Journal of Computational and Applied Mathematics 236 (2011) 434–446

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Better bases for radial basis function interpolation problems

R.K. Beatson [a,*], J. Levesley [b], C.T. Mouat [a]

[a] Department of Mathematics and Statistics, University of Canterbury, Private Bag 8020, Christchurch, New Zealand

[b] Department of Mathematics, University of Leicester, University Road, Leicester, LE17RH, United Kingdom

ARTICLE INFO

Keywords:
RBFs
Bases
Preconditioning
Loss of significance

ABSTRACT

Radial basis function interpolation involves two stages. The first is fitting, solving a linear system corresponding to the interpolation conditions. The second is evaluation. The systems occurring in fitting problems are often very ill-conditioned. Changing the basis in which the radial basis function space is expressed can greatly improve the conditioning of these systems resulting in improved accuracy, and in the case of iterative methods, improved speed, of solution. The change of basis can also improve the accuracy of evaluation by reducing loss of significance errors. In this paper new bases for the relevant space of approximants, and associated preconditioning schemes are developed which are based on Floater's mean value coordinates. Positivity results and scale independence results are shown for schemes of a general type. Numerical results show that the given preconditioning scheme usually improves conditioning of polyharmonic spline and multiquadric interpolation problems in $\mathcal{R}^2$ and $\mathcal{R}^3$ by several orders of magnitude. The theory indicates that using the new basis elements (evaluated indirectly) for both fitting and evaluation will reduce loss of significance errors on evaluation. Numerical experiments confirm this showing that such an approach can improve overall accuracy by several significant figures.

1. Introduction

Radial basis function interpolation involves two stages. The first is fitting, solving a linear system corresponding to the interpolation conditions. The second is evaluation. In this paper we introduce new bases for radial basis interpolation problem which lead to a computationally inexpensive method for preconditioning the linear systems associated with fitting. The new basis, when evaluated indirectly, also greatly improves the accuracy of evaluation of the fitted RBF (radial basis function). We also explore positivity properties of the basis.

Radial basis functions have enjoyed great success in a wide variety of data fitting applications such as surface modelling from point clouds, custom manufacture of artificial limbs, ore grade estimation, and flow modelling. They are particularly advantageous when the data is scattered rather than gridded, the former situation occurring frequently with geophysical data. Unfortunately, as is well known, the matrix of the usual formulation of radial basis function interpolation problems in terms of the natural basis is frequently badly conditioned, even when the number of nodes is small. Indeed many authors have commented on the numerical difficulties that solving this system presents [1–5]. In this paper we develop a computationally inexpensive change of basis based on Floater's mean value coordinates. In the planar case forming the differences underlying the change of basis requires only $\mathcal{O}(N \log N)$ operations, where N is the number of interpolation nodes. This leads naturally to an inexpensive preconditioning method for the interpolation system.

* Corresponding author.
E-mail address: r.beatson@math.canterbury.ac.nz (R.K. Beatson).

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.06.030

A radial basis function (RBF) with centres $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$ is a function of the form

$$s(\cdot) = \sum_{i=1}^{N} \lambda_i \Phi(\cdot - \boldsymbol{x}_i) + c_1 p_1(\cdot) + \cdots + c_\ell p_\ell(\cdot), \tag{1}$$

where Φ is a fixed, usually radial, function and $\{p_1, \ldots, p_\ell\}$ is a basis for π^d_{k-1}. Often the side conditions

$$\sum_{i=1}^{N} \lambda_i q(\boldsymbol{x}_i) = 0, \quad \text{for all } q \in \pi^d_{k-1}, \tag{2}$$

are imposed. These can be viewed either as taking away the extra degrees of freedom created by the polynomial part in (1), or alternatively of enforcing some decay near infinity.

Given a set of data values $\{f_1, \ldots, f_N\}$ corresponding to the centres $\mathcal{X}$, the interpolation problem is to find a function of the form (1) satisfying the side conditions (2) and the interpolation conditions

$$s(\boldsymbol{x}_i) = f_i, \quad 1 \le i \le N.$$

Thus, the standard (pointwise) interpolation problem can be written in matrix form as

$$\begin{bmatrix} A & P \\ P^T & O_\ell \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{0} \end{bmatrix}, \tag{3}$$

where 0_ℓ is the $\ell \times \ell$ zero matrix,

$$A_{ij} = \Phi(\boldsymbol{x}_i - \boldsymbol{x}_j), \qquad P_{ij} = p_j(\boldsymbol{x}_i), \tag{4}$$

and $\boldsymbol{f} = [f_1, \ldots, f_N]^T$.

We will need the following definitions.

Definition 1.1. A set of linear functionals μ_i, $1 \le i \le m$ will be called unisolvent for π^d_{k-1} if

$q \in \pi^d_{k-1}$ and $\mu_j(q) = 0$ for all $1 \le j \le m$ implies q is the zero polynomial.

A set of points $\mathcal{X}$ is said to be unisolvent for π^d_{k-1} when the corresponding set of point evaluations has this property.

Definition 1.2. A continuous function $\Phi : \mathcal{R}^d \to \mathcal{R}$ will be called (pointwise) conditionally positive definite of order k on $\mathcal{R}^d$ if

(i) Φ is even.
(ii) For all choices of a positive integer N and of a set $\mathcal{X}$ of N distinct points in $\mathcal{R}^d$, the quadratic form $\boldsymbol{\lambda}^T A \boldsymbol{\lambda}$ is nonnegative for all vectors $\boldsymbol{\lambda}$ such that

$$\sum_{j=1}^{N} \lambda_j q(\boldsymbol{x}_j) = 0, \quad \text{for all } q \in \pi^d_{k-1}. \tag{5}$$

Φ is called strictly conditionally positive definite of order k if the inequality above is strict whenever $\boldsymbol{\lambda} \neq \boldsymbol{0}$.

It is well known that the matrix

$$A_\Phi = \begin{bmatrix} A & P \\ P^T & O \end{bmatrix} \tag{6}$$

of the usual formulation (3) of the interpolation problem is invertible when Φ is strictly conditionally positive (negative) definite of order k, and the points $\mathcal{X}$ are unisolvent for π^d_{k-1}.

The paper is arranged as follows. In Section 2 a general framework for preconditioning based upon a *complete set of differences* is set out. In Section 3 the framework is applied in the example case of natural cubic spline interpolation in $\mathcal{R}^1$. In Section 4 some positivity and decay properties of basis functions Ψ_j generated by difference preconditioning are discussed. In Section 5 scale independence properties of the preconditioned problem are shown. In Section 6 the specifics of a difference preconditioner based on mean value coordinates are presented. In Section 7 numerical results related to the condition numbers seen when the method is applied to random sets of centres are presented. Finally, in Section 8 we present both theory and numerics showing that the new basis significantly improves accuracy in the combination of fitting and evaluation. A key idea in this section is to evaluate the new basis elements indirectly, rather than directly via the $\Phi(\cdot - \boldsymbol{x}_j)$'s.

2. A general framework for preconditioning

In this section we consider a general framework for preconditioning RBF interpolation problems based on a strictly conditionally positive definite Φ.

We write ℓ for $\dim(\pi^d_{k-1})$ and assume throughout the rest of the paper $\mathcal{X}$ is unisolvent for π^d_{k-1}. This unisolvency is a necessary and sufficient condition for the matrix P occurring in the usual formulation of RBF interpolation to have full rank.

We will precondition using matrices Q whose columns correspond to difference functionals. The difference functionals are somewhat hidden in the previous works [6,7] but emphasizing them makes many arguments to follow more transparent.

Definition 2.1. A functional Δ_j with form

$$\Delta_j g = \sum_{i=1}^{N} q_{ij} g(\boldsymbol{x}_i), \quad \text{for all } g : \mathcal{R}^d \to \mathcal{R}, \tag{7}$$

for some constants $\{q_{ij}\}$, will be called a *difference functional* (based on nodes $\mathcal{X}$).

Definition 2.2. Let $\ell = \dim(\pi^d_{k-1})$. A set $\mathcal{G} = \{\Delta_1, \ldots, \Delta_{N-\ell}\}$, of difference functionals will be called a k-complete set of difference functionals if

(1) Each functional Δ_j in the set annihilates π^d_{k-1}.
(2) The set of difference functionals $\mathcal{G}$ has full rank in the sense that the $N \times (N-\ell)$ matrix Q with jth column the coefficients of the difference functional Δ_j has rank $N - \ell$.

Remark 2.1. A consequence of the definition is that the columns of Q form a basis for the orthogonal complement of the column space of the matrix P, with $P_{ij} = p_j(\boldsymbol{x}_i)$, occurring in the usual formulation of RBF interpolation.

Then with A defined as in (4) AQ is the matrix with rj entry

$$\sum_{i=1}^{N} \Phi(\boldsymbol{x}_r - \boldsymbol{x}_i) q_{ij} = \Delta_j^{\boldsymbol{y}} \Phi(\boldsymbol{x}_r - \boldsymbol{y}),$$

where the notation $\Delta_j^{\boldsymbol{y}}$ indicates the functional Δ_j operating on the $\boldsymbol{y}$ variable. Now $B = Q^T AQ$ has ij entry

$$B_{ij} = \sum_{r=1}^{N} q_{ri} (AQ)_{rj} = \Delta_i^{\boldsymbol{x}} \Delta_j^{\boldsymbol{y}} \Phi(\boldsymbol{x} - \boldsymbol{y}). \tag{8}$$

The general framework for preconditioning takes the form

Outline algorithm:

Given a set $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ of points in $\mathcal{R}^d$ unisolvent for π^d_{k-1} and a function Φ which is strictly conditionally positive definite of order k

Step 1: Choose, or construct, a k-complete set of difference functionals $\{\Delta_1, \ldots, \Delta_{N-\ell}\}$.

Step 2: Letting $\boldsymbol{\lambda} = Q\boldsymbol{\mu}$ and premultiplying (3) by Q^T gives the new symmetric positive definite system which could be solved for $\boldsymbol{\mu}$, or equivalently $\boldsymbol{\lambda}$,

$$B\boldsymbol{\mu} = Q^T \boldsymbol{f} \quad \text{where } B = Q^T AQ. \tag{9}$$

Here B is positive definite since, for $\boldsymbol{\mu} \neq \mathbf{0}$,

$$\boldsymbol{\mu}^T B \boldsymbol{\mu} = \boldsymbol{\mu}^T Q^T AQ \boldsymbol{\mu} = \boldsymbol{\lambda}^T A \boldsymbol{\lambda} > 0,$$

where in the last step we have used that $\boldsymbol{\lambda} \neq \mathbf{0}$, that $P^T \boldsymbol{\lambda} = \left(P^T Q\right) \boldsymbol{\mu} = \mathbf{0}$, and the strict conditional positive definiteness of Φ.

Step 3: Equilibrate, that is perform diagonal scaling on B by letting D be the $(N-\ell) \times (N-\ell)$ diagonal matrix with $D_{ii} = 1/B_{ii}$. Set $\widetilde{Q} = QD^{1/2}$ and

$$S = D^{1/2} B D^{1/2} = \widetilde{Q}^T A \widetilde{Q}.$$

Step 4: Solve for $\boldsymbol{\lambda}$ by first solving $S\boldsymbol{\mu} = \widetilde{Q}^T \boldsymbol{f}$ for $\boldsymbol{\mu}$. Then set $\boldsymbol{\lambda} = \widetilde{Q}\boldsymbol{\mu}$.

Step 5: Find the ℓ polynomial coefficients by choosing $c_1, \ldots, c_\ell$ so that

$$s(\boldsymbol{x}) = c_1 p_1 + \cdots + c_\ell p_\ell + \sum_{j=1}^{N} \lambda_j \Phi(\boldsymbol{x} - \boldsymbol{x}_j),$$

interpolates to f at any ℓ points from $\mathcal{X}$ which are unisolvent for π^d_{k-1}. This involves solving an $\ell \times \ell$ linear system. This procedure is possible since $\boldsymbol{f} - A\boldsymbol{\lambda}$ is in the column space of P, that is $\widetilde{Q}^{\perp}$, as

$$\widetilde{Q}^T (\boldsymbol{f} - A\boldsymbol{\lambda}) = \widetilde{Q}^T \boldsymbol{f} - \widetilde{Q}^T A \widetilde{Q} \boldsymbol{\mu} = \mathbf{0}.$$

Different preconditioners of this overall form have been previously explored for example in [1,8].

Within the construction we can view the functions

$$\Psi_i(\cdot) = \Delta_i^{\boldsymbol{y}} \Phi(\cdot - \boldsymbol{y}), \quad 1 \leq i \leq N - \ell,$$

as part of a new, hopefully better, basis for the space from which RBF interpolants will be drawn. The basis will be completed by adding the ℓ elements of some basis for the polynomials. The Ψ functions already satisfy the constraints (2), whereas for $k > 0$ the functions $\Phi(\cdot - \boldsymbol{x}_i)$ do not and hence the $\Phi(\cdot - \boldsymbol{x}_i)$ functions do not even belong to the RBF space.

The choice remaining within the macro algorithm is the particulars of the construction of the difference functionals. In order for any scheme to be practical the differences should be reasonably cheap to construct. The schemes to be presented meet this criteria. For example basing the scheme on a Delaunay triangulation forming the differences will take only $\mathcal{O}(N \log N)$ operations in the planar case. Further, in our view, the differences should be as local as possible. A motivation for this being the results of the next two sections which show that local support of the difference Δ_i leads to *approximate locality* in the corresponding new basis elements Ψ_i.

3. Example: cubic splines in one dimension

In this section we apply the macro algorithm of the previous section to natural cubic spline interpolation, viewed as RBF interpolation with $\Phi(\boldsymbol{x}) = |\boldsymbol{x}|^3$. We will see that this preconditioning method is very successful on this example.

In one dimension cubic spline interpolation can be viewed as RBF interpolation built upon the basic function $\Phi(x) = |x|^3$, a strictly conditionally positive definite function of order 2. Viewed this way the natural cubic spline interpolant based upon points $x_1 < x_2 < \cdots < x_N$ is an interpolant of the form

$$s = q_1 + \sum_{j=1}^{N} a_j \Phi(\cdot - x_j),$$

where $q_1 \in \pi_1$ and the weights satisfy the constraints

$$\sum_{j=1}^{N} a_j = \sum_{j=1}^{N} a_j x_j = 0.$$

The interpolation problem posed this way in terms of the *bad basis* will have matrix of form (3), where the matrix A will be full. Thus the expression of the problems in the natural way in terms of Φ functions is very much suboptimal. It is essentially the same as expressing the spline in terms of the powers and truncated powers. This truncated power expression can be useful for theory. However, expressing the problem in terms of B-splines is vastly superior for computations in that it yields a tridiagonal system with bounded condition number. The preconditioning scheme described in the previous section greatly improves the RBF matrix system for the cubic spline problem yielding a tridiagonal system with bounded condition number.

In the cubic spline case we take Δ_j as a second divided difference

$$[x_j, x_{j+1}, x_{j+2}]g = \frac{g(x_{j+2})}{(x_{j+2} - x_j)(x_{j+2} - x_{j+1})} + \frac{g(x_j)}{(x_{j+2} - x_j)(x_{j+1} - x_j)} - \frac{g(x_{j+1})}{(x_{j+2} - x_{j+1})(x_{j+1} - x_j)}, \quad 1 \le j \le N - 2.$$

Thus $q_{j,j}$, $q_{j+1,j}$, $q_{j+2,j}$ are the only nonzero entries in the jth column of the $N \times (N-2)$ matrix Q. Hence, Q is zero above its main diagonal and if b_t is the first nonzero entry in the $N-2$ vector $\boldsymbol{b}$ then

$$(Q\boldsymbol{b})_t = q_{t,t} b_t \neq 0.$$

Consequently Q has full rank. Therefore applying the analysis of the previous section the matrix $B = Q^T A Q$ is symmetric positive definite. In this case

$$B_{ij} = [x_i, x_{i+1}, x_{i+2}]^x [x_j, x_{j+1}, x_{j+2}]^y \Phi(x - y).$$

But a suitably normalized linear B-spline, M_i, nonzero on (x_i, x_{i+2}), is the Peano kernel of the second divided difference. Therefore for $x \ge x_{j+2}$

$$[x_j, x_{j+1}, x_{j+2}]^y \Phi(x - y) = \int_{x_j}^{x_{j+2}} M_j(y) \Phi''(x - y)\, \mathrm{d}y = \int_{x_j}^{x_{j+2}} M_j(y) 6(x - y)\, \mathrm{d}y = \alpha_j x + \beta_j.$$

Hence for $i \ge j + 2$, $B_{ij} = [x_i, x_{i+1}, x_{i+2}] \left\{\alpha_j x + \beta_j\right\} = 0$. Similar arguments show $B_{ij} = 0$ for $i \le j - 2$. It follows that B, and its equilibrated form S, are tridiagonal.

4. Some positivity and decay properties

In this section we explore positivity properties of the new basis. These go together with decay properties arising from the differencing. Throughout $|\cdot|$ means the 2-norm.

Lemma 4.1. *Let $\phi \in C[0,\infty)$ be convex and nondecreasing. Then $\Phi(\boldsymbol{x}) = \phi(|\boldsymbol{x}|)$ is a convex function from $\mathcal{R}^d$ into $\mathcal{R}$.*

Proof. For $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{R}^d$, and $0 \le t \le 1$

$$\begin{aligned}\Phi\left((1-t)\boldsymbol{x}+t\boldsymbol{y}\right) &= \phi(|(1-t)\boldsymbol{x}+t\boldsymbol{y}|)\\ &\le \phi\left((1-t)|\boldsymbol{x}|+t|\boldsymbol{y}|\right),\end{aligned}$$

where in the last step we have used the triangle inequality and that ϕ is nondecreasing. Then applying the convexity of ϕ

$$\begin{aligned}\Phi\left((1-t)\boldsymbol{x}+t\boldsymbol{y}\right) &\le (1-t)\phi(|\boldsymbol{x}|)+t\phi(|\boldsymbol{y}|)\\ &= (1-t)\Phi(\boldsymbol{x})+t\Phi(\boldsymbol{y}).\end{aligned}$$

Hence Φ is convex. □

Lemma 4.1 implies that the generalized multiquadric $\Phi(\boldsymbol{x}) = (|\boldsymbol{x}|^2 + c^2)^\beta$ is convex for $\beta \ge 1/2$.

In what follows $\mathcal{H}(\mathcal{A})$ denotes the convex hull of the set $\mathcal{A}$.

The lemma below shows that combining shifts of a convex basic function Φ with a certain type of difference functional leads to a nonnegative function Ψ.

Lemma 4.2 (*Positivity and Approximate Laplacians*)**.** *Suppose $\boldsymbol{x}_0 \in \mathcal{H}(\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_s\}) \subset \mathcal{R}^d$. Further suppose*

$$\theta_i \ge 0 \quad \text{for all } i, \sum_{i=1}^{s}\theta_i = 1, \quad \text{and} \quad \boldsymbol{x}_0 = \sum_{i=1}^{s}\theta_i\boldsymbol{x}_i.$$

Let $\Phi : \mathcal{R}^d \to \mathcal{R}$ be convex. Define

$$\Psi(\boldsymbol{x}) = \triangle^{\boldsymbol{y}}\Phi(\boldsymbol{x}-\boldsymbol{y}) := \left\{\sum_{i=1}^{s}\theta_i\Phi(\boldsymbol{x}-\boldsymbol{x}_i)\right\} - \Phi(\boldsymbol{x}-\boldsymbol{x}_0).$$

Then $\Psi(\boldsymbol{x}) \ge 0$ for all $\boldsymbol{x}$.

Proof. The lemma is an immediate consequence of Jensen's inequality. More explicitly, from the hypotheses

$$\boldsymbol{x}-\boldsymbol{x}_0 = \sum_{j=1}^{s}\theta_j\left(\boldsymbol{x}-\boldsymbol{x}_j\right), \quad \text{for all } \boldsymbol{x}.$$

Then defining $\boldsymbol{y}_j = \boldsymbol{x}-\boldsymbol{x}_j$ the convexity of Φ implies

$$\Phi\left(\sum_{j=1}^{s}\theta_j\boldsymbol{y}_j\right) \le \sum_{j=1}^{s}\theta_j\Phi(\boldsymbol{y}_j). \quad \square$$

The positivity of the Ψ function can be seen in Fig. 1.

The combinations in the lemma above can be viewed as difference functionals $\Delta_j = \sum_{t=0}^{N} q_{tj}\delta_{\boldsymbol{x}_t}$ applied to $\Phi(\boldsymbol{x}-\boldsymbol{y})$ viewed as a function of $\boldsymbol{y}$. These functionals annihilate $p(\boldsymbol{y})$ whenever p is a linear polynomial. Thus they are in a sense generalized second differences. Recalling that second derivatives applied to homogeneous functions lower the asymptotic rate of growth by 2 we expect these generalized differences to do the same. Proofs can be based on far field expansions where the difference will usually annihilate the first few coefficients when terms are grouped in decreasing order of growth at infinity. Thus, for the ordinary multiquadric nonzero differences annihilating π_1^d will give a $\Psi_j(|\boldsymbol{x}|)$ of growth $\mathcal{O}(|\boldsymbol{x}|^{-1})$ at infinity, and for the ordinary thin-plate spline will give a $\Psi_j(\boldsymbol{x})$ of growth $\mathcal{O}(|\log\boldsymbol{x}|)$ at infinity. One can clearly see the $\mathcal{O}(|\boldsymbol{x}|^{-1})$ growth in Fig. 1(b) above. The corresponding figure for the ordinary thin-plate spline, Fig. 2, shows slow growth in $|\Psi(\boldsymbol{x})|$ for large $|\boldsymbol{x}|$, far slower than the $\mathcal{O}(|\boldsymbol{x}|^2\log|\boldsymbol{x}|)$ growth of $\Phi(\boldsymbol{x}-\boldsymbol{x}_i)$. These decay phenomena were discussed in [7]. They were rediscovered in [9].

The combinations discussed above lead to collections of functions Ψ_i that decay much faster at infinity than the original functions $\Phi_j(\boldsymbol{x}) = \Phi(\boldsymbol{x}-\boldsymbol{x}_j)$ did. For the multiquadrics these functions Ψ_i are even nonnegative. In $\mathcal{R}^1$ the decay is sometimes even fast enough so that $\{\Psi_i\}$ forms a partition of unity; see [10]. Unfortunately in $\mathcal{R}^d$ with $d > 1$ and scattered data things are more difficult and the decay is usually insufficient to form infinite partitions of unity.

The use of these basis elements in preconditioning is discussed in the sections below. To gain full benefit from them loss of significance errors upon evaluation must be reduced by evaluating them indirectly, rather than directly in terms of the $\Phi(\cdot-\boldsymbol{x}_j)$'s (see Section 8).

5. Scalability

In this section we show that for certain functions Φ, including the polyharmonic splines, the preconditioned interpolation matrix $S = \widetilde{Q}^T A\widetilde{S}$ produced by any interpolation scheme of the type described in Section 2 is independent of the scale.

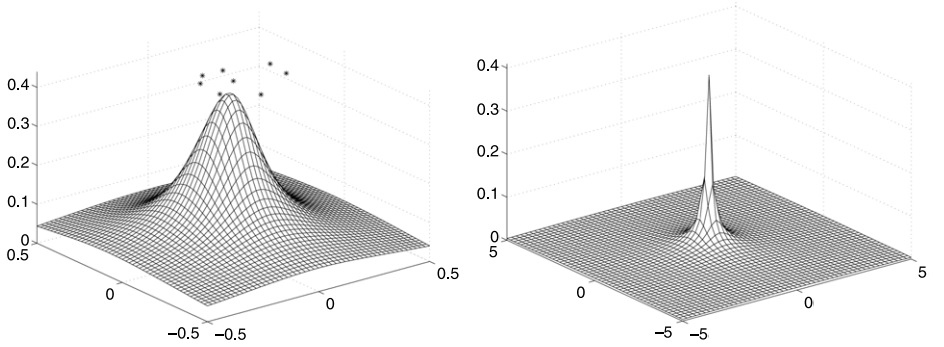

Fig. 1. Two views of a Ψ-element formed by combining shifts of the ordinary multiquadric. Note the nonnegativity and the $\mathcal{O}(|x|^{-1})$ decay at infinity.

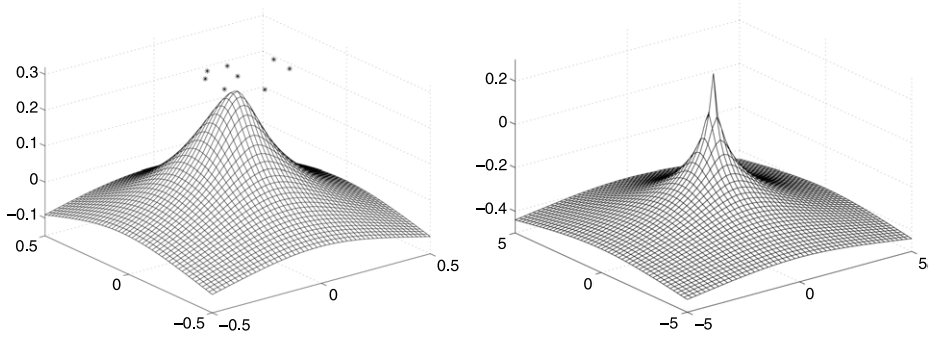

Fig. 2. Two views of a Ψ-element formed by combining shifts of the ordinary TPS basic function $|\boldsymbol{x}|^2 \log |\boldsymbol{x}|$. Note the slow growth in $|\Psi(\boldsymbol{x})|$ as $|\boldsymbol{x}| \to \infty$.

Consequently its condition number, and the relative spread of its eigenvalues, are also scale independent. This is a very desirable property for applications, where the units used should not impact the quality of the final fit.

The Beppo–Levi space $\mathcal{B}L^k(\mathcal{R}^d)$ is the space of all functions f such that for each multiindex α with $|\alpha| = k$, the distributional derivative $D^{\boldsymbol{\alpha}} f \in L^2(\mathcal{R}^d)$. The solution of the polyharmonic spline interpolation problem can be characterized as the unique solution of the variational problem.

Problem 5.1. Given nodes $\mathcal{X}$ unisolvent for π^d_{k-1} and function values $\{f_1, \ldots, f_N\}$ find a function s_h in the Beppo–Levi space $\mathcal{B}L^k(\mathcal{R}^d)$ minimizing

$$E(g) = \sum_{|\boldsymbol{\alpha}|=k} \binom{k}{\boldsymbol{\alpha}} \int_{\mathcal{R}^d} \left(D^{\boldsymbol{\alpha}} g\right)^2 \mathrm{d}\boldsymbol{x},$$

over all functions $g \in \mathcal{B}L^k(\mathcal{R}^d)$ which take the values $\{f_1, \ldots, f_N\}$ at the points of $h\mathcal{X}$.

It is clear from this formulation that the solution is independent of the scale. That is the solution s_h for nodes $h\mathcal{X}$, and the solution s_1 for nodes $\mathcal{X}$, transform into each other simply by scaling the domain. More precisely, $s_1(\cdot) = s_h(h\cdot)$.

Unfortunately, the usual formulation (3) used for numerically fitting interpolatory polyharmonic splines does not share this desirable scale independence. Indeed the condition number of the matrix involved in fitting the RBF can vary wildly under uniform scaling; see e.g. Table 2 of [8]. It is important to avoid such scale dependent bad conditioning for fitting methods such as domain decomposition (see e.g. [8]) and the two-stage method (see e.g. [11]) where solutions to systems on many different scales are required. In contrast the method presented in Section 2 is scale independent when applied to polyharmonic splines, as will be shown in Corollary 5.6.

Lemma 5.2. *Let $\mathcal{X} = \{x_1, \ldots, x_N\}$ be unisolvent with respect to π^d_{k-1}. Suppose the difference functionals Δ_i and Δ_j of form (7) annihilate π^d_{k-1}. Define $T : C(\mathcal{R}^d) \to \mathcal{R}$ by*

$$Tg = \Delta_i^{\boldsymbol{x}} \Delta_j^{\boldsymbol{y}} g(\boldsymbol{x} - \boldsymbol{y}). \tag{10}$$

Then T annihilates π^d_{2k-1}.

Proof. Consider $p_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \boldsymbol{x}^{\boldsymbol{\alpha}}$, where $\boldsymbol{x} \in \mathcal{R}^d$ and $\boldsymbol{\alpha} \in \mathcal{Z}^d_+$ with $|\boldsymbol{\alpha}| < 2k$. From the Binomial Theorem we have, for some numbers a_α,

$$p_{\boldsymbol{\alpha}}(\boldsymbol{x}-\boldsymbol{y}) = \sum_{0\le\boldsymbol{\beta}\le\boldsymbol{\alpha}} a_{\boldsymbol{\beta}}\boldsymbol{x}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}\boldsymbol{y}^{\boldsymbol{\beta}} = \sum_{0\le\boldsymbol{\beta}\le\boldsymbol{\alpha}} a_{\boldsymbol{\beta}} p_{\boldsymbol{\alpha}-\boldsymbol{\beta}}(\boldsymbol{x}) p_{\boldsymbol{\beta}}(\boldsymbol{y}), \quad \boldsymbol{x},\boldsymbol{y}\in\mathcal{R}^d.$$

Therefore

$$\begin{aligned} T\,p_{\boldsymbol{\alpha}} &= T\left\{\sum_{0\le\boldsymbol{\beta}\le\boldsymbol{\alpha}} a_{\boldsymbol{\beta}} p_{\boldsymbol{\alpha}-\boldsymbol{\beta}}(\boldsymbol{x}) p_{\boldsymbol{\beta}}(\boldsymbol{y})\right\} \\ &= \sum_{0\le\boldsymbol{\beta}\le\boldsymbol{\alpha}} a_{\boldsymbol{\beta}}\left\{\Delta_i^{\boldsymbol{x}} p_{\boldsymbol{\alpha}-\boldsymbol{\beta}}(\boldsymbol{x})\right\}\left\{\Delta_j^{\boldsymbol{y}} p_{\boldsymbol{\beta}}(\boldsymbol{y})\right\}. \end{aligned}$$

From the hypothesis, and because either $|\boldsymbol{\alpha}-\boldsymbol{\beta}| \le k-1$ or $|\boldsymbol{\beta}| \le k-1$, either $\Delta_i p_{\boldsymbol{\alpha}-\boldsymbol{\beta}}$ or $\Delta_j p_{\boldsymbol{\beta}}$ is zero. Hence $Tp_{\boldsymbol{\alpha}}$ is zero. The result follows since $\{p_{\boldsymbol{\alpha}} : 0 \le |\boldsymbol{\alpha}| < 2k\}$ is a basis for π^d_{2k-1}. □

Theorem 5.3. *Let the continuous even function $\Phi : \mathcal{R}^d \to \mathcal{R}$ be such that $\Phi(h\boldsymbol{x}) = h^\gamma \Phi(\boldsymbol{x}) + p_h(\boldsymbol{x})$ for all $h > 0$ and $\boldsymbol{x} \in \mathcal{R}^d$, where $\gamma \in \mathcal{R}$ and $p_h \in \pi^d_{2k-1}$. Let $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be a unisolvent set of points with respect to π^d_{k-1} and let Δ_i, Δ_j be as in Lemma 5.2. Define differences for the scale $h > 0$ by*

$$\Delta_{j,h} g = \sigma(h)\Delta_j^{\boldsymbol{x}} g(h\cdot), \quad 1 \le j \le N-\ell,$$

where $\sigma(h) \ne 0$. Define the functional $T_h\Phi$ by

$$T_h\Phi = \Delta_{i,h}^{\boldsymbol{x}}\Delta_{j,h}^{\boldsymbol{y}}\Phi(\boldsymbol{x}-\boldsymbol{y}) = \sigma(h)^2\Delta_i^{\boldsymbol{x}}\Delta_j^{\boldsymbol{y}}\Phi(h\boldsymbol{x}-h\boldsymbol{y}),$$

and write T for T_1. Then for $h > 0$, $T_h\Phi = \sigma(h)^2 h^\gamma T\Phi$.

Remark 5.4. The differences $\Delta_{i,h}$ operate on function values at the points $h\mathcal{X}$ rather than function values at the points $\mathcal{X}$. The weights of these differences are scaled by some nonzero quantity $\sigma(h)$. It will be seen that the diagonal scaling means the exact form of the function $\sigma(h)$ has no influence.

Proof. From the definition we have

$$\begin{aligned} T_h\Phi &= \sigma(h)^2\Delta_i^{\boldsymbol{x}}\Delta_j^{\boldsymbol{y}}\Phi(h\boldsymbol{x}-h\boldsymbol{y}), \\ &= \sigma(h)^2\Delta_i^{\boldsymbol{x}}\Delta_j^{\boldsymbol{y}}\left\{h^\gamma\Phi(\boldsymbol{x}-\boldsymbol{y}) + p_h(\boldsymbol{x}-\boldsymbol{y})\right\}, \\ &= \sigma(h)^2 h^\gamma T\Phi + q(h)^2 T p_h, \end{aligned}$$

for some $p_h \in \pi^d_{2k-1}$ where T is as in Lemma 5.2. From that lemma $Tp_h = 0$ and the theorem follows. □

Let the columns of the $N \times (N-\ell)$ matrix Q_h be formed from the coefficients in the differences $\Delta_{1,h}, \ldots, \Delta_{N-\ell,h}$. Let A_h be the $N \times N$ matrix with ij-entry $\Phi(h\boldsymbol{x}_i - h\boldsymbol{x}_j)$. The corresponding matrices with scale $h = 1$ are defined in Section 2 and denoted by Q and A.

Theorem 5.5. *Let $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be unisolvent with respect to π^d_{k-1}, and suppose the differences $\Delta_1, \ldots, \Delta_{N-\ell}$, form a k-complete set of differences. Define differences $\Delta_{i,h}$ for scale $h > 0$ as in Theorem 5.3 and also suppose the function Φ is as in that theorem. Then*

(i)

$$B_h := Q_h^T A_h Q_h = \sigma(h)^2 h^\gamma B_1 = \sigma(h)^2 h^\gamma B.$$

(ii) *Let S_h be produced from B_h by equilibration. Then S_h is a constant matrix independent of h.*

Proof. From Theorem 5.3 and the condition on Φ

$$\Delta_{i,h}^{\boldsymbol{x}}\Delta_{j,h}^{\boldsymbol{y}}\Phi = \sigma(h)^2 h^\gamma \Delta_i^{\boldsymbol{x}}\Delta_j^{\boldsymbol{y}}\Phi(\boldsymbol{x}-\boldsymbol{y}), \quad 1 \le i,j \le N-\ell,$$

which is the componentwise form of the first statement of the theorem.

Turn now to the second part the theorem. Since the equilibrated matrix is $S_h = D_h B_h D_h$ where D_h is diagonal with $(D_h)_{ii} = (B_h)_{ii}^{-1/2}$,

$$\begin{aligned} (S_h)_{ij} &= (D_h)_{ii}(B_h)_{ij}(D_h)_{jj} \\ &= \frac{1}{\sigma(h)^2 h^\gamma\sqrt{B_{ii}B_{jj}}}\sigma(h)^2 h^\gamma B_{ij} \\ &= \frac{B_{ij}}{\sqrt{B_{ii}B_{jj}}}, \end{aligned}$$

which is independent of h. □

Corollary 5.6. *Any preconditioning method for strictly conditionally positive definite kernels of order k of the type described in Theorem 5.5 applied to the basic functions*

$$\Phi(\boldsymbol{x}) = (-1)^{\lceil \beta/2 \rceil} |\boldsymbol{x}|^{\beta}, \quad \text{with } 0 \le \beta < 2k,\ \beta \notin 2\mathcal{N}, \tag{11}$$

or

$$\Phi(\boldsymbol{x}) = (-1)^{j+1} |\boldsymbol{x}|^{2j} \log |\boldsymbol{x}|, \quad \text{with } j \in \mathcal{N},\ 1 \le j < k, \tag{12}$$

produces for the nodes $h\mathcal{X}$ an equilibrated interpolation matrix S which is independent of the scale $h > 0$.

Remark 5.7. This corollary covers all the popular thin-plate spline kernels for $\mathcal{R}^d$.

Proof. Consider first the power kernel with parameter β of Eq. (11). It follows from a result of Micchelli [12] that this kernel is strictly conditionally positive definite of order $\lceil \beta/2 \rceil$. Theorem 5.5 then applies with $\gamma = \beta$ and p_h the zero polynomial.

Now consider the basic function $\Phi(\cdot) = (-1)^{j+1} |\cdot|^{2j} \log|\cdot|$, $1 \le j < k$. It follows from results of Micchelli [12] that Φ is strictly conditionally positive definite of order $j+1$, and therefore k, on $\mathcal{R}^d$. Note that,

$$\Phi(h\boldsymbol{x}) = h^{2j}\Phi(\boldsymbol{x}) + (-1)^{j+1} h^{2j} \log(h) |\boldsymbol{x}|^{2j}. \tag{13}$$

Then applying Theorem 5.5 with $\gamma = 2j$ gives the result. □

6. Preconditioning using mean value coordinates

In this section we discuss the construction of a particular preconditioning scheme of the type discussed in Section 2. In this scheme the differences are derived from Floater's mean value coordinates. The construction is appealing in that for "interior" points $\boldsymbol{x}_j$ of $\mathcal{X}$ it is local. That is, for such points the difference functional Δ_j and the entries in the jth column of Q, depend only on the geometry of the nodes near $\boldsymbol{x}_j$ and not on any properties of nodes far away. The discussion below concentrates on preconditioning interpolation problems in $\mathcal{R}^2$. The generalization to problems in $\mathcal{R}^3$ is very similar and the details will be omitted.

Mean value coordinates are generalized barycentric coordinates for polygons with an arbitrary number of sides. They were originally introduced by Floater and have been explored in a series of papers; see [13–15]. We will consider their application to preconditioning RBF interpolation equations.

A competitor to the mean value coordinates for the preconditioning application are the boundary over distance weights of Sibson and Stone [1], explored in [6,7]. The boundary over distance method works extremely well. However, a mean value coordinate based preconditioner has some advantages.

- The boundary over distance preconditioner requires a Voronoi tessellation, equivalently a Delaunay triangulation, of the points. The mean value coordinate based preconditioner does not require a Delaunay triangulation, or indeed any triangulation. Therefore it may have advantages in problems where the mesh moves with time.
- The new basis elements Ψ_j associated with an interior point in the case of the boundary over distance preconditioner involve $\Phi(\cdot - \boldsymbol{x}_i)$'s corresponding to centres in the Voronoi neighbours. In the mean value coordinate case the centres generating $\Phi(\cdot - \boldsymbol{x}_i)$'s involved with a new basis element Ψ_j can be chosen according to any heuristic one likes.

The kernel of a polygon is the set of points $\boldsymbol{v}$ such that for each vertex $\boldsymbol{v}_i$ the line segment $[\boldsymbol{v}, \boldsymbol{v}_i]$ is a subset of the closure of the interior of the polygon. A key property of mean value coordinates is that if $[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ are the vertices in counterclockwise order of a starshaped non-self-intersecting polygon, and $\boldsymbol{v}$ is in the kernel of that polygon, then the mean values coordinates $\lambda_j(\boldsymbol{v})$ satisfy

(i) $\lambda_j(\boldsymbol{v}) \ge 0$ for all $1 \le j \le n$.
(ii) $\sum_{j=1}^{n} \lambda_j(\boldsymbol{v}) = 1$.
(iii) $\sum_{j=1}^{n} \lambda_j(\boldsymbol{v}) \boldsymbol{v}_j = \boldsymbol{v}$.

It is clear that given an interior vertex $\boldsymbol{v}$ of a triangulation the polygon formed by joining vertices in its one ring in counterclockwise order is starshaped with $\boldsymbol{v}$ in the kernel. Here the one ring of a vertex is the set of all vertices of the triangulation one edge away from the given vertex as illustrated in Fig. 3(a). Sometimes the work of forming a triangulation may be too much, or it may be desirable to force a specific choice of the neighbours $\boldsymbol{v}_j$, in terms of which $\boldsymbol{v}$ is expressed. Therefore the question arises of when an (unordered) set of nearby points is suitable for the mean value coordinate construction. An answer is given below.

Observation 6.1. *Consider a finite set $\mathcal{V} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ of distinct points in $\mathcal{R}^2$. Let $\boldsymbol{v} \notin \mathcal{V}$ be in the convex hull of $\mathcal{V}$. Order the points in $\mathcal{V}$ in counterclockwise order about $\boldsymbol{v}$, and within that sort (i.e. for points at the same angle) by increasing radial distance from $\boldsymbol{v}$. Label the ordered points as $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$. Then the polygon $[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ with edges $[\boldsymbol{v}_1, \boldsymbol{v}_2]$, $[\boldsymbol{v}_2, \boldsymbol{v}_3], \ldots, [\boldsymbol{v}_{n-1}, \boldsymbol{v}_n], [\boldsymbol{v}_n, \boldsymbol{v}_1]$, is starlike and non-self-intersecting with $\boldsymbol{v}$ in its kernel.*

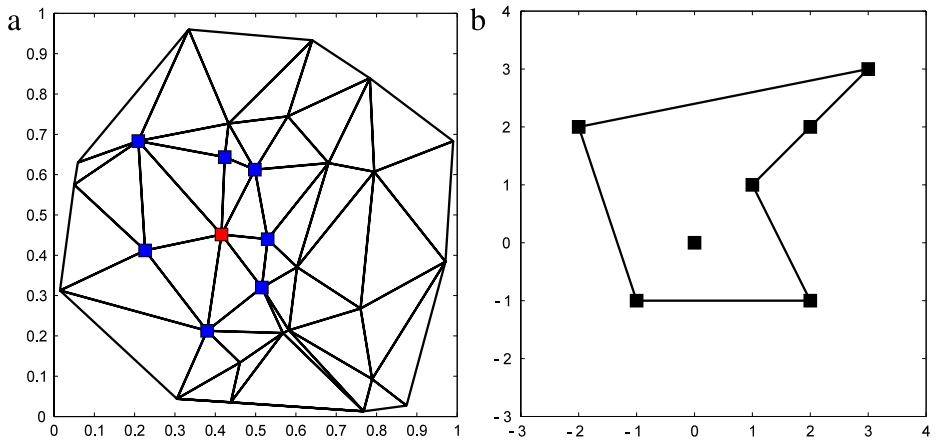

Fig. 3. (a) The one ring of a vertex in a triangulation. (b) A typical geometry of nodes in the proof of Observation 2.

Remark 6.2. If $\boldsymbol{v}$ is not in the convex hull of $\mathcal{V}$ then it is easy to construct examples where the polygon constructed from the edges $[\boldsymbol{v}_1, \boldsymbol{v}_2]$, $[\boldsymbol{v}_2, \boldsymbol{v}_3]$, $\ldots$, $[\boldsymbol{v}_n, \boldsymbol{v}_1]$ is self-intersecting.

Proof. Assume without loss of generality that $\boldsymbol{v} = 0$. Write each $\boldsymbol{v}_j$ as a complex number

$$\boldsymbol{v}_j = r_j e^{i\theta_j}, \quad r_j > 0, \text{ and } 0 \le \theta_j < 2\pi.$$

Further, assume without loss of generality that

$$0 \le \theta_1 \le \theta_2 \le \cdots \le \theta_n.$$

This results in a layout of nodes as illustrated in Fig. 3(b).

We treat the subscripts in a circular manner. Therefore when $j = 1$, $\boldsymbol{v}_{j-1}$ is $\boldsymbol{v}_n$, and when $j = n$, $\boldsymbol{v}_{j+1}$ is $\boldsymbol{v}_1$. $\delta\theta_j$ is defined to be the angle $\widehat{\boldsymbol{v}_j\, 0\, \boldsymbol{v}_{j+1}}$, that is the angle in the closed radial segment SS_j obtained when one rotates the ray from 0 through $\boldsymbol{v}_j$ counterclockwise to obtain the ray from 0 through $\boldsymbol{v}_{j+1}$.

Suppose there is a j so that $\delta\theta_j > \pi$. Then all the points $\boldsymbol{v}_r$ lie in the complement of the interior of the radial segment SS_j. Therefore all the $\boldsymbol{v}_j$ lie on one side of the line through 0 perpendicular to the ray through 0 along the middle of the radial sector SS_j. Hence $\boldsymbol{v} = \boldsymbol{0}$ is not in the convex hull of the points $\mathcal{V}$, which is a contradiction.

It follows that all the angles $\delta\theta_j$ are less than or equal to π. Hence all the line segments $[\boldsymbol{v}_j, \boldsymbol{v}_{j+1}]$ lie in the corresponding radial segments SS_j. Further, the interior of the line segment $[\boldsymbol{v}_j, \boldsymbol{v}_{j+1}]$ lies within the interior of SS_j if $\theta_j \ne \theta_{j+1}$. Consequently the polygon $[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ is not self-intersecting. It is by construction star like about 0. □

In the simplest case our mean value coordinate preconditioning matrix Q is assembled as follows. Let $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\} \subset \mathcal{R}^2$ be a set of vertices not all lying on a single straight line. Assume that $\mathcal{T}$ is a triangulation of the convex hull $\mathcal{H}(\mathcal{X})$ with vertices in $\mathcal{X}$. The assumption of non-collinearity means $\mathcal{H}(\mathcal{X})$ has nonzero area. Select three of the extreme points of $\mathcal{H}(\mathcal{X})$ as special points. Numerical experiments (not included) show that it is beneficial to choose these three points so that the area of the corresponding triangle is large. Relabel the points of $\mathcal{X}$ so that the special points are $\boldsymbol{x}_{N-2}$, $\boldsymbol{x}_{N-1}$ and $\boldsymbol{x}_N$.

Algorithm forming a 2-complete set of differences for $\mathcal{X} \subset \mathcal{R}^2$

Proceed though the vertices $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-3}$

if $\boldsymbol{x}_j$ is an interior point of $\mathcal{H}(\mathcal{X})$
Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_t$ be the vertices of the one ring of $\boldsymbol{x}_j$. Let $\lambda_1(\boldsymbol{v}), \lambda_2(\boldsymbol{v}), \ldots, \lambda_t(\boldsymbol{v})$ be the corresponding mean value coordinates specifying $\boldsymbol{v} = \boldsymbol{x}_j$ as a convex combination of the vertices in its 1-ring. Then define Δ_j by

$$\Delta_j g = -g(\boldsymbol{v}) + \sum_{r=1}^{t} \lambda_r(\boldsymbol{v}) g(\boldsymbol{v}_r) =: \sum_{i=1}^{N} q_{ij} g(\boldsymbol{x}_i).$$

else when $\boldsymbol{x}_j$ *is on the boundary of* $\mathcal{H}(\mathcal{X})$

Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_t$ be the vertices in the one ring of $\boldsymbol{x}_j$. Introduce a false point $\boldsymbol{v}_{t+1}$ such that $\boldsymbol{v} = \boldsymbol{x}_j$ is the centroid of the points $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{t+1}$. The point $\boldsymbol{v}_{t+1}$ is outside $\mathcal{H}(\mathcal{X})$ as some of the points $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_t$ do not lie on a supporting hyperplane for $\mathcal{H}(\mathcal{X})$ through $\boldsymbol{x}_j$. Calculate the mean value coordinates $\lambda_1(\boldsymbol{v}), \ldots, \lambda_t(\boldsymbol{v}), \lambda_{t+1}(\boldsymbol{v})$ of $\boldsymbol{v} = \boldsymbol{x}_j$ with respect to $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_t, \boldsymbol{v}_{t+1}\}$. Also express the false point $\boldsymbol{v}_{t+1}$ as an affine combination of the special points $\boldsymbol{x}_{N-2}, \boldsymbol{x}_{N-1}$ and $\boldsymbol{x}_N$. Thus

$$\boldsymbol{v}_{t+1} = \mu_2 \boldsymbol{x}_{N-2} + \mu_1 \boldsymbol{x}_{N-1} + \mu_0 \boldsymbol{x}_N,$$

where $\mu_2 + \mu_1 + \mu_0 = 1$. Define Δ_j by

$$\Delta_j g = -g(\boldsymbol{v}) + \sum_{r=1}^{t} \lambda_r(\boldsymbol{v}) g(\boldsymbol{v}_r) + \lambda_{t+1} \sum_{i=0}^{2} \mu_i \, g(\boldsymbol{x}_{N-i}) =: \sum_{i=1}^{N} q_{ij} g(\boldsymbol{x}_i).$$

end if

Lemma 6.3. *The differences defined above form a* 2*-complete set of differences.*

Proof. We defer the proof that each difference annihilates linears and turn first to the question of completeness. Write

$$Q = \begin{bmatrix} E \\ F \end{bmatrix},$$

where E is $(N-3) \times (N-3)$.

Now consider E^T. We will show that E^T is invertible by a discrete maximum principle type argument. Suppose E^T is not invertible. Then there is a nontrivial solution $\boldsymbol{c}$ to $E^T \boldsymbol{c} = \boldsymbol{0}$.

Let j be such that $|c_j| = \|\boldsymbol{c}\|_\infty$. Let $\boldsymbol{r}_j$ be the jth row of E^T. If the corresponding vertex $\boldsymbol{x}_j$ is an internal vertex then the equation $\boldsymbol{r}_j \boldsymbol{c} = 0$ says that c_j is a convex combination of the components c_r of $\boldsymbol{c}$ at neighbouring vertices $\boldsymbol{x}_r$, which has modulus $\|\boldsymbol{c}\|_\infty$. Hence the values c_r of $\boldsymbol{c}$ at all the neighbouring vertices must equal c_j. Iterating this argument we find that there is at least one vertex $\boldsymbol{x}_r$ on the boundary of $\mathcal{H}(\mathcal{X})$ at which c_r achieves its maximum modulus. We are already in this case if the original vertex $\boldsymbol{x}_j$ is a boundary vertex.

Let $\boldsymbol{x}_j$ be such a boundary vertex. By construction $e_{jj} = -1$ and the other components in the jth column of E are nonnegative and sum to less than 1. (The entries in the jth column of F are not necessarily nonnegative.) Hence since $|c_j| = \|\boldsymbol{c}\|_\infty \neq 0$ it follows that $\boldsymbol{r}_j \boldsymbol{c} \neq 0$, which is a contradiction. Hence E^T must be invertible, contrary to our assumption, and Q has full rank.

Second, using the notations introduced in the first part if $\boldsymbol{v} = \boldsymbol{x}_j$ is an interior vertex

$$\sum_{i=1}^{N} q_{ij} = -1 + \sum_{r=1}^{t} \lambda_r(\boldsymbol{v}) = 0,$$

and

$$\sum_{i=1}^{N} q_{ij} \boldsymbol{x}_i = -\boldsymbol{v} + \sum_{r=1}^{t} \lambda_r(\boldsymbol{v}) \boldsymbol{v}_r = 0.$$

Hence the difference $\Delta_j g = \sum_{i=1}^{N} q_{ij} g(\boldsymbol{x}_i)$ annihilates linears. The proof for columns corresponding to vertices $\boldsymbol{x}_j$ on the boundary of $\mathcal{H}(\mathcal{X})$ is only very slightly more complicated. The stated result follows. □

We note also that the mean value coordinates are unchanged under uniform scalings. Therefore the differences constructed from them by the algorithm above satisfy the assumptions required for scale invariance results Theorem 5.5 and Corollary 5.6, with $\sigma(h) = 1$.

7. Numerical results: condition numbers

Condition numbers for the matrices occurring in unpreconditioned and preconditioned interpolation problems in a large sets of random trials are shown in Tables 1–3. As can be seen from the tables, the preconditioning procedure results in a dramatic improvement in the conditioning of multiquadric and biharmonic radial basis function interpolation problems in $\mathcal{R}^2$ and $\mathcal{R}^3$.

The entries in the tables are percentiles of the distribution of 2-norm condition numbers observed in random trials. For N chosen in turn as 100, 200, 400, 800, and for each basic function Φ, 20,000 random trials were conducted. In each trial N centres were chosen uniformly at random in $[0, 1]^2$ or $[0, 1]^3$ as appropriate, a Delaunay triangulation of the data

Table 1
Percentiles for the distribution of 2-norm condition numbers of interpolation matrices observed for N random interpolation nodes, for various choices of N. The basic function is the thin-plate (biharmonic) in $\mathcal{R}^2$, $\Phi(\boldsymbol{x}) = |\boldsymbol{x}|^2 \log |\boldsymbol{x}|$.

N	Min	1%	10%	Median	90%	99%	Max
100	2.39(4)	4.09(4)	6.71(4)	1.72(5)	8.11(5)	6.3(6)	1.41(9)
Precon.	6.79(0)	1.06(1)	1.69(1)	4.02(1)	1.22(2)	4.06(2)	6.47(3)
200	1.20(5)	2.33(5)	3.98(5)	1.03(6)	5.14(6)	4.10(7)	2.16(9)
Precon.	8.96(0)	1.73(1)	3.19(1)	8.36(1)	2.69(2)	8.80(2)	2.63(4)
400	6.69(5)	1.39(6)	2.45(6)	6.63(6)	3.40(7)	2.78(8)	1.56(10)
Precon.	1.43(1)	3.06(1)	6.17(1)	1.73(2)	5.80(2)	2.00(3)	2.09(4)
800	4.31(6)	9.20(6)	1.63(7)	4.48(7)	2.44(8)	1.98(9)	2.60(11)
Precon.	1.75(1)	5.78(1)	1.24(2)	3.68(2)	1.28(3)	4.49(3)	4.71(4)

Table 2
Percentiles for the distribution of 2-norm condition numbers observed for N random interpolation nodes, for various choices of N. The basic function is the ordinary multiquadric in $\mathcal{R}^2$, $\Phi(\boldsymbol{x}) = \sqrt{|\boldsymbol{x}|^2 + c^2}$, with $c = 1/\sqrt{N}$.

N	Min	1%	10%	Median	90%	99%	Max
100	2.17(5)	5.10(5)	1.15(6)	4.54(6)	3.58(7)	4.19(8)	1.16(11)
Precon.	1.42(1)	2.41(1)	3.49(1)	6.97(1)	2.46(2)	1.42(3)	8.02(5)
200	1.44(6)	3.43(6)	7.47(6)	2.87(7)	2.25(8)	2.53(9)	1.37(12)
Precon.	2.30(1)	3.49(1)	5.06(1)	1.03(2)	3.62(2)	2.03(3)	1.91(5)
400	8.69(6)	2.23(7)	4.77(7)	1.78(8)	1.38(9)	1.61(10)	1.04(13)
Precon.	3.44(1)	4.93(1)	7.30(1)	1.49(2)	5.35(2)	3.20(3)	1.31(6)
800	5.82(7)	1.41(8)	2.97(8)	1.10(9)	8.42(9)	8.81(10)	9.83(12)
Precon.	4.60(1)	7.02(1)	1.06(2)	2.16(2)	7.71(2)	4.60(3)	2.22(5)

Table 3
Percentiles for the distribution of 2-norm condition numbers of interpolation matrices observed for N random interpolation nodes, for various choices of N. The basic function is the biharmonic in $\mathcal{R}^3$, $\Phi(\boldsymbol{x}) = |\boldsymbol{x}|$.

N	Min	1%	10%	Median	90%	99%	Max
100	1.17(3)	1.39(3)	1.68(3)	2.34(3)	4.16(3)	8.88(3)	7.11(4)
Precon.	5.05(0)	6.78(0)	8.46(0)	1.15(1)	1.65(1)	2.29(1)	3.92(1)
200	3.40(3)	4.20(3)	5.10(3)	7.26(3)	1.31(4)	2.69(4)	2.02(5)
Precon.	7.72(0)	9.88(0)	1.24(1)	1.71(1)	2.46(1)	3.43(1)	1.16(2)
400	1.02(4)	1.28(4)	1.57(4)	2.26(4)	4.16(4)	9.10(4)	4.30(5)
Precon.	1.00(1)	1.44(1)	1.81(1)	2.52(1)	3.64(1)	4.99(1)	1.21(2)
800	3.09(4)	3.97(4)	4.87(4)	7.05(4)	1.28(5)	2.75(5)	1.83(6)
Precon.	1.53(1)	2.13(1)	2.66(1)	3.68(1)	5.29(1)	7.26(1)	1.13(2)

constructed, and 2-norm condition numbers of the unpreconditioned and preconditioned interpolation matrices calculated. The results were then sorted and the various percentiles (e.g. the median) recorded. An entry $x.yz(e)$ in the tables indicates the number $x.yx \times 10^e$. Also in each table the rows starting with *Precon* are the results for the equilibrated $(N-\ell) \times (N-\ell)$ interpolation matrix of the preconditioned problem corresponding to the unpreconditioned results of the previous row.

8. Numerical results: accuracy of evaluation—indirect evaluation of the Ψ functions

Solving using the preconditioned system is a good first step towards avoiding the problems of accurate calculation with RBFs. However, used alone, it is often not enough. Numerical experiments reported in Table 4 show that proceeding by fitting with the "good" Ψ_i basis, and then converting back to the "bad" $\Phi(\cdot - \boldsymbol{x}_i)$ basis before evaluation, can result in throwing away most of the expected gain from preconditioning. In this section, we show that evaluating the new basis elements indirectly, significantly reduces this problem.

Evaluating a "small" RBF,

$$s(\boldsymbol{x}) = q(\boldsymbol{x}) + \sum_{j=1}^{N} \lambda_j \Phi(\boldsymbol{x} - \boldsymbol{x}_j),$$

given in terms of the "bad" basis can inherently involve taking differences of large numbers. This can happen because the coefficients are large and of differing sign. It can also happen when the coefficients are moderately sized and of differing sign, but the values of $\Phi(\boldsymbol{x} - \boldsymbol{x}_j)$'s are relatively large and almost equal. In both cases loss of significance errors will probably occur upon evaluation. Since the sum of the coefficients of the $\Phi(\cdot - \boldsymbol{x}_j)$ functions contributing to a single Ψ_i is zero, even evaluating a single Ψ_i directly via the $\Phi(\boldsymbol{x} - \boldsymbol{x}_j)$'s will incur loss of significance errors, when $\boldsymbol{x}$ is far from the relevant $\boldsymbol{x}_j$'s. See [5] for further discussion of the impact of basis upon the stability of evaluation.

Table 4
Infinity norm condition numbers, norms of coefficient vectors, and relative residuals, for various thin-plate spline interpolation problems based on the spiral points.

N	Method	Condition number	Function	Norm of $\boldsymbol{\lambda}$ or $\boldsymbol{\mu}$	Relative residual
100	Usual	7.02(11)	f	6.92(9)	1.37(−6)
	Ψ's to Φ's A	1.90(1)		6.92(9)	9.91(−7)
	Ψ's to Φ's B			1.06(5)	5.05(−7)
	Ψ's only			1.06(5)	2.40(−11)
	Usual	7.02(11)	g	8.74(7)	1.64(−8)
	Ψ's to Φ's A	1.90(1)		8.74(7)	1.14(−8)
	Ψ's to Φ's B			2.05(3)	5.06(−9)
	Ψ's only			2.05(3)	8.30(−13)
200	Usual	8.98(13)	f	4.43(11)	1.43(−4)
	Ψ's to Φ's A	8.48(1)		4.43(11)	5.16(−5)
	Ψ's to Φ's B			8.50(5)	2.33(−5)
	Ψ's only			8.50(5)	1.98(−10)
	Usual		g	1.85(10)	5.63(−6)
	Ψ's to Φ's A			1.85(10)	2.24(−6)
	Ψ's to Φ's B			3.65(4)	1.07(−6)
	Ψ's only			3.65(4)	6.84(−12)

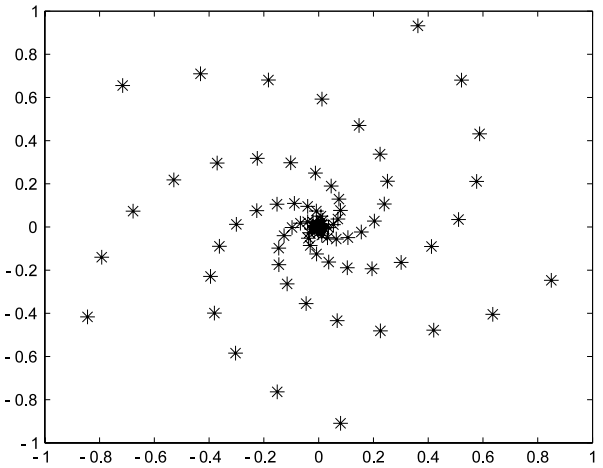

Fig. 4. Spiral points for $N = 100$.

One method to reduce the problem above is to stay with the basis of Ψ_i functions as much as possible, and to evaluate the Ψ_i's indirectly. In the experiments below the indirect, or more precise, evaluation method chosen is to calculate the values $\Psi_i(\boldsymbol{x})$ using far field expansions whenever the evaluation point $\boldsymbol{x}$ is sufficiently far from the $\boldsymbol{x}_j$'s involved in the definition of Ψ_i, and by direct evaluation of the relevant $\Phi(\boldsymbol{x} - \boldsymbol{x}_j)$'s otherwise. More precisely, far field approximations were used to evaluate $\Psi_i(\boldsymbol{x})$ whenever they could be guaranteed to give accuracy at least ϵ in evaluating $\mu_i\Psi_i(\boldsymbol{x})$. ϵ was chosen as 10^{-12}. In the numerical experiments reported below we did this, calculating even the preconditioned matrix with the Ψ_i functions.

In the experiments the nodes of interpolation were chosen as the spiral points

$$\boldsymbol{x}_i = \frac{(N+1-i)^3}{N^3}[\cos(1.2\,i), \sin(1.2\,i)], \quad 1 \le i \le N,$$

illustrated for $N = 100$ in Fig. 4. The geometry of these nodes results in the usual interpolation system (6) being very ill-conditioned. The first function to fit, f, is defined only at the interpolation nodes and is chosen to have a large high frequency component. It is

$$f(\boldsymbol{x}_i) = f_i = (-1)^i, \quad 1 \le i \le N.$$

The second function to fit is the smooth function

$$g(\boldsymbol{x}) = \exp(-(x^2 + y^2))\cos(x).$$

The results of fitting these functions with either 100 or 200 nodes are shown in Table 4. The table gives infinity norm condition numbers of the relevant interpolation matrices, and the infinity norm of the relevant parameter vector λ or μ. Evaluation accuracy was measured by the infinity norm relative residual in the RBF, s, at the interpolation nodes.

In the table rows labelled "Usual" show the results for fitting via system (3), and evaluating with the $\Phi(\cdot - \boldsymbol{x}_j)$ functions. The rows labelled "Ψ's to Φ's A" show the results of solving the preconditioned system and then converting s to the basis of Φ functions before evaluation. The rows labelled "Ψ's to Φ's B" show the results of solving the preconditioned system and then evaluating each individual Ψ_i function in terms of its constituent $\Phi(\cdot - \boldsymbol{x}_j)$'s. This groups terms in the sum differently than for the "Ψ's to Φ's A" rows. The rows labelled "Ψ's only" show the results of fitting with the preconditioned system and then evaluating the Ψ's indirectly.

It can be seen from the table that the evaluation accuracy was consistently better for the smooth signal compared with the high frequency signal. Further, as already mentioned above, fitting with the preconditioned system and then evaluating with the Φ functions, by either of the two methods described, typically improves the evaluation accuracy only slightly. Finally, the "Ψ's only" rows show that fitting with the preconditioned system and then evaluating the Ψ functions indirectly yields almost five more significant figures of accuracy.

We now turn to the question of showing analytically why the use of far field expansions to evaluate the Ψ_i's can give such an increase in accuracy. The short answer is that there are frame like bounds between the coefficients of far field expansions and the function they represent (restricted to the far field). Thus evaluating via these expansions automatically avoids loss of significance resulting from expressing a small quantity as the difference of two large quantities. It leads to greater accuracy provided, the forming of the coefficients of the far field expansion is itself not problematic.

8.1. Far field expansions for evaluation—polyharmonic case

The untruncated expansion for an $(m+1)$-harmonic RBF in $\mathcal{R}^2$ is, writing $z = x + \mathrm{i}\,y$,

$$g_p(z) = \Re\left\{\left(\sum_{j=0}^{m}\overline{z}^j\sum_{k=0}^{j}a_{jk}z^k\right)\log|z| + \sum_{j=0}^{m}\overline{z}^j\sum_{\sigma=\max(1-m,-j)}^{\infty}c_{j\sigma}z^{-\sigma}\right\};$$

see [16]. Handling terms that grow at infinity separately, and restricting to the biharmonic case for simplicity, the interesting part of the expansion can be rewritten in terms of polar coordinates as

$$g(r,\theta) + r^2h(r,\theta) = \sum_{k=0}^{\infty}\frac{d_{0,k}\cos(k\theta) + e_{0,k}\sin(k\theta)}{r^k} + \sum_{k=0}^{\infty}r^2\frac{d_{1,k+2}\cos((k+2)\theta) + e_{1,k+2}\sin((k+2)\theta)}{r^{k+2}}, \tag{14}$$

converging uniformly for $r \geq R$. Then using Parseval

$$\frac{1}{\pi}\int_{-\pi}^{\pi}g(r,\theta)^2\mathrm{d}\theta = \frac{1}{2}d_{0,0}^2 + \sum_{j=1}^{\infty}\frac{d_{0,j}^2 + e_{0,j}^2}{r^{2j}},$$

and

$$\frac{1}{\pi}\int_{-\pi}^{\pi}r^4h(r,\theta)^2\,\mathrm{d}\theta = \sum_{j=2}^{\infty}\frac{d_{1,j}^2 + e_{1,j}^2}{r^{2j-4}},$$

all quantities being decreasing in $r \geq R$. Thus the sizes of the sequences of coefficients are directly related to the norms of the functions $g(r,\theta)$ and $h(r,\theta)$. Also the terms multiplying the coefficients (except for the constant term) are bounded by negative powers of r. Thus there can be no catastrophic loss of significance in evaluating the far field expansion (14). It is an orthogonal expansion and thus inherently well conditioned.

References

[1] R. Sibson, G. Stone, Computation of thin-plate splines, SIAM J. Sci. Stat. Comput. 12 (1991) 1304–1313.
[2] W.R. Madych, Miscellaneous error bounds for multiquadric and other interpolators, Comput. Math. Appl. 24 (12) (1992) 121–138.
[3] N. Dyn, D. Levin, S. Rippa, Numerical procedures for surface fitting of scattered data by radial functions, SIAM J. Sci. Stat. Comput. 7 (1986) 639–659.
[4] F.J. Narcowich, J.D. Ward, Norm estimates for the inverses of a general class of scattered-data radial-function interpolation matrices, J. Approx. Theory 69 (1992) 84–109.
[5] S. Muller, R. Schaback, A Newton basis for kernel spaces, J. Approx. Theory 161 (2) (2008) 645–655.
[6] C.T. Mouat, R.K. Beatson, On the boundary over distance preconditioner for radial basis function interpolation, in: J. Levesley, I. Anderson, J.C. Mason (Eds.), Algorithms for Approximation IV, University of Huddersfield Press, 2002, pp. 252–259.
[7] C.T. Mouat, R.K. Beatson, Some properties of the boundary over distance preconditioner for radial basis function interpolation, Report UCDMS 2001/6, Department of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand.
[8] R.K. Beatson, W.A. Light, S. Billings, Fast solution of the radial basis function interpolation equations: domain decomposition methods, SIAM Journal on Scientific Computing 22 (2000) 1717–1740.
[9] M. Bozzini, L. Lenarduzzi, R. Schaback, Kernel B-splines and interpolation, Numer. Algorithms 41 (2006) 1–16.
[10] R.K. Beatson, N. Dyn, Multiquadric B-splines, J. Approx. Theory 87 (1996) 1–24.
[11] O. Davydov, A. Sestini, R. Morandi, Local RBF approximation for scattered data fitting with bivariate splines, in: Trends and Applications in Constructive Approximation, in: M.G. de Bruin, D.H. Mache, J. Szabados (Eds.), ISNM, vol. 151, Birkhäuser, 2005, pp. 91–102.
[12] C.A. Micchelli, Interpolation of scattered data: distance matrices and conditionally positive definite functions, Constr. Approx. 2 (1986) 11–22.
[13] M. Floater, Mean value coordinates, Comput. Aided Geom. Design 20 (2003) 19–27.
[14] K. Hormann, M. Floater, Mean value coordinates for arbitrary polygons, ACM Trans. Graph. 25 (2006) 1424–1441.
[15] M.S. Floater, G. Kos, M. Reimers, Mean value coordinates in 3D, Comput. Aided Geom. Design 22 (2005) 623–631.
[16] R.K. Beatson, W.A. Light, Fast evaluation of radial basis functions: methods for 2-dimensional polyharmonic splines, IMA J. Numer. Anal. 17 (1997) 343–372.