter c, $\phi(r) = \sqrt{r^2+c^2}$ and the latter is $\phi(r) = r^2 \log r$, where in the second case (1.3) and (1.4) with $k = 2$ are applied. The multiquadric is, subject to a sign change, conditionally positive definite of order $k = 1$, but it turns out that the original interpolation problem without augmentation by constants is also nonsingular because of special properties of the multiquadric function (Micchelli 1986).

One more well-known example comes up when we set $c = 0$ in the multiquadric example: then we have the so-called linear radial basis function $\phi(r) = r$ which also gives a nonsingular interpolation problem without augmentation by constants. All these examples are useful for various forms of interpolation and approximation and they all allow this interpolation procedure in all dimensions n and all sets of distinct centres, independently of the geometry of the points. In contrast to spline approximation in more than one dimension, for instance, there is no triangulation or tessellation of the data points required, nor is there any restriction on the dimensionality of the problem. We note, however, that these nonsingularity properties are strictly linked to the fact that we use Euclidean norms; for p-norms with

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

$p > 2$, including $p = \infty$, or $p = 1$, singularity can occur in any dimension for unfortunate choices of data points ξ (Baxter 1991).

While it has been known for a long time (see Hardy (1990) for references and the history of the approach) that approximants of the above form exist and approximate well if the centres are sufficiently close together and the function f is smooth enough, it took quite a while to underpin these empirical results theoretically, that is, get existence, uniqueness and convergence results, and extend the known classes of useful radial basis functions to further examples. Now, however, research into radial basis functions is a very active and fruitful area and it is timely to stand back and summarize its new developments in this article. Among the plethora of new papers that are published every year on radial basis functions we have made out and selected five major directions which have had important new developments and which we will review in this work.

One major feature, for example, that has been looked at recently (again) and that we address, is the accuracy of approximation with radial basis functions when the centres are scattered points in a domain, and, ultimately, form a dense subset of the domain – a subject, incidentally, well fitted to begin this review paper in the next section, because the whole theoretical development started some 20 years ago in France with Duchon's contributions (1976, 1978, 1979) to exactly this question in a special context.

Apart from the fundamental question of unique solvability of the ensuing linear system – which has by now been completely answered for large classes of radial basis functions that are conditionally positive definite, mostly following Micchelli (1986) – an important question is that of convergence and convergence rates of those interpolants to the function f that is being approximated by collocation to $f|_\Xi$ if f is in a suitable smoothness space. Duchon (1976, 1978, 1979) gave answers to this question that address the important special case of thin-plate splines $\phi(r) = r^2 \log r$ and its siblings, for instance the odd powers of r ('pseudo-cubics' $\phi(r) = r^3$, *etc.*), and recent research has improved some of his 20 year-old results in various directions, including inverse theorems, and theorems about optimality of convergence orders, the sets of centres still always being finite. Several of these relevant results we address in the next section. We do not comment further, however, on the nonsingularity results, as they have been reviewed often before (Powell (1992*a*), for instance).

Next, some theorems are given on the new classes of radial functions with compact support that are currently under investigation. They must, in the author's opinion, be seen as an alternative to the standard radial functions of global support, like thin-plate splines or the famous and extremely useful multiquadric $\phi(r) = \sqrt{r^2 + c^2}$ (Hardy 1990), but not as an excluding alternative, because the approximation orders they give are much less impressive than the dimension-dependent orders of the familiar radial

functions. Some possible applications, for instance, for numerical solutions of partial differential equations come to mind when those radial functions are used because they can act as finite elements, and they are being tested at the moment by several colleagues for this very purpose. Results are mildly encouraging, although the mathematical analysis is still lacking (see, *e.g.*, Fasshauer (1999)); some mathematical underpinning is given in Franke and Schaback (1998) and Pollandt (1997). We will explain further and give examples in Section 6.

In Section 4, several recent results about the efficient implementation of the radial basis function interpolation are given, especially iterative methods for the computation of interpolants when the number $|\Xi|$ of centres is very large. This work is called for when the radial functions are of global support and increasing with increasing argument, as they often are, because no direct or simple iterative methods will then work satisfactorily, the matrices being ill-conditioned, and sometimes highly so, with exponentially increasing condition numbers. This is particularly unfortunate because, in applications, large numbers of data occur frequently.

Section 5 is devoted to radial basis functions on spheres. Indeed, several radial basis functions were created to interpolate data given on the Earth's surface, for instance potential or temperature data (Hardy 1990). This has inspired many researchers to consider the question of how to approximate efficiently when the data are from a sphere and when the whole idea of distance defining the above approximants is adjusted properly to distances on spheres, that is, geodesic distances. Several groups are currently working on these approximations, which need not be interpolants, although the main focus is, as always with radial basis functions, on interpolation, and we give a brief review of some important new results.

The final section is devoted to applications, mostly to some initial attempts at the numerical solution of differential equations with radial basis function methods.

At the end of this introduction we remind the reader that everything said here is the result of a selection, not comprehensive, and explained, of course, from the author's personal point of view. It is certain that several relevant theorems have been omitted as a consequence. On the other hand, there will be a fairly comprehensive book by the author that gives more details and proofs of many results that are merely stated here.

2. Convergence rates

As usual in the study of methods for the approximation of functions, one of the central themes in the analysis of radial basis functions is their convergence behaviour when the centres become dense in a domain or in the whole underlying Euclidean space. This is highly relevant because it shows how

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

well we can approximate smooth functions even in the practical case when the centres become close together but do not cover a whole domain. Several of the results were initiated by the work of Duchon and this explains the title of the following subsection.

2.1. Improvements and extensions to Duchon's convergence theorems

Many of the theorems about convergence rates are related to the so-called variational approach and founded on ideas of Duchon (1976, 1978, 1979) who considered the – even today – important special cases of radial basis functions of thin-plate spline type in n dimensions, namely

$$\phi(r) = \begin{cases} r^{2k-n} \log r, & \text{if } 2k-n \text{ is an even integer,} \\ r^{2k-n}, & \text{if } 2k-n \text{ is not an even integer,} \end{cases} \tag{2.1}$$

where we admit nonintegral k but always demand $k > \frac{1}{2}n$. The typical case we always think of is the thin-plate spline in two dimensions, namely $n = k = 2$ and therefore $\phi(r) = r^2 \log r$. The aforementioned odd powers – linear or cubic, for instance – also belong to this class. Several further important cases such as multiquadrics or shifted logarithms are in fact derived from the above by composition with $\sqrt{r^2 + c^2}$: namely, $\phi(r) = r$ altered in this fashion gives multiquadrics and $\phi(r) = \log r$ provides the shifted logarithm $\log \sqrt{r^2 + c^2}$ (although in this case $2k = n$). The reason for this transformation is to gain infinite smoothness when c is positive (recall that all of the above are not smooth at the origin when composed with the Euclidean norm).

In this context, the approximants are usually taken from the 'native spaces' X of distributions (Jones (1982), for instance, for generalized functions or distributions) in n unknowns whose total kth degree partial derivatives are square-integrable: we call the space that depends on the choice of ϕ the space $X := D^{-k}L^2(\mathbb{R}^n)$. The Sobolev embedding theorem tells us that this space consists of continuous functions as long as $k > \frac{n}{2}$. This is the first and perhaps most important example of the general notion of native spaces, namely semi-Hilbert spaces X that 'belong' to the radial function and are defined by all distributions $f : \mathbb{R}^n \to \mathbb{R}$ that render a certain ϕ-dependent seminorm $\|f\|_\phi$ finite. In the present case, the seminorm is the homogeneous Sobolev norm of order k. We will return to this concept soon in somewhat more generality.

A beautiful convergence result that holds on a very general domain Ω and does not yet require explicitly native space seminorms (although they are used implicitly in the proof, as we shall see) is the following one of Powell (1994). So long as all domains as general as those in the statement of the theorem are admitted, this is the best possible bound achievable for this general class of domains, that is, the constant C on its right-hand side

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

cannot, for $h \to 0$, be replaced by $o(1)$. In order to state the result we let

$$h := \sup_{x \in \Omega} \inf_{\xi \in \Xi} \|x - \xi\|. \tag{2.2}$$

This notion is also used in the rest of this section, as is the notation $\|\cdot\|_{\infty,\Omega}$ for the Chebyshev norm restricted to Ω.

Theorem 1. Let ϕ be from the class (2.1) with $k = n = 2$ and Ω be bounded and not contained in a straight line. Let s be the radial basis function interpolant to $f|_\Xi$ satisfying (1.2) for $\Xi \subset \Omega$ where k keeps the same meaning as in the introduction, that is, linear polynomials are added to s and the appropriate side conditions (1.4) demanded. Then there is an h- and Ξ-independent C such that

$$\|s - f\|_{\infty,\Omega} \leq Ch\sqrt{\log(h^{-1})}, \qquad 0 < h < 1.$$

In fact we can be even more specific about the constant C in the above error estimate. It is, of course, independent of h and Ξ, but its dependence on f can be expressed by $C = \tilde{c}\|f\|_\phi$, where $\tilde{c}$ depends only on Ω and $\|f\|_\phi$ is the homogeneous Sobolev seminorm of order 2 of f, depending on the aforementioned native space. The general approach to convergence estimates of this form is always to bound the error $|s(x) - f(x)|$ by a fixed constant multiple of

$$\|f\|_\phi \sqrt{\Phi(\alpha)} \tag{2.3}$$

that depends on the radial basis function (2.1) and the dimension n, where $\alpha = (\alpha_\xi) \in \mathbb{R}^\Xi$ are the coefficients of the representation

$$x = \sum_{\xi \in \Xi} \alpha_\xi \xi, \qquad x \in \Omega,$$

and Φ is the so-called power functional

$$\Phi(\alpha) = \phi(0) - 2\sum_{\xi \in \Xi} \alpha_\xi \phi(\|x - \xi\|) + \sum_{\xi \in \Xi} \alpha_\xi \sum_{\zeta \in \Xi} \alpha_\zeta \phi(\|\xi - \zeta\|).$$

Consequently, the main work lies in bounding this power functional from above. If this is done judiciously we can obtain optimal bounds, because the bound of the error function by a suitable constant multiple of (2.3) is best possible (Wu and Schaback 1993).

Bejancu (1997) has generalized this result to arbitrary k and n, and his theorem includes the above result (see also Matveev (1997)). There are no further restrictions on Ω except, in general, its $\mathbb{P}_n^{k-1}$-unisolvency which we demand for the following theorem.

Theorem 2. Let ϕ be from the class (2.1). Let Ω be bounded and contain a $\mathbb{P}_n^{k-1}$-unisolvent subset. Let s be the radial basis function interpolant (1.3) to $f|_\Xi$ for $\Xi \subset \Omega$ as in the introduction where k keeps the same

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

meaning, that is, $\mathbb{P}_n^{k-1}$ polynomials p are added to s with the appropriate side conditions (1.4) on the $\lambda \in \mathbb{R}^{\Xi}$. Then there is an h-independent constant C such that

$$\|s-f\|_{\infty,\Omega} \leq C \begin{cases} h\sqrt{\log(1/h)}, & \text{if } 2k-n=2, \\ \sqrt{h}, & \text{if } 2k-n=1, \\ h, & \text{in all other cases,} \end{cases} \quad \text{and} \qquad 0<h<1.$$

Johnson (1998*b*) has the following improved convergence orders that do, however, require Ω to be a domain with Lipschitz continuous boundary and satisfying an interior cone condition (see any of the cited papers by Duchon for the standard definition of this concept). Now, (2.1) for general n and k are admitted and H^{2k} denotes the usual Sobolev space.

Theorem 3. Let $\tilde{\Omega} \subset \Omega$ be compact and $f \in H^{2k}(\Omega)$ be supported in $\tilde{\Omega}$. Then we have for any $\Xi \subset \Omega$ that contains a $\mathbb{P}_n^{k-1}$-unisolvent subset and an h-independent constant C and the interpolant s on Ξ the error estimate in the p-norm

$$\|s-f\|_{p,\Omega} \leq Ch^{2k+\min[n/p-n/2,0]}, \qquad 0<h<1, \tag{2.4}$$

where $1 \leq p \leq \infty$.

Note that the best rate occurs in (2.4) when $p=2$ but $p=\infty$ gives the inferior rate $2k-\frac{1}{2}n$.

2.2. Upper bounds on approximation orders and inverse theorems; saturation orders

It is a remarkable new development to have *upper bounds* on the obtainable convergence rate rather than the lower bounds thereon as above. Johnson (1998*a*) shows that the rates of Theorem 3 are almost optimal. (As is well known, the optimal ones in case of $\Xi = h\mathbb{Z}^n$ and $p=\infty$ are $2k$ (Buhmann 1990*a*, 1990*b*) – we will come back to this soon.) We still let the radial basis function be from the above class (2.1), although many of the results cited include multiquadric interpolation, for example.

Theorem 4. Let $1 \leq p \leq \infty$ and let Ω be the unit ball. Suppose $\{\Xi = \Xi_h \subset \Omega\}_{h>0}$ is a sequence of finite sets of centres with distance (2.2) for each. Then there exists an infinitely smooth f such that, for the best $L^p(\Omega)$-approximation s to f of the form (1.1) with the appropriate polynomials added, the error on the left-hand side of (2.4) is *not* $o(h^{2k+1/p})$ as h tends to zero.

Therefore, in Theorem 3 we are not far from the optimal result. We do get the optimal results of $\mathcal{O}(h^{2k})$ for uniform convergence ($p=\infty$) on grids $\Xi = h\mathbb{Z}^n$ as mentioned already, but also, as Bejancu (1999) shows us, on finite grids $\Xi = \Omega \cap h\mathbb{Z}^n$, where Ω is a cube with sides parallel to the axes.

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

This extension of the results on $\Xi = h\mathbb{Z}^n$ to $\Xi = \Omega \cap h\mathbb{Z}^n$ relies on the *locality* of the cardinal interpolants on $h\mathbb{Z}^n$ which can be expressed conveniently in Lagrange form

$$s(x) = \sum_{\xi\in\Xi} f(\xi) L\Big(h^{-1}(x-\xi)\Big) = \sum_{i\in\mathbb{Z}^n} f(hi) L\Big(h^{-1}x - i\Big), \qquad x \in \mathbb{R}^n.$$

In this expression, the Lagrange function L that satisfies $L(k) = \delta_{0k}$ for all multi-integers k, and may be expanded as

$$L(x) = \sum_{i\in\mathbb{Z}^n} \lambda_i \phi(\|x - i\|), \qquad x \in \mathbb{R}^n,$$

enjoys the remarkable property of the fast (exponential) decay of $|L(x)|$ (Madych and Nelson 1990), which renders the approximant local with respect to the data. This, in turn, gives rise to the same convergence orders when the domain of approximation is restricted from the whole of the Euclidean space to a cube with sides parallel to the axes.

Theorem 5. Let Ω, $\Xi = \Omega \cap h\mathbb{Z}^n$ and the interpolant s be as above, and let $\tilde{\Omega}$ be a compact subset of the interior of Ω. Then, for $f \in \mathrm{Lip}^{2k+1}(\Omega)$, and for an h-independent constant C,

$$\|s - f\|_{\infty,\tilde{\Omega}} \le C h^{2k}, \qquad 0 < h < 1.$$

Precise upper bounds on the approximation order that can be identified for $\Xi = (h\mathbb{Z})^n$ may be stated in a very general context even for h-dependent radial basis functions ϕ_h that are from a sequence of radial functions $\{\phi_h\}_{h>0}$. Therefore we study approximants from spaces

$$S_h(\phi_h) = \mathrm{span}\left\{\phi_h\Big(\frac{\cdot}{h} - j\Big) \mid j \in \mathbb{Z}^n\right\}.$$

A special case is, of course, $\phi_h \equiv \phi$ for all h, perhaps taken from one of our (2.1). For the statement of the next theorem we let Ω be the unit ball and $\sigma : \mathbb{R}^n \to \mathbb{R}$ a smooth (C^∞) cut-off function that is supported in Ω and satisfies $\sigma\,|_{\frac{1}{2}\Omega} = 1$. We recall that the notation $\check{f}$ stands for the inverse Fourier transform

$$\check{f}(x) = \frac{1}{(2\pi)^n} \int \mathrm{e}^{ix\cdot t} f(t)\,\mathrm{d}t,$$

integrals being over $\mathbb{R}^n$ unless stated otherwise.

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

Theorem 6. Let $1 \le p \le \infty$. Let $\hat{\phi}_h : \mathbb{R}^n \to \mathbb{R}$ be measurable and such that, for some $j_0 \in \mathbb{Z}^n \backslash \{0\}$ and some $\varepsilon \in (0,1)$:

(1) $\hat{\phi}_h(x) \neq 0$ for almost all $x \in \varepsilon\Omega$;
(2) the *inverse Fourier transform* $\left(\sigma\left(\frac{\cdot}{\varepsilon}\right) \frac{\hat{\phi}_h(\cdot+2\pi j_0)}{\hat{\phi}_h}\right)^{\vee}$ is absolutely integrable;
(3) for a measurable function $\widetilde{\rho}$ which is locally bounded and for all positive m,

$$\left\| h^{2k}\,\widetilde{\rho} - \frac{\hat{\phi}_h(h\cdot+2\pi j_0)}{\hat{\phi}_h(h\cdot)} \right\|_{\infty, m\Omega} = o(h^{2k}), \qquad h \to 0.$$

Then the L^p-approximation order from $S_h(\phi_h)$ cannot be more than $2k$, that is, the distance in the L^p-norm between the L^p-closure of $S_h(\phi_h)$ and the class of band-limited f whose Fourier transform is infinitely differentiable cannot be $o(h^{2k})$ as h tends to zero.

We note that the class of all band-limited f whose Fourier transform is infinitely differentiable is a class of very smooth local functions and if the L^p-approximation order to such smooth functions cannot be more than h^{2k}, it cannot be more than h^{2k} to any general nontrivial larger set of less smooth functions.

A typical example occurs when we again use radial basis functions of the form (2.1) where $\hat{\phi}(r)$ is a constant multiple of r^{-2k}. Then condition (1) of the above theorem is certainly true for $\hat{\phi}_h \equiv \hat{\phi}$. Moreover, condition (2) is true because the smoothness of the cut-off function, the smoothness of the function $\hat{\phi}(\|\cdot+2\pi j_0\|)$ in a neighbourhood of the origin for nonzero j_0 and the fact that $2k > n$ imply by Lemma 2.7 of Buhmann and Micchelli (1992) that condition (2) holds. Finally, we can take $\tilde{\rho} = \hat{\phi}^{-1}$ and get the required result from condition (3), namely that $\mathcal{O}(h^{2k})$ is best possible.

This is the obtainable (saturation) order and an inverse theorem of Schaback and Wendland (1998) tells us that all functions for which a better order is obtainable must be trivial in the sense of polyharmonic functions. For its statement we recall the standard notation Δ for the Laplace operator.

Theorem 7. Let Ω be as in Theorem 2, $\Xi \subset \Omega$ a finite centre-set with distance h as in (2.2) and ϕ as in (2.1). If, for any $f \in C^{2k}(\Omega)$, and all compact $\tilde{\Omega} \subset \Omega$,

$$\|s - f\|_{\infty, \tilde{\Omega}} = o(h^{2k}), \qquad 0 < h < 1,$$

then $\Delta^k f = 0$ on Ω.

A very similar, also inverse, but more general theorem from the same paper is the following one with which we close this section. In order to state it we come back to the notion of native spaces and recall that for a

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

radial basis function with positive distributional Fourier transform $\hat{\phi}(\|\cdot\|) : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, the square of the native space norm is

$$\|f\|_\phi^2 := \frac{1}{(2\pi)^n} \int \frac{1}{\hat{\phi}(\|t\|)} |\hat{f}(t)|^2 \, \mathrm{d}t \tag{2.5}$$

and the native space X is the space of all distributions f on $\mathbb{R}^n$ for which (2.5) is finite. In the case of the thin-plate splines or, more generally, (2.1), X agrees with $D^{-k}L^2(\mathbb{R}^n)$, because $\hat{\phi}(\|t\|)^{-1}$ is a constant multiple of $\|t\|^{2k}$, and by the Parseval–Plancherel theorem (Stein and Weiss 1971); this being a special case, other positive $\hat{\phi}$ are admitted as well. For the statement of the following theorem we recall that the radial basis function is *conditionally positive definite* of order k in $\mathbb{R}^n$ if the matrix A with centres from $\Xi \subset \mathbb{R}^n$ is nonnegative definite on the subspace of coefficient vectors $\lambda \in \mathbb{R}^\Xi$ that are orthogonal to $\mathbb{P}_n^{k-1}(\Xi)$. The functions (2.1) are all conditionally positive definite of order k subject to a possibly needed sign change. We use the notation $\tau(x) \sim t(x)$ if both $\tau(x)/t(x)$ and $t(x)/\tau(x)$ are uniformly bounded for the appropriate range of x.

Theorem 8. Let Ω be a bounded domain containing a $\mathbb{P}_n^{k-1}$-unisolvent subset. Let ϕ be conditionally positive definite of order k and satisfy

$$\hat{\phi}(r) \sim r^{-2k}, \qquad r > 0.$$

If $f \in C(\Omega)$ and there exists $\mu > k$ such that

$$\|s - f\|_{\infty,\Omega} \leq Ch^\mu, \qquad h \to 0,$$

for the radial basis function interpolants s defined on all finite sets of centres $\Xi \subset \Omega$ with h given by (2.2), then f is an element of X. (The constant C depends on f but not on h.)

An example for the application of this result is the radial basis function (2.1) with k there and in Theorem 8 being the same.

3. Compact support

This section deals with radial basis functions that are compactly supported, quite in contrast to everything else we have seen before in this article. All of the radial basis functions that we have considered so far have global support, and in fact many of them, such as multiquadrics, do not even have isolated zeros (thin-plate splines do, by contrast). Moreover, the radial basis functions $\phi(r)$ are usually increasing with growing argument $r \to \infty$, so that square-integrability, for example, and especially absolute integrability are immediately ruled out. In most cases, this poses no severe restrictions; we can, in particular, interpolate with these functions and get good convergence rates nonetheless, as we have seen in the previous section. There are, however, practical applications that demand local support of the basis functions

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

such as finite element applications (but note one of the approaches to partial differential equations in the last section which works specifically with globally supported ϕ) or applications with very quickly growing data, or where frequent evaluations with substantial amounts of centres are required, *etc.* Therefore there is now a theory of radial basis functions of compact support where the entire class of ϕs gives rise to positive definite interpolation matrices for distinct centres.

We note that, however, in the approximation theory community there is an ongoing discussion about the usefulness of radial basis functions of compact support because of their inferior convergence properties (unless we wish to again forego the advantages of compact support – see our discussion about scaling in the text below), and in comparison to standard, piecewise polynomial finite elements. The latter, however, are harder to use in grid-free environments and $n > 3$, because they require triangulations to be found first.

As we shall see in this section, there exist at present essentially two approaches to constructing univariate, compactly supported $\phi : \mathbb{R}_+ \to \mathbb{R}$ such that the interpolation problem is uniquely solvable with a positive definite collocation matrix $A = \{\phi(\|\xi - \zeta\|)\}_{\xi,\zeta\in\Xi}$. There will be no restriction on the geometry of the set Ξ of centres, but there are – in fact, there must be – bounds on the maximal spatial dimension n which is admitted for each radial function ϕ so that positive definiteness is retained. So they are still useful in grid-free and high-dimensional applications.

In contrast to the radial basis functions of the previous section, the approximation orders we state, unfortunately, are not nearly as good as the maximal ones available of the well-known globally supported radial functions such as (2.1). This indeed puts a stricter limit to the usefulness of compactly supported radial basis functions than we are used to for the globally supported ones. Be that as it may, it is nonetheless interesting to study the question of when compactly supported radial functions give nonsingular and convergent interpolation schemes.

3.1. Wendland's functions

Initially, there were the approaches by Wu and by Wendland, where the radial basis functions consist of only one polynomial piece on the unit interval $[0, 1]$ and are otherwise zero. Although by this means they are piecewise polynomial seen as a univariate function, the resulting approximants are, of course, not. The whole idea is based on the earlier work by Askey (1973) who observed by considering Fourier transforms that the truncated power function $\phi(r) = \phi_0(r) = (1 - r)_+^\ell$ gives rise to positive definite interpolation matrices A for $\ell \geq [n/2] + 1$. Already here we see, incidentally, an

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

upper bound on the spatial dimension n if we fix the degree of the piecewise polynomial.

In order to derive a large class of compactly supported radial basis functions starting from Askey's result, Schaback and Wu (1995) introduced the two so-called operators on radial functions

$$Df(x) := -\frac{1}{x}f'(x), \qquad x > 0,$$

and

$$If(x) := \int_x^\infty rf(r)\,\mathrm{d}r, \qquad x > 0,$$

that are defined for suitably differentiable or asymptotically decaying f, respectively, and are inverse to each other. Additive polynomial terms do not arise on integration since we always restrict ourselves to compactly supported functions. Next, Wendland (1995) and Wu (1995) use the fact that the said interpolation matrix A for the truncated power ϕ_0 remains positive definite if the basis function

$$\phi(r) = \phi_{n,k}(r) = I^k\phi_0(r), \qquad r \geq 0, \tag{3.1}$$

is used when $\ell = k+[n/2]+1$. The way to establish that fact is by considering the Fourier transform of the n-variate radially symmetric function $\phi(\|\cdot\|)$, which is also radially symmetric and computed by the *univariate* Hankel transform

$$\hat{\phi}(r) = (2\pi)^{n/2}r^{1-n/2}\int_0^\infty s^{n/2}J_{n/2-1}(rs)\phi(s)\,\mathrm{d}s, \qquad r = \|x\| \geq 0, \tag{3.2}$$

where $J_{n/2-1}$ is a Bessel function. This is the radial part of the radially symmetric Fourier transform of $\phi(\|\cdot\|)$. Hence we have to show that positivity of the Fourier transform, which is necessary and sufficient for the positive definiteness of the matrix A by Bochner's theorem, prevails when the operator I above is applied, as long as the restrictions on dimension are observed. This proof is carried out by studying the action of I on the Hankel transform (3.2) and by direct computation and use of identities of Bessel functions.

Starting from this, Wendland (1995, 1998), in particular, developed an entire theory of the radial basis functions of compact support which are piecewise polynomial and are positive definite. This theory encompasses recursions for their coefficients when they are expanded in linear combinations of powers and truncated powers of lower order, convergence results, and minimality of polynomial degree for given dimension and smoothness. Two results below serve as examples for the whole theory. The first states the minimality of degree for given smoothness and dimension and the second is a convergence result.

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

Theorem 9. The radial function (3.1) gives rise to a positive definite interpolation matrix A with radial basis function $\phi = \phi_{n,k}$, and with distinct centres Ξ in $\mathbb{R}^n$. Further, among all such functions for dimension n and smoothness C^{2k}, it is of minimal polynomial degree.

Theorem 10. Let ϕ be defined by (3.1), let f be a function in the Sobolev space $H^{k+(n+1)/2}(\mathbb{R}^n)$, and let k be at least 1 when $n = 1$ or $n = 2$. Then, for a compact domain Ω with centres $\Xi \subset \Omega$, the interpolant (1.1) satisfies

$$\|s - f\|_{\infty,\Omega} = \mathcal{O}\Big(h^{k+1/2}\Big), \qquad h \to 0,$$

where h is given by (2.2).

Examples are, for $\ell \geq [n/2] + 1$,

$$\begin{aligned} \phi(r) &= (1-r)_+^{\ell+1}((\ell+1)r+1), \\ \phi(r) &= (1-r)_+^{\ell+2}((\ell^2+4\ell+3)r^2+(3\ell+6)r+3), \end{aligned}$$

in a form proposed by Fasshauer.

3.2. Further contributions

The second class of radial basis functions of compact support (Buhmann 1998, 2000) are reminiscent of the famous thin-plate splines, albeit truncated in a suitable way, and of a certain convolution form. It contains, for instance, the following cases. (We state the value of the function only on the unit interval; elsewhere it is zero. It can, of course, be scaled suitably depending on the distances between the centres.)

Namely, two examples that give twice and three-times continuously differentiable functions, respectively, in three dimensions and two dimensions are as follows. We state them first because they are useful to illustrate the goal of our later result. The parameter choices for the theorem below $\alpha = \delta = \frac{1}{2}$, $\rho = 1$, and $\lambda = 2$, give, for $n = 3$,

$$\phi(r) = 2r^4 \log r - \frac{7}{2}r^4 + \frac{16}{3}r^3 - 2r^2 + \frac{1}{6}, \qquad 0 \leq r \leq 1,$$

while the choices $\alpha = \frac{3}{4}$, $\delta = \frac{1}{2}$, $\rho = 1$, and $\lambda = 2$, provide, for $n = 2$,

$$\phi(r) = \frac{112}{45}r^{\frac{9}{2}} + \frac{16}{3}r^{\frac{7}{2}} - 7r^4 - \frac{14}{15}r^2 + \frac{1}{9}, \qquad 0 \leq r \leq 1.$$

Another two-dimensional ($n = 2$) example which is twice times continuously differentiable is

$$\phi(r) = \frac{1}{18} - r^2 + \frac{4}{9}r^3 + \frac{1}{2}r^4 - \frac{4}{3}r^3 \log r, \qquad 0 \leq r \leq 1.$$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

Theorem 11. Let $0 < \delta \leq \frac{1}{2}$, $\rho \geq 1$ be reals, and suppose $\lambda \neq 0$ and $\alpha > -1$ are also real quantities with

$$\lambda \in \begin{cases} (-\frac{1}{2}, \infty), & \alpha \leq \min[\frac{1}{2}, \lambda - \frac{1}{2}], & \text{if } n = 1, \quad \text{or} \\ [1, \infty), & -\frac{1}{2} < \alpha \leq \frac{1}{2}\lambda, & \text{if } n = 1, \quad \text{and} \\ (-\frac{1}{2}, \infty), & \alpha \leq \min[\frac{1}{2}(\lambda - \frac{1}{2}), \lambda - \frac{1}{2}], & \text{if } n = 2, \quad \text{and} \\ (0, \infty), & \alpha \leq \frac{1}{2}(\lambda - 1), & \text{if } n = 3, \quad \text{and} \\ (\frac{1}{2}(n-5), \infty), & \alpha \leq \frac{1}{2}(\lambda - \frac{1}{2}(n-1)), & \text{if } n > 3. \end{cases}$$

Then the radial basis function

$$\phi(r) = \int_0^\infty \Big(1 - r^2/\beta\Big)_+^\lambda \beta^\alpha (1 - \beta^\delta)_+^\rho \, \mathrm{d}\beta, \qquad r \geq 0, \tag{3.3}$$

has a positive Fourier transform and therefore, by Bochner's theorem, gives rise to positive definite interpolation matrices A with centres Ξ from $\mathbb{R}^n$. Moreover, $\phi(\|\cdot\|) \in C^{1+\lceil 2\alpha \rceil}(\mathbb{R}^n)$.

There is also a convergence estimate for the above radial functions available which includes scaling of the radial basis function, because, as the distances between the centres become smaller, we wish to decrease the support of the radial function as well. Otherwise we would lose the advantages of compact support since the support relative to the distance between the centres would increase. Note that, in the statement of the following theorem, the approximand f is continuous by the Sobolev embedding theorem precisely as long as $1 + \alpha$ is positive by the conditions in the previous theorem (this is therefore a necessary condition).

Theorem 12. Let ϕ be as in the previous theorem and suppose additionally $\rho > 1$, $2\alpha \leq \lambda - n/2 - 3 + \lfloor \rho \rfloor$. Let Ξ be a finite set in a compact domain Ω. Let s be the scaled interpolant (1.1)

$$s(x) = \sum_{\xi \in \Xi} \lambda_\xi \phi(\eta^{-1}\|x - \xi\|), \qquad x \in \mathbb{R}^n,$$

to $f \in L^2(\mathbb{R}^n) \cap D^{-n/2-1-\alpha} L^2(\mathbb{R}^n)$, with the interpolation conditions $s|_\Xi = f|_\Xi$ satisfied. Then the uniform convergence estimate

$$\|f - s\|_{\infty,\Omega} \leq C h^{1+\alpha} \eta^{-n/2-1-\alpha}$$

holds for $h \to 0$ and positive bounded η, the positive constant C being independent of both h and η.

It is interesting to observe that the function classes of the first subsection of this section can be integrated into the more general class discussed in this subsection. Namely, when the operator D is applied to our radial

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

functions (3.3), it gives

$$D^\lambda \phi(r) = \Gamma(\lambda+1)2^\lambda \int_{r^2}^1 \beta^{\alpha-\lambda}(1-\sqrt{\beta})^\rho \,\mathrm{d}\beta, \qquad 0 \le r \le 1,$$

for $\delta = \frac{1}{2}$ and any integral nonnegative λ. Therefore, by one further application of the differentiation operator and explicit evaluation,

$$D^{\lambda+1}\phi(r) = \Gamma(\lambda+1)2^{\lambda+1} r^{2\alpha-2\lambda}(1-r)_+^\rho = \Gamma(\lambda+1)2^{\lambda+1}(1-r)_+^\rho, \quad r \ge 0,$$

for $\alpha = \lambda$. Now let $\lambda = k-1$, $k \ge 1$,$\rho = [\frac{1}{2}n] + k + 1$ and recall that, on the other hand, we know that the radial functions $\phi_{n,k}$ of Wendland are such that

$$D^k \phi_{n,k}(r) = (1-r)_+^{[\frac{1}{2}n]+k+1}, \qquad r \ge 0.$$

Therefore, the functions generated in this fashion can also be derived from (3.3), albeit with a parameter α that does not fulfil the conditions of Theorem 11, so it is a type of continuation. The functions from Wu (1995) also belong to that class as they were shown by Wendland to be special cases of his functions. Therefore, since the class of Wendland functions described here contains those introduced by Wu (1995), our class covers both. Finally, because we know that all those functions have *positive Fourier transforms*, there follow immediately new results about the positivity of Hankel transforms (3.2) which extend the work by Misiewicz and Richards (1994), and this is in spite of the fact that the parameter α here does not yield the conditions of Theorem 11.

4. Iterative methods for implementation

Among the most useful radial basis functions which provide good, accurate approximations are the thin-plate splines, for instance in two dimensions, and multiquadrics. However, these two, like many other examples of radial basis functions with *global support*, lead to linear interpolation systems that require special considerations if $|\Xi|$ is more than a few hundred when we calculate the coefficients λ_ξ. Otherwise the computational cost of $\mathcal{O}(|\Xi|^3)$ is prohibitively expensive for direct methods, storage being another major obstacle if the size of the set of centres is too large – even with the fastest and biggest workstations currently available.

There are several approaches currently under intense investigation, mainly in Cambridge and in Christchurch, New Zealand, for dealing with this problem, two of which we shall explain here to exemplify what can be done today when we have 50000 centres, say. This is the BGP (Beatson–Goodsell–Powell) method of local Lagrange functions, another highly relevant and successful class of methods being the fast multipole approaches (Beatson and Newsam (1992); see also Beatson and Light (1997) and Beatson, Cherrie and Mouat (1999)). Like the fast multipole methods, the BGP algorithm

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

(Beatson, Goodsell and Powell 1995) is iterative and depends on an initial structuring of the centres Ξ before the start of the iteration, although that one is less complicated than the fairly sophisticated hierarchical structure demanded for the multipole schemes, especially when $n > 2$.

4.1. The BGP method

We explain the method for $n = 2$, but we remark that the case $n = 3$ is under investigation and will probably be admissible and have suitable software soon. The special further complication for $n = 3$ is the important three-dimensional structuring of the data; conceptually there are no changes in the algorithm we describe below. We also take thin-plate splines only, although the method has actually been already tested successfully on multiquadrics as well (Faul and Powell 1999), and once again there are almost no changes except for a different k which plays an important rôle below. This is because the method makes, as we shall see, extensive use of the variational properties explained in Section 2 of this paper, and those can be extended to multiquadrics, for instance, although they were first found in connection with thin-plate splines.

The interpolant we wish to compute still has the same form as in the introduction for $k = n = 2$ and $\phi(r) = r^2 \log r$ but we denote the actual, sought interpolant by s^*, whereas s denotes in this section only the active approximation to s^* at each stage of the algorithm. This is useful for describing our iterative scheme. The basic idea of the algorithm is derived from a Lagrange formulation of the interpolant

$$s^*(x) = \sum_{\xi\in\Xi} f(\xi) L_\xi(x), \qquad x \in \mathbb{R}^n, \tag{4.1}$$

instead of (1.1), where each Lagrange function L_ξ satisfies the Lagrange conditions

$$L_\xi(\zeta) = \delta_{\zeta\xi}, \qquad \xi, \zeta \in \Xi, \tag{4.2}$$

and is of the form

$$L_\xi(x) = \sum_{\zeta\in\Xi} \lambda_{\zeta\xi} \phi(\|x - \zeta\|) + p_\xi(x), \qquad x \in \mathbb{R}^n. \tag{4.3}$$

Here, $p_\xi \in \mathbb{P}_2^1$ and $\lambda_{\cdot,\xi} \perp \mathbb{P}_2^1(\Xi)$ for each ξ as usual. Clearly, the computation of such full Lagrange functions would be just as expensive as solving the full usual linear interpolation system of equations. Therefore the idea is to replace (4.2) by local Lagrange conditions which require for each ξ only that the identity holds for some $q = 30$ points, say, ζ that are nearby ξ. Hence we take Lagrange functions that are still of the form (4.3) but with at most q nonzero coefficients. We end up with an approach to our interpolation

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

problem that resembles a domain decomposition approach, because we divide the set of centres into subsets and regard the interpolation method as a local method on those subsets.

So that we can associate with each ξ the Lagrange function L_ξ and its 'active' point set of centres, we call the latter $\mathcal{L}_\xi \subset \Xi$, $|\mathcal{L}_\xi| = q$. In addition to having q elements, $\mathcal{L}_\xi$ must contain a unisolvent subset with respect to $\mathbb{P}_n^{k-1}$ as required before. We also extract another set of centres with a unisolvent subset Σ of approximately the same size q from Ξ for which no local Lagrange functions are computed because that set is sufficiently small to allow direct solution of the interpolation problem.

In order that an iterative method can be applied, we order the centres $\Xi \setminus \Sigma = \{\xi_i\}_{i=1}^m$ for which local Lagrange functions are computed and have the final requirement that, for all $i = 1, 2, \ldots, m$,

$$\xi_i \in \mathcal{L}_{\xi_i} \subset \Xi \setminus \{\xi_1, \xi_2, \ldots, \xi_{i-1}\}.$$

We require that the q points in the set above are those among the centres that follow ξ_i which are the closest ones to ξ_i. There may be ties in the necessary ordering procedure that can be broken randomly. The Lagrange conditions that must be satisfied by the local Lagrange functions are now the locally restricted conditions

$$L_\xi(\zeta) = \delta_{\zeta\xi}, \qquad \zeta \in \mathcal{L}_\xi. \tag{4.4}$$

We employ the same notation for the local Lagrange functions as for the full ones, as the latter will no longer occur in this section. Using these local Lagrange functions, (4.1) is naturally no longer a representation of the exact approximant s^*, but only an approximation thereof, and here is where the iteration and iterative correction come in. Of course the Lagrange functions, that is, their coefficients, are computed in advance once and for all and stored before the beginning of the iterations. (This is an $\mathcal{O}(|\Xi|)$ process.) A useful comparison of this approach with the Newton formulation of univariate polynomial interpolants is made by Powell (1999).

In those iterations, we make an iterative refinement of the approximation at each step of the algorithm by correcting its residual through updates

$$s(x) \longrightarrow s(x) + L_\xi(x) \times c_\xi(x),$$

where

$$c_\xi(x) = \frac{1}{\lambda_{\xi\xi}} \sum_{\zeta \in \mathcal{L}_\xi} \lambda_{\zeta\xi}(f(\zeta) - s(\zeta)). \tag{4.5}$$

We shall see in the proof of the next theorem why $\lambda_{\xi\xi}$ is positive and we may therefore divide by $\lambda_{\xi\xi}$. The correction (4.5) is added for all $\xi \in \Xi \setminus \Sigma$ for each sweep of the algorithm. The final stage of each sweep consists of

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

the correction

$$s(x) \longrightarrow s(x) + \sigma(x),$$

where σ is the full standard thin-plate spline solution of the interpolation problem with centres Σ computed with a direct method

$$\sigma(\xi) = f(\xi) - s(\xi), \qquad \xi \in \Sigma.$$

Here also, s denotes the current approximation to s^* after all intermediate steps. This finishes the sweep of the algorithm, and, if we started with a trivial approximation $s = s_0 = 0$ to s^*, the sweep replaces s_m by s_{m+1} that goes into the next sweep. The stopping criterion can be, for instance, that we terminate the algorithm if all the residuals are sufficiently small. The main work is clearly the computation of the residuals, the coefficients of the local Lagrange made available before the start. For that, a method such as the one described in the next subsection about multipole methods can be used.

It is quite a novelty that there is a convergence proof of this algorithm, although it had been known for some time that the method performs exceptionally well in practical computations even if Ξ contains as many as 50000 points. It turns out that it is not unusual to have an increase of accuracy of one digit per sweep of the algorithm, that is, each sweep reduces $\max|s(\xi) - f(\xi)|$ by a factor of ten, which indicates very fast convergence.

We provide the essentials of the proof of this theorem (Faul and Powell 1999) – which is in fact not very long – because it is highly instructive about the working of the algorithm and involves many concepts that are typical for the use and analysis of radial basis functions, such as native spaces and their seminorms and semi-inner products.

Theorem 13. Let $\{s_j\}_{j=0}^{\infty}$ be a sequence of approximations to s^* generated by $s_0 = 0$ and the above algorithm. Let the radial function be the thin-plate spline function and $n = 2$. Then $\lim_{j\to\infty} s_j = s^*$.

Proof. We show first that the native space norm $\|s^* - s_j\|_\phi$ is monotonically decreasing with increasing j. Let $(\cdot,\cdot)$ be the semi-inner product that corresponds to the native space norm, namely

$$(u, v) = \frac{1}{(2\pi)^n} \int \frac{1}{\hat{\phi}(\|t\|)}\, \hat{u}(t)\overline{\hat{v}(t)}\, \mathrm{d}t. \tag{4.6}$$

Here, the requirements $\|u\|_\phi < \infty$, $\|v\|_\phi < \infty$ are sufficient for the above integral to exist, by the Cauchy–Schwarz inequality. In the thin-plate spline case and in two dimensions, this is by Parseval–Plancherel a multiple of

$$\int \frac{\partial^2 u(x,y)}{\partial x^2}\frac{\partial^2 v(x,y)}{\partial x^2} + 2\frac{\partial^2 u(x,y)}{\partial x \partial y}\frac{\partial^2 v(x,y)}{\partial x \partial y} + \frac{\partial^2 u(x,y)}{\partial y^2}\frac{\partial^2 v(x,y)}{\partial y^2}\, \mathrm{d}x\, \mathrm{d}y$$

and has as kernel the space of linear polynomials.

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

We start with $s = s_0 = 0$. Each full sweep of the algorithm replaces s_j by s_{j+1}, but within each sweep there are further $m+1$ updates . By those updates, $s_{j,0} = s_j$ is replaced by $s_{j,1}$, and so on until $s_{j,m}$ is replaced by $s_{j+1} = s_{j,m+1} = s_{j,m} + \sigma$ where σ is defined above. At the $(j-1)$st stage, for each index i, $\|s^* - s_{j-1,i-1}\|_\phi$ is replaced by

$$\|s^* - s_{j-1,i-1} - \theta_{\xi_i} L_{\xi_i}\|_\phi,$$

where $\theta_{\xi_i} = c_{\xi_i}$. It is elementary that we get

$$\|s^* - s_{j-1,i-1} - \theta L_{\xi_i}\|_\phi^2 = \min!$$

when the linear parameter θ assumes the value

$$\theta = \frac{(s^* - s_{j-1,i-1}, L_{\xi_i})}{\|L_{\xi_i}\|_\phi^2}. \tag{4.7}$$

We claim that (4.7) is the same as θ_{ξ_i} and in the proof of this fact we shall also show as a by-product that the denominator in (4.7) is positive. Indeed, recalling that the additional polynomials used are in the kernel of the inner product (4.6), we get from the reproducing kernel identity

$$\begin{aligned} \|L_{\xi_i}\|_\phi^2 &= \left(\sum_{\zeta \in \mathcal{L}_{\xi_i}} \lambda_{\zeta\xi_i} \phi(\|\cdot - \zeta\|), \sum_{\tau \in \mathcal{L}_{\xi_i}} \lambda_{\tau\xi_i} \phi(\|\cdot - \tau\|) \right) \\ &= \sum_{\zeta \in \mathcal{L}_{\xi_i}} \lambda_{\zeta\xi_i} \sum_{\tau \in \mathcal{L}_{\xi_i}} \lambda_{\tau\xi_i} \phi(\|\zeta - \tau\|) \\ &= \lambda_{\xi_i\xi_i}, \end{aligned}$$

by (4.4) and because p_ζ is annihilated by the side conditions on the coefficients $\lambda_{\zeta\xi_i}$. This also shows, incidentally, that the important inequality

$$\lambda_{\xi_i\xi_i} > 0 \tag{4.8}$$

holds because $\|L_{\xi_i}\|_\phi^2$ is not zero. Otherwise L_{ξ_i} would be in the kernel of our semi-inner product and not able to satisfy the cardinality conditions. Moreover, by the same token, we get from the reproducing kernel properties

$$\begin{aligned} (s^* - s_{j-1,i-1}, L_{\xi_i}) &= \sum_{\zeta \in \mathcal{L}_{\xi_i}} \lambda_{\zeta\xi_i} (s^* - s_{j-1,i-1}, \phi(\|\cdot - \zeta\|)) \\ &= \sum_{\zeta \in \mathcal{L}_{\xi_i}} \lambda_{\zeta\xi_i} (s^*(\zeta) - s_{j-1,i-1}(\zeta)) \\ &= \sum_{\zeta \in \mathcal{L}_{\xi_i}} \lambda_{\zeta\xi_i} (f(\zeta) - s_{j-1,i-1}(\zeta)). \end{aligned}$$

Next, we have to consider the alteration to $\|s^* - s_{j-1,m}\|_\phi$ as soon as σ is added to $s_{j-1,m}$. In order to prove the monotonic decrease of $\|s^* - s_{j-1,m}\|_\phi$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

when this happens, we need to prove that the following inner product vanishes

$$\begin{aligned}\left(s^*-s_{j-1,m}-\sigma,\sum_{\tau\in\Sigma}\hat{\lambda}_\tau\phi(\|\cdot-\tau\|)+\hat{p}\right)\\ =\quad\left(s^*-s_{j-1,m}-\sigma,\sum_{\tau\in\Sigma}\hat{\lambda}_\tau\phi(\|\cdot-\tau\|)\right)\\ =\quad 0.\end{aligned}$$

In the above, the coefficients $\hat{\lambda}_\tau$ are real and $\hat{p}\in\mathbb{P}_n^{k-1}$. This orthogonality relation is established using the same facts as those we required for showing $\|L_{\xi_i}\|_\phi^2=\lambda_{\xi_i\xi_i}$ and therefore we do not repeat the arguments.

In summary, we now know that $\|s_{j,i}-s^*\|_\phi$ tends to a limit for *all* fixed i, and j increasing, because it is bounded below by zero and monotonically decreasing, as it is a subsequence of the monotonically decreasing doubly indexed sequence $\{\|s_{j,i}-s^*\|_\phi\}_{j=0,0\le i\le m}^{\infty}$. Moreover, by the definition of our semi-inner product

$$\|s^*-s_{j-1,i-1}-\theta_{\xi_i}L_{\xi_i}\|_\phi^2=\|s_{j-1,i-1}-s^*\|_\phi^2-\frac{(s^*-s_{j-1,i-1},L_{\xi_i})^2}{\|L_{\xi_i}\|_\phi^2}.$$

Therefore, for all centres

$$(s^*-s_{j-1,i-1},L_{\xi_i})\to 0,\qquad j\to\infty,\ \xi_i\in\Xi\setminus\Sigma, \tag{4.9}$$

because $\|s_{j,i}-s^*\|_\phi$ converges. In particular, it follows that

$$(s^*-s_j,L_{\xi_1})\to 0,\qquad j\to\infty. \tag{4.10}$$

Moreover, $s_{j,1}=s_j+(s^*-s_j,L_{\xi_1})L_{\xi_1}/\|L_{\xi_1}\|_\phi^2$, so that we also get

$$\left(s^*-s_j-\frac{(s^*-s_j,L_{\xi_1})L_{\xi_1}}{\|L_{\xi_1}\|_\phi^2},L_{\xi_2}\right)\to 0,\qquad j\to\infty.$$

Therefore

$$(s^*-s_j,L_{\xi_2})\to 0,\qquad j\to\infty,$$

and indeed

$$(s^*-s_j,L_\xi)\to 0,\qquad j\to\infty,\ \xi\in\Xi\setminus\Sigma. \tag{4.11}$$

Finally, we observe

$$(s^*-s_j,L_{\xi_i})=\sum_{\zeta\in\mathcal{L}_{\xi_i}}\lambda_{\zeta\xi_i}(s^*(\zeta)-s_j(\zeta))\to 0,\qquad j\to\infty. \tag{4.12}$$

Recalling that s^*-s_j restricted to Σ vanishes anyway, and recalling that $\lambda_{\xi_i\xi_i}$ is positive, we now go backwards and start from $i=m$ with an induction argument, whereupon (4.8), (4.9) and (4.12) imply $s_j(\zeta)\to s^*(\zeta)$ as $j\to\infty$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

for all $\zeta \in \Xi$. This implies $s_j \to s^*$ for $j \to \infty$, as demanded, because the space spanned by the translates $\phi(\|\cdot - \xi\|)$ plus the polynomial space is finite-dimensional. □

4.2. Fast multipole methods

Another approach for approximation and iterative refinement to the radial basis function interpolants is that of *fast multipole methods*. These algorithms are based on analytic expansions of the underlying radial functions for large argument. We see that this is possible for all non-compactly supported radial basis functions that have been mentioned, because they are smooth away from their centres. For the compactly supported ones, the approach for fast evaluation and solution of the linear system is, of course, not required. The salient ideas date back to a paper of Greengard and Rokhlin (1987) where the methods are used to solve numerically integral equations – this is related to our radial basis function approximations because the sums and coefficients ('weights') can be viewed as discretizations of integrals.

The methods require structuring the data in a hierarchical way before the onset of the iteration, and computing so-called far-field expansions (Laurent series expansions for large arguments, as already alluded to) in advance. The goal is to reduce the cost of a single approximate evaluation of the radial function sum down to $\mathcal{O}(1)$ and storage to $\mathcal{O}(|\Xi|)$. The far-field expansions exploit the fact that radial basis functions of the form (2.1) are analytic except at the origin, even when made multivariate through composition with Euclidean norms. Therefore, they can be approximated well away from the origin by a truncated Laurent expansion. The accuracy (that is, the length of the truncated expansion) to this can be preset and made to match any chosen accuracy ε, for instance a small positive multiple of the machine precision of the computer in use, but of course the cost of the method, that is, the multiplier in the operation count, rises with larger accuracy. In tandem with the hierarchical structure of the centre-set, which we shall explain in some detail below, this allows approximative evaluation of many radial basis function terms whose centres are close to each other simultaneously by a single finite Laurent expansion that is inexpensive, as long as the argument x is far from the said cloud of centres. By contrast, all radial functions whose translates are close to x are computed exactly and explicitly. We shall call those centres the near field below.

The hierarchical structure of the centres is fundamental to the algorithm because it decides and orders what is far and what is near any given x where we wish to evaluate. It is built up as follows. We assume for the sake of an easy exposition that $\Xi \subset [0,1]^2$ and that the ξ are fairly uniformly distributed in that unit square so that uniform subdivisions are acceptable. We form a *quad-tree* of centres which contains as a root Ξ and as the next

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

children the intersections of the four quarter squares of the unit square with Ξ. They are in turn divided into four grandchildren each in the same fashion. We stop this process at a predetermined level that in part determines the accuracy of the calculation in the end. Each member of this family is called a panel. We point out that there are implementations where the divisions are only by two and not by four (Powell 1993). There is no principal reason for making this choice that forms a binary tree instead of a quad-tree and both approaches have been tried successfully.

Now we must define the far field and the near field for each x where $s(x)$ is computed. Given x, all centres ξ that are in the near field give rise to explicit evaluations of $\phi(\|x-\xi\|)$. The near field consists simply of contributions from all points which are not 'far', according to the following definition: we say x is far away from a panel T and therefore from all centres in that panel, if there is at least one more panel between x and T, and if this panel is on the same level of parenthood as T itself.

Next, we have to decide how to group the far points. All panels Q are in the *evaluation list* of a panel T if Q is either at the same or a coarser (higher) level than T and every point in T is far away from Q, and if, finally, T contains a point that is *not* far away from the parent of Q. Thus the far field of an x whose closest centre is in a panel T is the sum of all

$$s_Q(x) = \sum_{\xi \in Q} \lambda_\xi \phi(\|x-\xi\|) \tag{4.13}$$

such that Q is in the evaluation list of T. For each (4.13), a common Laurent series is computed. Since we do not know the value of x that is to be inserted into (4.13), we compute the *coefficients* of the Laurent series and insert x later on.

In a set-up process, we compute these expansions of the radial function for large argument, that is, their coefficients, and store them; when x is provided at the evaluation stage, we combine those expansions for all centres from each evaluation list, and in an additional, final step we can simplify further by approximating the whole far field by one Taylor series.

A typical Laurent series expansion to thin-plate spline terms, which we still use as a paradigm for algorithms with more general classes of radial basis functions, is as follows. To this end, we work in two dimensions $n = 2$ and identify two-dimensional real space with one-dimensional complex space.

Lemma 1. Let z and t be complex numbers, and let

$$\phi_t(z) := \|t-z\|^2 \log \|t-z\|.$$

Then, for all $\|z\| > \|t\|$, and denoting the real part by $\mathrm{Re}\,$,

$$\phi_t(z) = \mathrm{Re}\left\{ (\bar{z}-\bar{t})(z-t)\left(\log z - \sum_{k=1}^{\infty} \frac{1}{k} \left(\frac{t}{z}\right)^k \right) \right\},$$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

which is the same as

$$(\|z\|^2 - 2\mathrm{Re}\,(\bar{t}z) + \|t\|^2)\log\|z\| + \mathrm{Re}\left\{\sum_{k=0}^{\infty}(a_k\bar{z} + b_k)z^{-k}\right\},$$

where $b_k = -\bar{t}a_k$ and $a_0 = -t$ and $a_k = t^{k+1}/[k(k+1)]$ for positive k. Moreover, if the above series is truncated after $p+1$ terms, the remainder is bounded above by

$$\frac{\|t\|^2}{(p+1)(p+2)}\frac{c+1}{c-1}\left(\frac{1}{c}\right)^p$$

with $c = \|z/t\|$.

In summary, the principal steps of the whole algorithm are as follows.

Set-up

(1) Perform the repeated subdivision of the square down to $\log|\Xi|$ levels and sort the elements of Ξ into the finest level panels.
(2) Form the Laurent series expansions (*i.e.*, compute their coefficients) for all fine level panels R.
(3) Translate centres of expansions so that the expansions can be re-used for other centres and, by working up the tree towards coarser levels, form analogous expansions for all less refined levels.
(4) Working down the tree from the coarsest level, compute Taylor expansions of the *whole* far field for each panel Q.

Evaluation at x

(1) Locate the finest level panel Q containing the centre closest to x.
(2) Evaluate $s(x)$ by computing near field contributions explicitly and by using the Taylor approximation of the entire far field. For the far field, we use all panels that are far away from x and are not subsets of any coarser panels already considered.

The computational cost without set-up is $\mathcal{O}(|\Xi|)$ because of the hierarchical structure. The set-up cost is $\mathcal{O}(\log|\Xi|)$, because of the tree structure, but the constant contained in this estimate may be large because of the computation of the various expansions and the design of the tree. This is so although each expansion is an $\mathcal{O}(1)$ procedure. In practice it turns out that the method is superior to direct computation if $|\Xi|$ is at least of the order of 200 points.

In which way is this algorithm now related to the computation of interpolation coefficients? It is related in one way because the efficient evaluation of the linear combination of thin-plate spline translates is required in the first algorithm presented in this section. There the residuals $f_\xi - s(\xi)$ played an important rôle, $s(\xi)$ being the current linear combination of thin-plate

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

splines at a centre and the current approximation to the solution we require. Therefore, to make the BGP algorithm efficient, fast evaluation of s is needed, partly because our radial basis functions are conditionally positive definite (the preconditioning making the matrices positive (semi-)definite matrices) and partly because the condition numbers severely influence the convergence properties of conjugate gradient methods.

However, the importance of fast availability of residuals is not only restricted to the BGP method. Other iterative methods for computing radial basis function interpolants such as conjugate gradient methods are also in need of these residuals. In order to apply those, however, a preconditioning method is usually needed.

There are various possible improvements to the algorithm of this subsection. For instance, we can use an adaptive method for subdividing into the hierarchical structure of panels, so that the panels may be of different sizes, but always contain about the same number of centres. This is particularly advantageous when the data are highly non-uniformly distributed.

5. Interpolation on spheres

Because of the many applications that suit radial basis functions in geodesy, there is already a host of papers that specialize radial basis function approximation and interpolation to spheres. Freeden and co-workers (1981, 1986, 1995) have made a very large number of contributions to this aspect of approximation theory of radial basis functions. There are excellent and long review papers available from the work of this group (see the cited references) and we will therefore be relatively brief in this section. Of course, we no longer use the conventional Euclidean norm in connection with a univariate radial function when we approximate on the $(n-1)$ sphere S^{n-1} within $\mathbb{R}^n$ but apply so-called geodesic distances. Therefore the standard notions of positive definite functions and conditional positive definiteness no longer apply, and one has to study new concepts of (conditionally) positive definite functions on the $(n-1)$ sphere. This started with Schoenberg (1942), who characterized positive definite functions on spheres as those ones whose expansions in series of Gegenbauer polynomials always have nonnegative coefficients. Xu and Cheney (1992) studied strict positive definiteness on spheres and gave necessary and sufficient conditions. This was further generalized by Ron and Sun in 1996

Recent papers by Jetter, Stöckler and Ward (1999) and Levesley, Light, Ragozin and Sun (1999) use native spaces, (semi-)inner products and reproducing kernels (*cf.* Saitoh (1988) for a treatise on the theory of reproducing kernels) to derive approximation orders in a very similar fashion to the work summarized in Section 2. They all apply the *spherical harmonics* $\{Y_k^{(\ell)}\}_{k=1}^{d_\ell}$ that form an orthonormal basis of the d_ℓ-dimensional space of polynomials

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

on the sphere in $\mathbb{P}_n^\ell(S^{n-1}) \cap \mathbb{P}_n^{\ell-1}(S^{n-1})^\perp$. They are called spherical harmonics because they are the restrictions of polynomials of total degree ℓ to the sphere and are in the kernel of the Laplace operator Δ. The dimension d_ℓ is computable, and this is not difficult, but it is not important to us at this stage.

Then, a *native space* X is defined for all expansions of functions on the sphere in spherical harmonics: namely, the native space's elements are

$$f(x) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} \hat{f}_{\ell k} Y_k^{(\ell)}(x), \qquad x \in S^{n-1}, \tag{5.1}$$

whose coefficients $\hat{f}_{\ell k}$ satisfy certain square summability conditions with prescribed positive real weights $a_{\ell k}$. In other words, the native space is defined by

$$X = \left\{ f \Big| \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} \frac{|\hat{f}_{\ell k}|^2}{a_{\ell k}} < \infty \right\}. \tag{5.2}$$

The native space X given in (5.2) and functions (5.1) will give rise to a reproducing kernel that is positive definite, but if we enlarge the space by starting the first sum in (5.2) only at $\ell = \kappa$ and thereby weakening conditions, we can also get conditionally positive definite (reproducing) kernels (Levesley et al. 1999) for the ensuing spaces X_κ. Then the native space will be a semi-Hilbert space, that is, the inner product that we shall describe shortly has a nullspace

$$K = \left\{ f \Big| f = \sum_{\ell=0}^{\kappa-1} \sum_{k=1}^{d_\ell} \hat{f}_{\ell k} Y_k^{(\ell)} \right\}.$$

Now, a standard choice for the positive weights for defining the space X which are often independent of k is $a_{\ell k} = (1+\lambda_\ell)^{-s}$. This gives rise to the Sobolev space $H^s(S^{n-1})$. Here the $\lambda_\ell = \ell(\ell+n-2)$ are eigenvalues of the Laplace–Beltrami operator.

The inner product that the native space is equipped with can be described by

$$\langle f, g \rangle = \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} \frac{1}{a_{\ell k}} \hat{f}_{\ell k} \hat{g}_{\ell k},$$

where the coefficients are defined through (5.1); they are still assumed to be positive. The reproducing kernel that results from this Hilbert space X with the above inner product and that corresponds to the function of our previous radial basis functions in the native space is, when x and y are on

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

the sphere,

$$\phi(x,y)=\sum_{\ell=0}^{\infty}\sum_{k=1}^{d_\ell}a_{\ell k}Y_k^{(\ell)}(x)Y_k^{(\ell)}(y),\qquad x,y\in S^{n-1}. \tag{5.3}$$

This can be simplified by the famous addition theorem (Stein and Weiss 1971) to $\phi(x,y)=\phi(x^Ty)$, where

$$\phi(t)=\frac{1}{\omega_{n-1}}\sum_{\ell=0}^{\infty}d_\ell a_{\ell k}P_\ell(t), \tag{5.4}$$

ω_{n-1} being the measure of the unit sphere, if the coefficients are constant with respect to k. Here, the d_ℓ are as above and P_ℓ is a Gegenbauer polynomial (Abramowitz and Stegun 1972) normalized by $P_\ell(1)=1$. Therefore, we now use (5.3) or (5.4) for interpolation on the sphere, in the same place and with the same centres Ξ as before, but they are from the sphere themselves of course. Convergence estimates are available from all three sources mentioned above that vary in approaching the convergence question. Using the mesh norm

$$h=\sup_{x\in S^{n-1}}\inf_{\xi\in\Xi}\arccos(x^T\xi),$$

Jetter et al. prove the following theorem. The notation $|\Xi|$ is for the cardinality of the set Ξ as before.

Theorem 14. Let X and Ξ be as above with the given mesh norm h. Let κ be a positive integer such that $h\le 1/(2\kappa)$. Then, for any $f\in X$, there is a unique interpolant s in

$$\text{span}\,\{\phi(\xi,\cdot)\mid\xi\in\Xi\}$$

that interpolates f on Ξ and satisfies the error estimate

$$\|s-f\|_\infty^2\le\frac{5(|\Xi|+1)}{\omega_{n-1}}\|f\|_\phi^2\sum_{\ell=\kappa+1}^{\infty}\left(d_\ell\max_{k=1,\dots,d_\ell}a_{\ell k}\right). \tag{5.5}$$

Corollary 1. Let the assumptions of the previous theorem hold and suppose further that $|\Xi|+1\le C_1\kappa^{n-1}$ and

$$\frac{C_2}{1+\kappa}\le h\le\frac{1}{2\kappa}.$$

Then the said interpolant s provides

$$\|s-f\|_\infty=\mathcal{O}\left(\left(\frac{h}{C_2}\right)^{(\alpha-n)/2}\right)$$

or

$$\|s-f\|_\infty=\mathcal{O}\left(\frac{\exp(-\alpha C_2/2h)}{h^{(n-1)/2}}\right),$$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

respectively, if $d_\ell \times \max_{k=1}^{d_\ell} a_{\ell k}$ is bounded by a constant multiple of $(1+\ell)^{-\alpha}$ for an $\alpha > n$ or by a constant multiple of $\exp(-\alpha(1+\ell))$ for a positive α, respectively.

An error estimate of Levesley et al., which includes conditionally positive definite kernels for $\kappa = 1$ or $\kappa = 2$, is as follows, for $n = 2$ (see also Freeden and Hermann (1986) for a similar, albeit slightly weaker result).

Theorem 15. Let X, Ξ, h and κ be as above, let s be the minimal norm interpolants to $f \in X_\kappa$ on Ξ. When ϕ is twice continuously differentiable on $[1-\varepsilon, \varepsilon]$ for some $\varepsilon \in (0,1)$, then

$$\|s - f\|_\infty \leq Ch^2 \|f\|_\phi,$$

so that, in particular, a polynomial $p \in \mathbb{P}_3^{\kappa-1}(S^2)$ is added to the s used in Theorem 14.

6. Applications

6.1. Numerical solution of partial differential equations

Given that radial basis functions are known to be useful to approximate multivariate functions efficiently, it is suitable to apply them to approximate solutions of partial differential equations numerically. Three approaches have been tried and tested in this direction, namely collocation techniques, variational formulations and boundary element methods, all in order to solve elliptic partial differential equations with boundary values given. There are various reasons why radial basis functions are useful for these three approaches. The first of these is useful because we know much about existence and accuracy of radial basis function interpolants, especially when the data are scattered, which is useful for non-grid collocation. The second resembles typical finite element applications, where usually radial basis functions of compact support are used to mimic the standard finite element approach with multivariate piecewise polynomials. Finally, boundary element methods are suitable in several cases when radial basis functions are known to be fundamental solutions (Green's functions) of elliptic partial differential operators, most notably powers of the Laplace operator. An example is the thin-plate spline radial basis function and the bi-Laplacian operator. After all, in boundary element methods, explicit solutions of the associated homogeneous problem are required in advance, for which it is immensely helpful to have Green's functions to work with.

Naturally, an important decision is the choice of radial basis function, especially whether globally or locally supported ones should be used. In a Galerkin approach, locally supported elements are almost always employed. Further, the use of radial basis functions becomes particularly interesting when nonlinear partial differential equations are solved or non-grid

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

approaches are used, for instance because of non-smooth domain boundaries, where non-uniform knot placement is always important to modelling the solution to good accuracy.

We begin with a description of the collocation approach, which is the first approach one is tempted to try because of our knowledge of interpolation properties of the radial basis functions. The first important decision is whether to use the well-known, standard globally supported radial functions such as multiquadrics or the new compactly supported ones that are described earlier in this review. Since the approximation properties of the latter are not as good as the former ones, we have a trade-off between accuracy on one hand and sparsity of the collocation matrix on the other hand. Compactly supported ones, if scaled suitably, give banded collocation matrices, while the globally supported ones give no sparsity to speak of. When we use the compactly supported radial functions we have, in fact, another trade-off which we have not mentioned so far, because even their scaling pits accuracy against sparseness of the matrix. If we impose no scaling to the radial basis function as in Theorem 12, we do have satisfactory convergence, as shown in Theorem 10, but basically the radial basis function behaves as a globally supported one, with essentially full matrices, since the centres become dense in each of the radial functions' supports. In the other extreme case, when we scale so that there is always a uniformly bounded number of centres inside each support, we run the risk of losing convergence altogether – but the interpolation matrix may be diagonal (and nonsingular, of course). There are several approaches to fixing this problem, and we will mention two of them while describing algorithms.

One typical linear partial differential equation problem suitable for collocation techniques reads

$$\begin{aligned} Lu(x) &= f(x), \quad x \in \Omega \subset \mathbb{R}^n, && (6.1)\\ Bu|_{\partial\Omega} &= q, && (6.2) \end{aligned}$$

where Ω is a domain with suitably smooth – at least Lipschitz-continuous – boundary $\partial\Omega$ and f, q are prescribed functions. Here L is a linear differential operator and B a boundary operator. We will soon come to some specific nonlinear examples in the context of boundary element techniques.

The usual approach to collocation is then for centres Ξ that are partitioned in two disjoint sets Ξ_1 and Ξ_2, the former from the domain, the latter from its boundary, to solve the Hermite–Birkhoff interpolation system

$$\begin{aligned} \Lambda_\xi u_\Xi &= f(\xi), \quad \xi \in \Xi_1,\\ \Lambda_\zeta u_\Xi &= q(\zeta), \quad \zeta \in \Xi_2. && (6.3) \end{aligned}$$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

The approximants u_Ξ are defined by the sums

$$u_\Xi(x) = \sum_{\xi \in \Xi_1} c_\xi \Lambda_\xi \phi(\|x - \xi\|) + \sum_{\zeta \in \Xi_2} d_\zeta \Lambda_\zeta \phi(\|x - \zeta\|).$$

The Λ_ξ and Λ_ζ are suitable discrete functionals to describe our operators L and B on the discrete set of centres. In the above display the operators are applied with respect to the variable x. Examples for such discrete functionals to replace the operators L and B are obtained, for example, by replacing derivatives by symmetric differences or one-sided differences for the boundary in case of Neumann problems. Thus we end up with a square symmetric system of linear equations whose collocation matrix is nonsingular if, for instance, the radial basis function is positive definite, smooth enough for application of the operators Λ, and the discrete linear functionals are linearly independent functionals in the dual space of the native space of the radial basis functions (see Section 2 and Wu (1992) for the details). An error estimate is given in Franke and Schaback (1998). For those error estimates, it has been noted that more smoothness of the radial basis function is required than for a comparable finite element setting in order to get the same approximation orders, but clearly, the radial basis function setting has the distinct advantage of availability in any dimension and the absence of grids or triangulations.

If a compactly supported radial basis function is used, the necessary scaling leads to the aforementioned trade-off between accuracy and bandwidth of the matrix. In fact, the *conditioning* of the collocation matrix is also affected, becoming worse with smaller η, with $\phi(\cdot/\eta)$ being used, although the matrix is positive definite. A Jacobi preconditioning by the diagonal values helps here, so the matrix A is replaced by $P^{-1}AP^{-1}$ where $P = \sqrt{\mathrm{diag}(A)}$, the diagonal elements of the matrix being positive. Moreover, one can use a *multilevel method* (Narcowich, Schaback and Ward (1999), Fasshauer (1999)) where numerical approximations $\{u_k\}_{k=0}^N$ are computed on nested sets of centres $\Xi_k \supset \Xi_{k-1}$, $k = 1, 2, \ldots, N$, $\Xi_N = \Xi$, and, within each sweep of the algorithm, a new approximation to the desired solution is computed as follows. For instance, to solve $Lu = f$ on the domain Ω with Dirichlet boundary conditions only, at each step k of one sweep one computes, starting with $u_0 = 0$,

$$L\tilde{u}_k = (f - Lu_{k-1})$$

and sets $u_k = \tilde{u}_k + u_{k-1}$. Unfortunately, little is known about the convergence behaviour of such a multilevel method.

In the event that a Galerkin method is applied, for instance, to the Helmholtz equation with Neumann conditions when $L = -\Delta + I$ and $B = \frac{\partial}{\partial n}$, we end up with a square system of linear equations, the stiffness equations for u_Ξ,

$$a(u_\Xi, \phi(\|\cdot - \xi\|)) = (f, \phi(\|\cdot - \xi\|))_{L^2(\Omega)}, \qquad \xi \in \Xi,$$

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

with

$$a(u,v) = \int_\Omega (\nabla u)^T (\nabla v) + uv.$$

If ϕ is a radial basis function of compact support such that its Fourier transform satisfies the decay estimate $|\hat{\phi}(r)| = \mathcal{O}(r^{-2k})$, then Franke and Schaback (1998) establish the convergence estimate

$$\|u - u_\Xi\|_{H^1(\Omega)} \leq Ch^{\sigma-1}\|u\|_{H^\sigma(\Omega)},$$

where h is given by (2.2) and $k \geq \sigma > n/2 + 1$.

We now outline the third method, that is, a boundary element (BEM) method, following Pollandt (1997). The dual reciprocity method uses the second Green's formula and a fundamental solution $\phi(\|\cdot\|)$ of the Laplace operator Δ, in order to reformulate a boundary value problem as a boundary integral problem over a space of one dimension lower. This will then lead through discretization to a linear system with a full matrix for collocation by radial basis functions, in the way familiar from other applications of radial basis function interpolation. The radial basis function that occurs in that context is this fundamental solution, and, naturally, it is highly relevant in this case that the Laplace operator is rotationally invariant and has radial functions as Green's functions. We give a concrete example. Namely, for a nonlinear problem on a domain $\Omega \subset \mathbb{R}^n$ with Dirichlet boundary conditions such as the following one with a nonlinear right-hand side

$$\Delta u(x) = u^2(x), \quad x \in \Omega \subset \mathbb{R}^n, \tag{6.4}$$

$$u|_{\partial\Omega} = q, \tag{6.5}$$

the goal is to approximate the solution u of the elliptic partial differential equation on the domain by g plus a boundary term $\tilde{r}$ that satisfies $\Delta\tilde{r} \equiv 0$ on the domain. To this end, one gets after an application of Green's formula the equation on the boundary that $u(x)$ is the same as

$$\int_\Omega u(y)^2\phi(\|x-y\|)\,\mathrm{d}y - \int_{\partial\Omega} \phi(\|x-y\|)\frac{\partial}{\partial n_y}u(y) - u(y)\frac{\partial}{\partial n_y}\phi(\|x-y\|)\,\mathrm{d}\Gamma_y \tag{6.6}$$

for $x \in \Omega$, where $\frac{\partial}{\partial n_y}$ is the normal derivative with respect to y on $\Gamma = \partial\Omega$. The radially symmetric ϕ is still the fundamental solution of the Laplace operator used in the formulation of the differential equation above. Further, one gets after two applications of Green's formula the equation

$$\frac{1}{2}\,(u(x) - g(x)) + \int_{\partial\Omega} \phi(\|x-y\|) \times \frac{\partial}{\partial n_y}\,(u(y) - g(y)) - (q(y) - g(y))\,\frac{\partial}{\partial n_y}\,\phi(\|x-y\|)\,\mathrm{d}\Gamma_y = 0, \quad x \in \partial\Omega.$$

We will later use this equation to approximate the boundary part of the

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

solution, that is, the part satisfying the boundary conditions. Now we *assume* that there are real coefficients λ_ξ such that the infinite expansion (which will be truncated later on)

$$u^2(y) = \sum_\xi \lambda_\xi \, \widetilde{\phi}\,(\|y-\xi\|), \qquad y \in \Omega,$$

holds, and set

$$g(y) = \sum_\xi \lambda_\xi \, \widetilde{\Phi}\,(\|y-\xi\|), \qquad y \in \Omega,$$

so that $\Delta g = u^2$ everywhere with no boundary conditions. Therefore we are first solving a homogeneous problem. Here $\widetilde{\phi}$ is a suitable radial basis function, which is to be distinguished from the fundamental solution ϕ, with $\Delta\widetilde{\Phi}\,(\|\cdot\|) = \widetilde{\phi}(\|\cdot\|)$, and the centres ξ are from Ω. We now replace the infinite sums by finite ones (*i.e.*, we approximate the infinite expansions by finite sums for a suitable Ξ), so that exchanging summation and the Laplace operator and term-by-term differentiation in the sum is admitted:

$$u^2(y) = \sum_{\xi\in\Xi} \lambda_\xi \, \widetilde{\phi}\,(\|y-\xi\|), \qquad y \in \Omega, \tag{6.7}$$

and

$$g(y) = \sum_{\xi\in\Xi} \lambda_\xi \, \widetilde{\Phi}\,(\|y-\xi\|), \qquad y \in \Omega.$$

We require that the equation in the display after (6.6) holds for finitely many points $x = \zeta_j \in \partial\Omega$, $j = 1, 2, \ldots, t$, only. Then we solve for the coefficients λ_ξ by requiring that (6.7) holds for $y = \xi$, for all $\xi \in \Xi$. This fixes the λ_ξ by interpolation (collocation in the language of differential equations), whereas the equation after (6.6) determines the normal derivative $\frac{\partial}{\partial n_y}\, u(y)$ on $\partial\Omega$, where we are replacing $\frac{\partial}{\partial n_y}\, u(y)$ by another approximant, a polynomial spline $\tau(y)$, for instance. Thus the spline is found by requiring the above equation for all $x = \zeta_j \in \partial\Omega$, $j = 1, 2, \ldots, t$, and choosing a suitable t. Finally, an approximation $\widetilde{u}(x)$ to $u(x)$ is determined on Ω by the identity

$$\widetilde{u}(x) := g(x) + \int_{\partial\Omega} (q(y) - g(y)) \, \frac{\partial}{\partial n_y} \, \phi(\|x-y\|) \, \mathrm{d}\Gamma_y - \int_{\partial\Omega} \phi(\|x-y\|) \, \left(\tau(y) - \frac{\partial g(y)}{\partial n_y} \right) \mathrm{d}\Gamma_y, \quad x \in \Omega, \tag{6.8}$$

where the boundary term $\tilde{r}$ corresponds to the second term on the right-hand side of the display.

Now, all expressions on the right-hand side are known. This is an outline of the approach but we have skipped several important details. Nonetheless, one can clearly see how radial basis functions appear in this algorithm;

https://doi.org/10.1017/S0962492900000015 Published online by Cambridge University Press

indeed, it is most natural to use them here, since many of them are fundamental solutions of Laplace operators or their iterates in certain dimensions. In the above example and $n = 2$, $\phi(r) = (2\pi)^{-1} \log r$, $\widetilde{\phi}(r) = r^2 \log r$ (thin-plate splines) and $\widetilde{\Phi}(r) = \frac{1}{16}\, r^4 \log r - \frac{1}{32}\, r^4$ are the correct choices. As the collocation matrices that appear in the boundary integral equation we are left with in the method are dense and as interpolation is not absolutely necessary in this approach, it can be substituted by quasi-interpolation. Quasi-interpolation has asymptotically essentially the same approximation behaviour (*i.e.*, convergence speed) as interpolation but does not require any data-dependent linear systems to be solved; see also Buhmann (1990*a*). Convergence theorems for the method are available in the paper by Pollandt.

6.2. Other applications

We outline a few further practical applications of radial basis functions. They include, for example, mappings of two- or three-dimensional images such as portraits or underwater sonar scans into other images for comparison. Here interpolation comes into play because some special features of an image may have to be preserved while others need not be mapped exactly, thus enabling a comparison of some features that may differ while at the same time retaining others. Such so-called 'markers' can be, for example, certain points of the skeleton in an x-ray that has to be compared to another one, taken at another time. The same structure appears if we wish to compare sonar scans of a harbour at different times, the rocks being suitable as markers this time. Thin-plate splines turned out to be excellent for such very practical applications (Barrodale and Zala 1997, 1999). Work of this kind led to the invention of the methods for fast evaluation of thin-plate splines and other radial basis functions discussed above, because, after the interpolation, the computed interpolant had to be evaluated on a very fine square grid in two dimensions for analysis or display (Powell 1993).

Measurements of gravitational potential or temperature on the Earth's surface at 'scattered' meteorological stations or measurements on other multidimensional objects, may give rise to interpolation problems that require scattered data. Multiquadric approximations are performing well for this type of use (Hardy 1990). Much work on radial basis functions on spheres which we have only touched upon in this article originates from those applications of geophysics.

Many applications involve high-dimensional interpolation or approximation problems when data are coming through many 'channels', for instance electrodes measuring brain activity from nerve cells. Typical applications from neural physics produce 50–100 dimensional data taken from those electrodes recording measurements from the brain and require post-processing for smoothing, for example. Radial basis functions have been used extens-

ively for this (*e.g.*, Eckhorn (1999), Hochreiter and Schmidhuber (1999), Anderson, Das and Keller (1998)).

The approximation to so-called learning situations by neural networks usually leads to very high-dimensional interpolation problems with scattered data. Girosi (1992) mentions radial basis functions as a very suitable approach to this, partly because of their availability in arbitrary dimensions, and their smoothness. A typical application is in fire detectors. An advanced type of fire detector has to consider several measured parameters, such as colour, spectrum, intensity, movement of an observed object from which it must decide whether, for instance, it is looking at a fire in the room or not, because the apparent fire is reflected sunlight. There is a learning procedure before the implementation of the device, where several prescribed situations (these are the data) are tested and the values zero (no fire) and one (fire) are interpolated, so that the device can 'learn' to interpolate between these standard situations for general cases later when it is used in real life. Radial basis function methods have been tried very successfully for this application because they are excellent tools for high-dimensional problems that will undoubtedly find many more applications in real life, such as polynomial splines have done in at least the last 30 years and still do now.

REFERENCES

M. Abramowitz and I. Stegun, eds (1972), *Handbook of Mathematical Functions*, National Bureau of Standards, Washington.

R. W. Anderson, S. Das and E. L. Keller (1998), 'Estimation of spatiotemporal neural activity using radial basis function networks', *J. Comput. Neuroscience* **5**, 421–441.

N. Arad, N. Dyn and D. Reisfeld (1994), 'Image warping by radial basis functions: applications to facial expressions', *Graphical Models and Image Processing* **56**, 161–172.

R. Askey (1973), 'Radial characteristic functions', MRC Report 1262, University of Wisconsin, Madison.

I. Barrodale and C. A. Zala (1997), 'MJDP–BCS industrial liaison: applications to defence science', in *Approximation and Optimization: Tributes to M. J. D. Powell* (M. D. Buhmann and A. Iserles, eds), Cambridge University Press, Cambridge, pp. 31–46.

I. Barrodale and C. A. Zala (1999), 'Warping aerial photographs to orthomaps using thin plate splines', *Adv. Comput. Math.* **11**, 211–227.

B. J. C. Baxter (1991), 'Conditionally positive functions and p-norm distance matrices', *Constr. Approx.* **7**, 427–440.

R. K. Beatson and W. A. Light (1992), 'Quasi-interpolation in the absence of polynomial reproduction', in *Numerical Methods of Approximation Theory* (D. Braess and L. L. Schumaker, eds), Birkhäuser, Basel, pp. 21–39.

R. K. Beatson and W. A. Light (1997), 'Fast evaluation of radial basis functions: methods for 2-dimensional polyharmonic splines', *IMA J. Numer. Anal.* **17**, 343–372.

R. K. Beatson and G. N. Newsam (1992), 'Fast evaluation of radial basis functions: I', *Comput. Math. Appl.* **24**, 7–19.

R. K. Beatson, G. Goodsell and M. J. D. Powell (1995), 'On multigrid techniques for thin plate spline interpolation in two dimensions', in Vol. 32 of *Lectures in Applied Mathematics*, pp. 77–97.

R. K. Beatson, J. B. Cherrie and C. T. Mouat (1999), 'Fast fitting of radial basis functions: methods based on preconditioned GMRES iteration', *Adv. Comput. Math.* **11**, 253–270.

A. Bejancu (1997), 'The uniform convergence of multivariate natural splines', DAMTP Technical Report, University of Cambridge.

A. Bejancu (1999), 'Local accuracy for radial basis function interpolation on finite uniform grids', *J. Approx. Theory* **99** 242–257.

M. D. Buhmann (1990*a*), 'Multivariate interpolation in odd-dimensional Euclidean spaces using multiquadrics', *Constr. Approx.* **6**, 21–34.

M. D. Buhmann (1990*b*), 'Multivariate cardinal-interpolation with radial-basis functions', *Constr. Approx.* **6**, 225–255.

M. D. Buhmann (1998), 'Radial functions on compact support', *Proceedings of the Edinburgh Mathematical Society* **41**, 33–46.

M. D. Buhmann (2000), 'A new class of radial basis functions with compact support', to appear in *Math. Comput.*

M. D. Buhmann and C. A. Micchelli (1991), 'Multiply monotone functions for cardinal interpolation', *Adv. Appl. Math.* **12**, 359–386.

M. D. Buhmann and C. A. Micchelli (1992), 'On radial basis approximations on periodic grids', *Mathematical Proceedings of the Cambridge Philosophical Society* **112**, 317–334.

J. Duchon (1976), '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.

J. Duchon (1978), 'Sur l'erreur d'interpolation des fonctions de plusieurs variables pars les D^m-splines', *Rev. Française Automat. Informat. Rech. Opér. Anal. Numer.* **12**, 325–334.

J. Duchon (1979), 'Splines minimizing rotation-invariant semi-norms in Sobolev spaces', in *Constructive Theory of Functions of Several Variables* (W. Schempp and K. Zeller, eds), Springer, Berlin/Heidelberg, pp. 85–100.

N. Dyn, F. J. Narcowich and J. D. Ward (1997), 'A framework for interpolation and approximation on Riemannian manifolds', in *Approximation and Optimization: Tributes to M. J. D. Powell* (M. D. Buhmann and A. Iserles, eds), Cambridge University Press, Cambridge, pp. 133–144.

R. Eckhorn (1999), 'Neural mechanisms of scene segmentation: recordings from the visual cortex suggest basic circuits for linking field models', *IEEE Trans. Neural Net.* **10**, 1–16.

R. Estrada (1998), 'Regularization of distributions', *Int. J. Math. Math. Sci.* **21**, 625–636.

G. Fasshauer (1999), 'Solving differential equations with radial basis functions: multilevel methods and smoothing', *Adv. Comput. Math.* **11**, 139–159.

A. C. Faul and M. J. D. Powell (1999), 'Proof of convergence of an iterative technique for thin plate spline interpolation in two dimensions', *Adv. Comput. Math.* **11**, 183–192.

M. Floater and A. Iske (1996), 'Multistep scattered data interpolation using compactly supported radial basis functions', *J. Comput. Appl. Math.* **73**, 65–78.

C. Franke and R. Schaback (1998), 'Solving partial differential equations by collocation using radial basis functions', *Comput. Math. Appl.* **93**, 72–83.

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

W. Freeden and P. Hermann (1986), 'Uniform approximation by harmonic splines', *Math. Z.* **193**, 265–275.

W. Freeden, M. Schreiner and R. Franke (1995), 'A survey on spherical spline approximation', Technical Report 95–157, University of Kaiserslautern.

F. Girosi (1992), 'Some extensions of radial basis functions and their applications in artificial intelligence', *Comput. Math. Appl.* **24**, 61–80.

M. von Golitschek and W. A. Light (2000), 'Interpolation by polynomials and radial basis functions on spheres', to appear in *Constr. Approx.*

L. Greengard and V. Rokhlin (1987), 'A fast algorithm for particle simulations', *J. Comput. Phys.* **73**, 325–348.

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

S. Hochreiter and J. Schmidhuber (1999), 'Feature extraction through Lococode', *Neural Computation* **11**, 679–714.

K. Jetter, J. Stöckler and J. D. Ward (1999), 'Error estimates for scattered data interpolation on spheres', *Math. Comput.* **68**, 733–747.

M. J. Johnson (1997), 'An upper bound on the approximation power of principal shift-invariant spaces', *Constr. Approx.* **13**, 155–176.

M. J. Johnson (1998*a*), 'A bound on the approximation order of surface splines', *Constr. Approx.* **14**, 429–438.

M. J. Johnson (1998*b*), 'On the error in surface spline interpolation of a compactly supported function', manuscript, University of Kuwait.

M. J. Johnson (1999), 'Approximation in $L^p(\mathbb{R}^d)$ from spaces spanned by the perturbed integer translates of a radial function', manuscript, University of Kuwait.

D. S. Jones (1982), *The Theory of Generalised Functions*, Cambridge University Press, Cambridge.

J. Levesley, W. A. Light, D. Ragozin and X. Sun (1999), 'A simple approach to the variational theory for interpolation on spheres', in *New Developments in Approximation Theory* (M. D. Buhmann, M. Felten, D. Mache and M. W. Müller, eds), Birkhäuser, Basel, pp. 117–143.

W. R. Madych and S. A. Nelson (1990), 'Polyharmonic cardinal splines', *J. Approx. Theory* **60**, 141–156.

O. V. Matveev (1997), 'On a method for interpolating functions on chaotic nets', *Math. Notes* **62**, 339–349. Translated from *Mat. Zametki* (1997) **62**, 404–417.

C. A. Micchelli (1986), 'Interpolation of scattered data: distance matrices and conditionally positive definite functions', *Constr. Approx.* **1**, 11–22.
J. K. Misiewicz and D. St. P. Richards (1994), 'Positivity of integrals of Bessel functions', *SIAM J. Math. Anal.* **25**, 596–601.
F. J. Narcowich, R. Schaback and J. D. Ward (1999), Multilevel interpolation and approximation, *Appl. Comput. Harm. Analysis* **7**, 243–261.
R. Pollandt (1997), 'Solving nonlinear equations of mechanics with the boundary element method and radial basis functions', *Internat. J. Numer. Methods Engrg* **40**, 61–73.
M. J. D. Powell (1992*a*), 'The theory of radial basis function approximation in 1990', in *Advances in Numerical Analysis II: Wavelets, Subdivision, and Radial Functions* (W. A. Light, ed.), Oxford University Press, Oxford, pp. 105–210.
M. J. D. Powell (1992*b*), 'Tabulation of thin plate splines on a very fine two-dimensional grid', in *Numerical Methods of Approximation Theory* (D. Braess and L. L. Schumaker, eds), Birkhäuser, Basel, pp. 221–244.
M. J. D. Powell (1993), 'Truncated Laurent expansions for the fast evaluation of thin plate splines', *Numer. Alg.* **5**, 99–120.
M. J. D. Powell (1994), 'The uniform convergence of thin-plate spline interpolation in two dimensions', *Numer. Math.* **67**, 107–128.
M. J. D. Powell (1999), 'Recent research at Cambridge on radial basis functions', in *New Developments in Approximation Theory* (M. D. Buhmann, M. Felten, D. Mache and M. W. Müller, eds), Birkhäuser, Basel, pp. 215–232.
A. Ron and X. Sun (1996), 'Strictly positive definite functions on spheres in Euclidean spaces', *Math. Comput.* **65**, 1513–1530.
S. Saitoh (1988), *Theory of Reproducing Kernels and its Applications*, Longman, Harlow.
R. Schaback and H. Wendland (1998), 'Inverse and saturation theorems for radial basis function interpolation', Technical Report, University of Göttingen.
R. Schaback and Z. Wu (1995), 'Operators on radial basis functions', *J. Comput. Appl. Math.* **73**, 257–270.
I. J. Schoenberg (1942), 'Positive definite functions on spheres', *Duke Math. J.* **9**, 96–108.
R. Sibson and G. Stone (1991), 'Computation of thin plate splines', *SIAM J. Sci. Statist. Comput.* **12**, 1304–1313.
E. M. Stein and G. Weiss (1971), *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton.
H. Wendland (1995), 'Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree', *Adv. Comput. Math.* **4**, 389–396.
H. Wendland (1998), 'Error estimates for interpolation by radial basis functions of minimal degree', *J. Approx. Theory* **93**, 258–272.
H. Wendland (1997), 'Sobolev-type error estimates for interpolation by radial basis functions', in *Surface Fitting and Multiresolution Methods* (A. LeMéhauté and L. L. Schumaker, eds), Vanderbilt University Press, Nashville, pp. 337–344.
Z. Wu (1992), 'Hermite–Birkhoff interpolation of scattered data by radial basis functions', *Approx. Theory Appl.* **8**, 1–10.
Z. Wu (1995), 'Multivariate compactly supported positive definite radial functions', *Adv. Comput. Math.* **4**, 283–292.

Z. Wu and R. Schaback (1993), 'Local error estimates for radial basis function interpolation of scattered data', *IMA J. Numer. Anal.* **13**, 13–27.
Y. Xu and E. W. Cheney (1992), 'Strictly positive definite functions on spheres', *Proc. Amer. Math. Soc.* **116**, 977–981.