Creating Surfaces from Scattered Data Using Radial Basis Functions

R. Schaback

Abstract. This paper gives an introduction to certain techniques for the construction of geometric objects from scattered data. Special emphasis is put on interpolation methods using compactly supported radial basis functions.

§1. Introduction

We assume a sample of multivariate scattered data to be given as a set $X = \{x_1, \dots, x_N\}$ of N pairwise distinct points $x_1, \dots, x_N$ in $\mathbb{R}^d$, called *centers*, together with N points $y_1, \dots, y_N$ in $\mathbb{R}^D$. An *interpolating curve, surface*, or *solid* to these data will be the range of a smooth function $s : \mathbb{R}^d \supset \Omega \to \mathbb{R}^D$ with

$$s(x_k) = y_k,\ 1 \le k \le N. \tag{1}$$

Likewise, an *approximating* curve, surface, or solid will make the differences $s(x_j) - y_j$ small, for instance in the discrete L_2 sense, i.e.

$$\sum_{k=1}^{N} \|s(x_k) - y_k\|_2^2$$

should be small. Curves, surfaces, and solids will only differ by their appropriate value of $d = 1, 2$, or 3. We use the term (geometric) *objects* to stand for curves, surfaces, or solids. Note that we assume objects to be defined via *explicit* representations. For *implicit* representations we refer the reader to section 4.

The main goal of this contribution is to describe a flexible class of objects that allows construction from scattered data in a very general way. Special emphasis is put on practical aspects, while theoretical background information will be contained in a forthcoming survey, as far as it is not contained in earlier surveys on similar topics [5], [6], [7], [8], [14], [16], [17], [21], [24]. To keep the

Mathematical Methods in CAGD III
M. Dæhlen, T. Lyche, and L. L. Schumaker (eds.), pp. 1–21.

ISBN xxx.

paper digestible for readers working in CAGD applications, we refrained from using advanced techniques like Fourier or Hankel transforms, distributions, or Bessel functions. Furthermore, we refer the reader to the cited surveys for obtaining additional references or historical remarks.

The theory of optimal recovery provides good theoretical reasons to assume s to be of the form

$$s(x) := \sum_{j=1}^{N} b_j \Phi(x - x_j) \tag{2}$$

with *control points* $b_1, \ldots, b_N \in \mathbb{R}^D$ and a function $\Phi : \mathbb{R}^d \to \mathbb{R}$ that is *positive definite* on $\mathbb{R}^d$ in the following sense: For all sets $X = \{x_1, \ldots, x_N\}$ of finitely many distinct points $x_1, \ldots, x_N$ in $\mathbb{R}^d$ the matrix

$$A = (\Phi(x_k - x_j))_{1 \le j,k \le N} \tag{3}$$

is *positive definite*. We write $\Phi \in \mathbf{PD}_d$ as shorthand for this property, and it is clear that it guarantees solvability of the system

$$s(x_k) = \sum_{j=1}^{N} b_j \Phi(x_k - x_j) = y_k, \; 1 \le k \le N, \tag{4}$$

defining an interpolating object. Of course it would suffice to require that (3) is nonsingular for each $x = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$, but so far there is no theory to handle this generalization, while there is a long history of positive (semi–) definite functions (see [33], but note the slightly different definition there). Another good reason to consider objects defined by (2) with $s \in \mathbf{PD}_d$ will be completely ignored in this contribution: all such objects minimize certain "energy functionals" over a large set of objects admissible for interpolation. Furthermore, there is a probabilistic background to this approach which has been exploited in detail by authors working on geophysical applications, using the term *Kriging* [2], [18], [35].

For geometric applications it would be very convenient if the solution of the interpolation conditions (1) were independent of Euclidean transformations of the data. This can be achieved if the function $\Phi : \mathbb{R}^d \to \mathbb{R}$ is *radial* in the sense

$$\Phi(x) = \phi(\|x\|_2), \; x \in \mathbb{R}^d \tag{5}$$

with a univariate function $\phi : \mathbb{R}_{\ge 0} \to \mathbb{R}$ and the Euclidean norm $\|.\|_2$ on $\mathbb{R}^d$. Furthermore, (5) often simplifies the evaluation of Φ due to the reduction to a univariate function ϕ. For convenience, we restrict ourselves from now on to *radial basis functions* Φ defined by some ϕ via (5).

There is an important extension to the class of objects defined by (2). First, one adds polynomials from the space $\mathbb{P}_m^d$ of d–variate polynomials of order at most m to generalize (2) to

$$s(x) = \sum_{j=1}^{N} b_j \Phi(x - x_j) + \sum_{\ell=1}^{M} c_\ell p_\ell(x) \tag{6}$$

with $M = \dim \mathbb{P}^d_m$, additional control points $c_1, \ldots, c_M \in \mathbb{R}^D$, and a basis $p_1, \ldots, p_M$ of $\mathbb{P}^d_m$. Then Φ usually is required to be *conditionally* positive *definite of order* m *on* $\mathbb{R}^d$, i.e. the matrix A of (3) only needs to be positive definite on the space of vectors $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N$ satisfying

$$\sum_{j=1}^{N} \alpha_j p_\ell(x_j) = 0, \; 1 \le \ell \le M. \tag{7}$$

The solution of the interpolation problem (1) for the augmented function (6) then requires the additional condition

$$\sum_{j=1}^{N} b_j p_\ell(x_j) = 0, \; 1 \le \ell \le M$$

corresponding to (7). This condition looks strange at first sight, but it occurs naturally if one considers minimization of certain error functionals.

Up to now, *conditionally* positive definite functions provided the most important cases of radial basis functions (see Table 1), and they have proven to be highly useful in multivariate scattered data analysis [11]. For $m > 0$ they usually behave like a positive power of $\|x\|_2$ for large $\|x\|_2$. Thus the matrix A of (3) will not be sparse, and there will be no off–diagonal decay unless A is suitably preconditioned [15]. Techniques for preconditioning are available but not easy to implement for scattered multivariate data [9]. Furthermore, the sum in (2) will be costly to evaluate for large N unless some rather complicated, but very efficient acceleration techniques are used [22], [23].

Multiquadrics $\phi(r) = (c^2 + r^2)^{\beta/2}$ for $\beta \in \mathbb{R}_{>-d} \setminus 2\mathbb{Z}$ and $2m > \beta$
Thin-plate splines $\phi(r) = r^\beta$ for $\beta \in \mathbb{R}_{>0} \setminus 2\mathbb{Z}$ and $2m > \beta$
Thin-plate splines $\phi(r) = (-1)^{\beta/2+1} r^\beta \log r$ for $\beta \in 2\mathbb{N}$, $2m > \beta$
Sobolew splines $\phi(r) = \frac{2\pi^k}{\Gamma(k)} K_{k-d/2}(2\pi r) \cdot r^{k-d/2}$ for $k > d/2$ and $m \ge 0$
using the Macdonalds (or spherical Bessel) function K_ν,
Gaussians $\phi(r) = e^{-cr^2}$ for $c > 0$ and $m \ge 0$.

Table 1. Conditionally positive definite functions of order m.

For these reasons we shall concentrate here on *compactly supported* positive definite radial functions $\Phi(\cdot) = \phi(\|\cdot\|_2)$. They will lead to a sparse matrix in (3) and to an easy evaluation of (2), if implemented properly (see section 5). However, such functions were discovered only very recently, and we give a fairly complete summary of their construction in section 2. Examples can be found in section 3 and 4, while the computational complexity is treated in section 5.

§2. Compactly supported positive definite radial functions

A straightforward technique for generation of compactly supported positive definite radial functions simply uses a d–variate convolution

$$\phi(\|x\|_2) = \int_{y\in\mathbb{R}^d} \psi(\|y\|_2)\cdot\psi(\|x-y\|_2)dy \tag{8}$$

of a continuous compactly supported nonzero function $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with itself. By Fourier transform arguments [32] the result is a radial function, and it is compactly supported by construction. Positive definiteness follows from

$$\sum_{j,k=1}^{N} \alpha_j\alpha_k\phi(\|x_j - x_k\|_2) = \int_{\mathbb{R}^d}\left(\sum_{j=1}^{N}\alpha_j\psi(\|x_j - y\|_2)\right)^2 dy$$

and the linear independence of finitely many distinct translates of a nonzero compactly supported continuous function.

By convolution (8) one can generate abundantly many nonnegative compactly supported, continuous, and positive definite radial functions. The main problem, however, is that for $d > 1$ there were no explicit constructions until recently.

To our knowledge the first example was provided in 1993 by H. Wendland [34], using the characteristic function $\psi = \chi_{[0,1]}$ and convolution in $\mathbb{R}^2$. We include the recipe here, because it is quite instructive.

The univariate hat function $\beta_2(t) = (1-|t|)_+$ is a second-order symmetric univariate B-spline with knots $-1, 0, 1$. It is the convolution $\beta_2 = \chi_{[-1/2,1/2]} * \chi_{[-1/2,1/2]}$ of two instances of the characteristic function $\chi_{[-1/2,1/2]}$ on $[-1/2, 1/2]$. If this approach is generalized to the multivariate setting, the result is a radial function only if convolutions of characteristic functions of *Euclidean* balls are taken. Using the unit balls of the Chebyshev or L_∞ norm yields the tensor product of instances of β_2, i.e.: a 2^d-sided pyramid that could be called *Chebyshev's hat* and that has knot lines joining the center with the corners of the domain.

We now proceed to calculate *Euclid's hat* by convolution $\beta_{2,d} := \chi_{B_d} * \chi_{B_d}$ where B_d is the Euclidean unit ball in $\mathbb{R}^d$. This involves to calculate the volume of the cap of the unit ball in $\mathbb{R}^d$ which is cut off by a plane at distance $t \leq 1$ to the origin. Thus, for $t \in [0,1]$, we have

$$\begin{aligned}\beta_{2,d}(2t) &= 2\int_t^1 vol(\sqrt{1-s^2}B_{d-1})ds\\ &= 2\omega_{d-1}\int_t^1 (1-s^2)^{(d-1)/2}ds,\end{aligned}$$

where ω_{d-1} is the volume of the unit ball B_{d-1} in $\mathbb{R}^{d-1}$. We substitute $t = \cos\psi,\ s = \cos\varphi$ and get

$$\beta_{2,d}(2\cos\psi) = 2\omega_{d-1}\int_0^\psi \sin^d\varphi d\varphi,\ 0 \leq \psi \leq \pi/2.$$

An easy integration by parts yields the recursion

$$dF_d(\psi) = (d-1)F_{d-2}(\psi) - \cos\psi \sin^{d-1}\psi,\ d \geq 2$$

for

$$F_d(\psi) = \int_0^\psi \sin^d \varphi d\varphi.$$

Clearly,

$$\begin{aligned} F_0(\psi) &= \psi, \\ F_1(\psi) &= \frac{1}{2}(\psi - \sin\psi\cos\psi), \end{aligned}$$

and all F_{2k+1} are odd polynomials in $\cos\psi$ due to the recursion. This implies that in spaces of odd dimension $d = 2k+1$ the Euclidean hat function is an odd polynomial

$$\begin{aligned} \beta_{2,2k+1}(2t) &= 2\omega_{2k}F_{2k}(\arccos t) \\ &= 2\omega_{2k}\frac{2k}{2k+1}F_{2k-1}(\arccos t) - \frac{1}{2k+1}t(1-t^2)^k \\ &= \frac{1}{2k+1}(2\pi\beta_{2,2k-1}(2t) - t(1-t^2)^k) \end{aligned}$$

for $t \in [0,1]$ with $\beta_{2,1}(2t) = 2(1-t)$. In spaces with even dimension $d = 2k$ we similarly get

$$\beta_{2,2k}(2t) = \frac{1}{2k}(2\pi\beta_{2,2k-2}(2t) - t\sqrt{1-t^2}(1-t^2)^{k-1})$$

with

$$\beta_{2,2}(2t) = 2(\arccos t - t\sqrt{1-t^2}).$$

This will always generate a function of t that can be written as a linear combination of $\arccos t/2$ and an odd polynomial in t multiplied by $\sqrt{1-t^2/4}$. In these formulae we made use of $\omega_n/\omega_{n-2} = 2\pi/n$. Figure 1 shows the graphs of $\beta_{2,d}(r)$ for dimensions $d = 1, 2, 3, 5$, and 100,

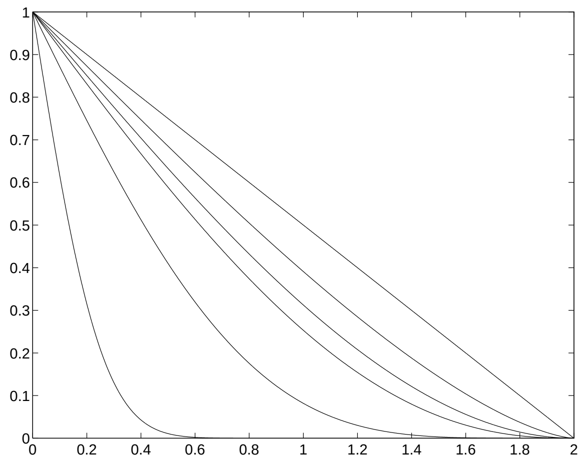


Fig. 1. Profile of Euclid's hats for dimensions $d = 1, 2, 3, 4, 10, 100$ from right to left.

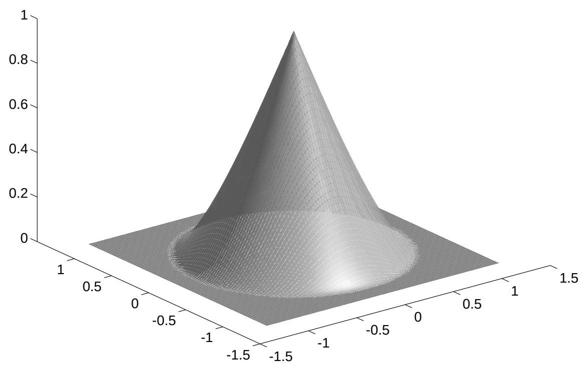


Fig. 2. Wendland's function $\beta_{2,2}$.

while Figure 2 shows a three-dimensional plot of $\beta_{2,2}$ after rescaling.

Note that the functions of Wendland are only continuous, but very easy to calculate for large space dimensions. To generate differentiable functions, we provide a second example. Let β_{2m} be the univariate B-spline of order $2m$ with uniform knot set $[-m, m] \cap \mathbb{Z}$. We do not need to specify the normalization here. It would be convenient to use $\beta_{2m}(\|x\|_2)$ on $\mathbb{R}^d$, but due to results of Wu [38], the plain B–Spline β_{2m} is not positive definite on any $\mathbb{R}^d$ for $d > 1$ when used as a radial d–variate function. This is why we now *radialize* the d-fold tensor product of β_{2m} in $\mathbb{R}^d$ to get a radial function $\rho_{2m,d}$ which is positive definite on $\mathbb{R}^d$, because its d-variate Fourier transform is positive, being the radialization of the d-variate tensor product of sinc^{2m}.

This approach yields the function

$$\rho_{2m,d}(r) := \frac{1}{\sigma_d} \int_{S_d} \prod_{j=1}^{d} \beta_{2m}(r y_j) dy, \; r \geq 0,$$

where σ_d is the surface area of the unit sphere $S_d \subset \mathbb{R}^d$. For $d = 2$ we parametrize S_d in the usual way and get, by use of symmetry,

$$\rho_{2m,2}(r) = \frac{8}{2\pi} \int_0^{\pi/4} \beta_{2m}(r \sin\varphi) \beta_{2m}(r \cos\varphi) d\varphi.$$

For positive arguments we represent β_{2m} as a piecewise polynomial via

$$\beta_{2m}(r) = \sum_{j=1}^{m} \chi_{[j-1,j]}(r) \sum_{k=0}^{2m-1} c_{jk} r^k$$

and evaluate $\rho_{2m,2}$ explicitly by application of the formulae

$$\rho_{2m,2}(r) = \frac{4}{\pi} \sum_{j,\ell=1}^{m} \sum_{k,n=0}^{2m-1} c_{jk} c_{\ell n} r^{k+n} F(r; j, \ell, k, n), \tag{9}$$

$$
\begin{aligned}
F(r;j,\ell,k,n) &:= \int_0^{\pi/4} \chi_{[j-1,j]}(r\cos\varphi)\chi_{[\ell-1,\ell]}(r\sin\varphi)\cos^k\varphi\sin^n\varphi d\varphi \\
&= \int_{\varphi_1(r;j,\ell)}^{\varphi_2(r;j,\ell)} \cos^k\varphi\sin^n\varphi d\varphi,
\end{aligned}
$$

where the limits of integration can be explicitly calculated and programmed. Since we want a precise evaluation without truncation errors, we make use of the fact that integrals of the form

$$\int_{\varphi_1}^{\varphi_2} \cos^k\varphi\sin^n\varphi d\varphi =: H_{k,n}(\varphi_1,\varphi_2)$$

are explicitly known and numerically available without integration. We use $\sin^2(x) = 1-\cos^2(x)$ to get the recursion

$$H_{k,n} = H_{k,n-2} - H_{k+2,n-2}$$

that reduces everything to $H_{k,1}$ and $H_{k,0}$, and then

$$
\begin{aligned}
H_{k,1}(\varphi_1,\varphi_2) &= \frac{1}{k+1}(\cos^{k+1}\varphi_1 - \cos^{k+1}\varphi_2) \\
kH_{k,0}(\varphi_1,\varphi_2) &= \sin\varphi_2\cos^{k-1}\varphi_2 - \sin\varphi_1\cos^{k-1}\varphi_1 + (k-1)H_{k-2,0}(\varphi_1,\varphi_2),
\end{aligned}
$$

while $H_{1,0}$ and $H_{0,0}$ are elementary. Though (9) contains $\mathcal{O}(m^4)$ terms, there can be only $\mathcal{O}(m)$ pairs (j,ℓ) that yield nonzero contributions for a specific value of r, because the arc $(r\cos\varphi, r\sin\varphi)$ meets only $\mathcal{O}(m)$ cubes $[j-1,j]\times[\ell-1,\ell]$. For each such pair (j,ℓ) we calculate $\varphi_1(r;j,\ell)$ and $\varphi_2(r;j,\ell)$ such that the interval $[\varphi_1,\varphi_2]$ precisely describes the set of values for φ that satisfy

$$\begin{pmatrix} j-1 & \le & r\cos\varphi & \le & j \\ \ell-1 & \le & r\sin\varphi & \le & \ell \\ 0 & \le & \varphi & \le & \pi/4. \end{pmatrix}$$

This can be done in a straightforward way, using arcsin and arccos functions appropriately. Once φ_1 and φ_2 are found, we calculate the whole set of values $F(r;j,\ell,k,n)$ for $0\le k,n\le 2m-1$ using the recursions.

This still is not a peculiarly efficient method to get $\rho_{2m,2}(r)$ for a given value of r. For large-scale applications one would prepare a sufficiently good piecewise polynomial or rational approximation to $\rho_{2m,2}$ beforehand, using the above method to generate precise values. We include Figure 3 for illustration.

Another example due to Iske [12] directly convolves the functions $(1-|t|)_+^\ell$ and again constructs continuous compactly supported radial basis functions on $\mathbb{R}^d$ for any value of d. However, it does not produce smooth functions, and thus we drop the details.

A rather neat construction due to Wu [37] starts with

$$f_\ell(x) = (1-x^2)_+^\ell,\ x\in\mathbb{R},\ \ell\ge 0,$$

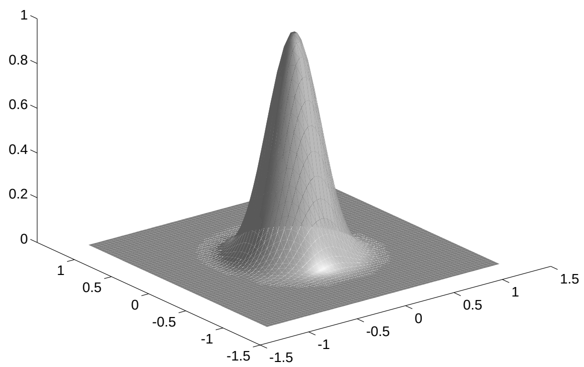


Fig. 3. Radialized cubic tensor product B-spline $\beta_4(r)$.

then takes univariate convolutions

$$\phi_{\ell,0}(x) = (f_\ell * f_\ell)(x)$$

and univariate derivatives

$$\phi_{\ell,k} = D^k \phi_{\ell,0}, \ \ D = -\frac{1}{x}\,\frac{d}{dx} \tag{10}$$

for $0 \leq k \leq \ell$. Curiously enough, this approach generates compactly supported functions

$$\phi_{\ell,k} \in \mathbf{PD}_{2k+1} \cap C^{2\ell-2k}(\mathbb{R}_{\geq 0})$$

which are nice piecewise polynomials of degree at most $4\ell - 2k + 1$ with a stable evaluation, having breakpoints only at zero and at the boundary of the support.

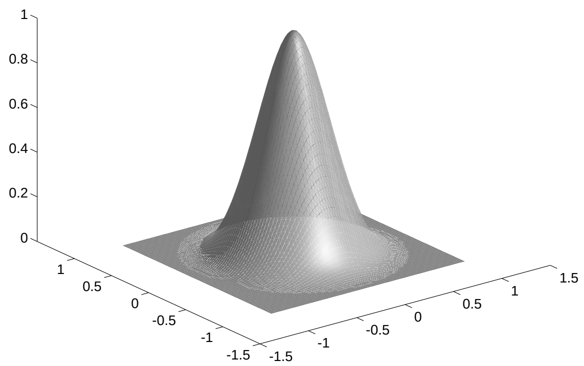


Fig. 4. Wu's function $\phi_{2,1}$ in $\mathbb{R}^2$.

See Table 2 for explicit formulae of $\phi_{\ell,k}(r)$ for the argument $r = \|.\|_2$ and with support normalized to $[0,1]$. Figure 4 provides a visualization of Wu's radial positive definite B–spline $\phi_{2,1}$ with C^2 smoothness. The C^0 function $\phi_{1,1}$ looks very much like Wendland's function in Figure 2, while the other functions from Wu's class look similar to $\phi_{2,1}$ when plotted in $\mathbb{R}^3$. Details are in [37], and the general theory will be contained in [31]. Note that $\doteq$ means equality up to a normalization factor in Table 2.

$$\phi_{0,0} = (1-r)_+ \in C^0 \cap \mathbf{PD}_1$$
$$\phi_{1,0} \doteq (1-r)_+^3(1+3r+r^2) \in C^2 \cap \mathbf{PD}_1$$
$$\phi_{1,1} = D\phi_{1,0} \doteq (1-r)_+^2(2+r) \in C^0 \cap \mathbf{PD}_3$$
$$\phi_{2,0} \doteq (1-r)_+^5(1+5r+9r^2+5r^3+r^4) \in C^4 \cap \mathbf{PD}_1$$
$$\phi_{2,1} = D\phi_{2,0} \doteq (1-r)_+^4(4+16r+12r^2+3r^3) \in C^2 \cap \mathbf{PD}_3$$
$$\phi_{2,2} = D^2\phi_{2,0} \doteq (1-r)_+^3(8+9r+3r^2) \in C^0 \cap \mathbf{PD}_5$$
$$\phi_{3,0} \doteq (1-r)_+^7(5+35r+101r^2+147r^3+101r^4+35r^5+5r^6) \in C^6 \cap \mathbf{PD}_1$$
$$\phi_{3,1} = D\phi_{3,0} \doteq (1-r)_+^6(6+36r+82r^2+72r^3+30r^4+5r^5) \in C^4 \cap \mathbf{PD}_3$$
$$\phi_{3,2} = D^2\phi_{3,0} \doteq (1-r)_+^5(8+40r+48r^2+25r^3+5r^4) \in C^2 \cap \mathbf{PD}_5$$
$$\phi_{3,3} = D^3\phi_{3,0} \doteq (1-r)_+^4(16+29r+20r^2+5r^3) \in C^0 \cap \mathbf{PD}_7$$

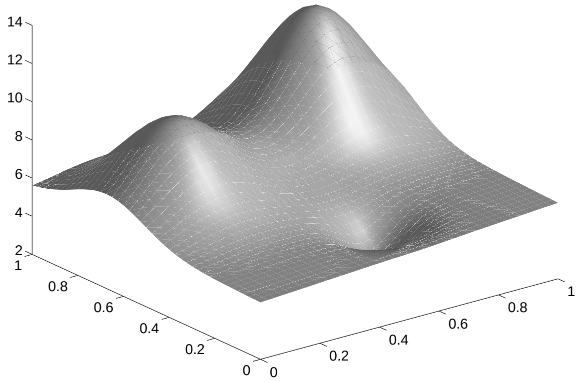
Table 2. Wu's compactly supported positive definite functions

We close this section by pointing at a remarkable property of Wu's functions $\phi_{\ell,k}$. They are in $C^{2\ell-2k}$ around zero, but in $C^{2\ell-k}$ at the boundary of their support. For $k > 0$ the singularities at zero will outweigh those at the boundary of the support. When superimposed for construction of interpolants, there will be $C^{2\ell-2k}$ singularities at the data locations, while there are $C^{2\ell-k}$ singularities on spheres around the data locations. As the next section will show, this property has a very positive effect on the visual appearance of the solutions.

§3. Some Preliminary Examples

Before turning to more specialized applications, we first want to show some examples of surface generation from scattered data using Wu's functions. To be compatible with earlier numerical experiments, we concentrate on the well-known function (Figure 5) introduced by Franke, scaled here to the domain $[0,1]^2$. We take 5×5 data points on a regular grid in $[0,1]^2$ and start with thin–plate splines (Figure 6) for comparison.

It is essential to note that the following examples with Wu's functions have a support radius of 0.499. This means that there are only 9 interpolation points at most in each support, and even less points in supports near the boundary. Enlarging the support always improves the quality of the results, but this is not intended here, because we want to illustrate what happens for reasonably small supports. The C^0 peaks of the interpolant with $\phi_{1,1}$ are clearly visible in Figure 7, but the reproduction in the flat areas is comparably



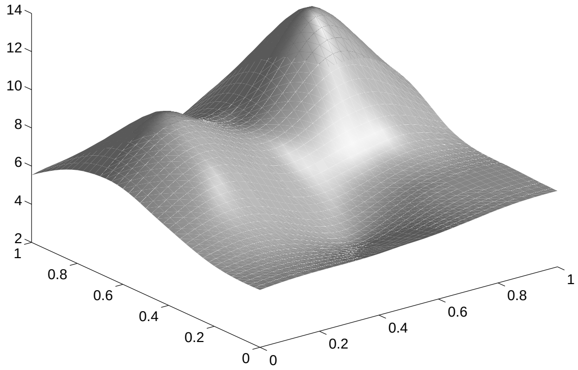
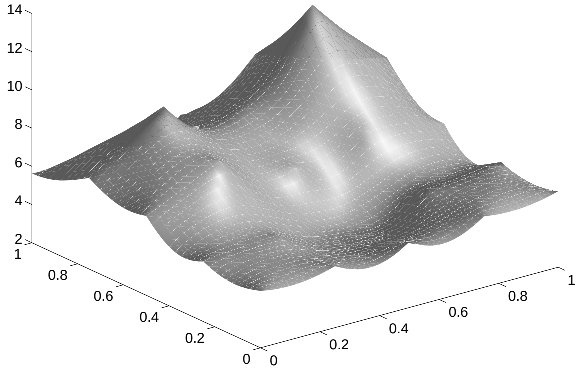


Fig. 5 und 6. Franke's function and Thin-plate spline interpolant on 5×5 points.



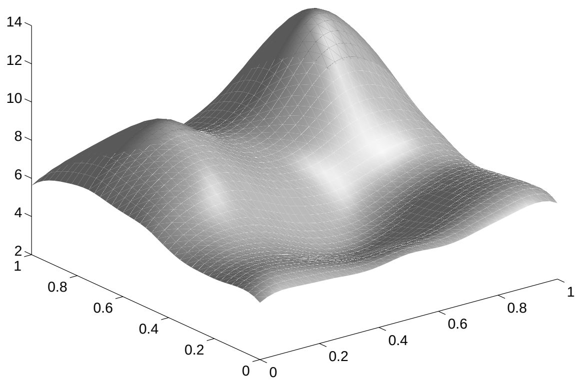
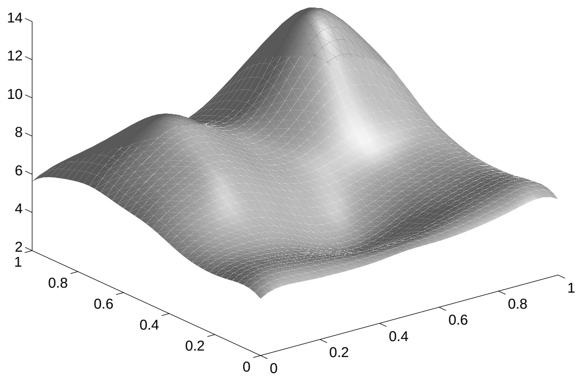


Fig. 7 and 8. Wu's C^0 interpolant on 5×5 points and Wu's C^2 interpolant on 5×5 points.



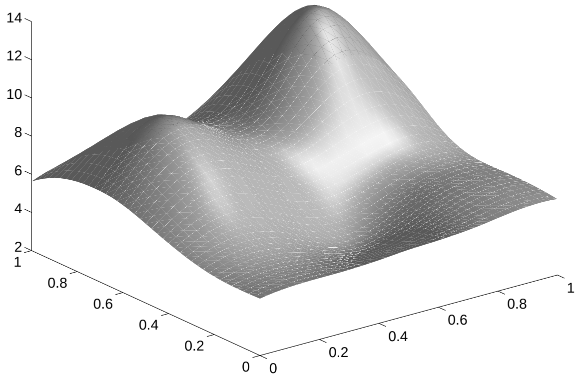


Fig. 9 and 10. Wu's C^4 interpolant on 5×5 points, support 0.499 and 1.5.

good, and for larger supports the flatness improves. Figures 8 and 9 show C^2 and C^4 interpolants, but the differentiability is C^3 and C^5 except at the data points. These are not visible. Note that for support radius = 0.499 there is a certain degradation near the boundary due to lack of points in the supports. Figure 10 shows how a larger support irons the boundary wiggles.

We do not directly intend to beat thin–plate splines in accuracy of reproduction at low numbers of data points. Our main goal is to show that comparable accuracy can be obtained by compactly supported functions whose support covers only a minor part of the data. If this works, one can treat extremely large problems very efficiently, making full use of sparsity of the system and locality of the representation. Thus the figures should be interpreted as local portions of much larger problems. For those the compactly supported radial basis functions will pay off with respect to computational efficiency, maintaining a comparable level of accuracy of reproduction. In particular, the pictures show local reproductions with supports containing at most 9 points, and the corresponding matrices will have maximally 9 entries in each row or column, while the reproduction quality is comparable to that of nonlocal radial basis functions like the thin-plate spline.

§4. Additional Features

In this section we give some hints to possible extensions and applications. These are mainly *in statu nascendi*, but the reader is invited to experiment with them and to accumulate further practical experience with compactly supported radial basis functions.

A widespread application of multivariate interpolation is *identification, comparison, morphing* or *warping* of images or other geometric objects. The basic idea is that an object $\mathcal{O}_1$ is to be linked to an object $\mathcal{O}_2$ via a transformation $F : \mathcal{O}_1 \to \mathcal{O}_2$ such that for certain corresponding points $x_j \in \mathcal{O}_1$ and $y_j \in \mathcal{O}_2$ one has $F(x_j) = y_j$. This multivariate interpolation problem can rather easily be implemented and solved using radial basis functions, and the locality provided by a compactly supported function should pay off in computational efficiency. See section 5 for efficiency considerations.

Another application arises from the first-stage construction of geometric objects from scattered measurements. An explicit representation can be computed directly via multivariate interpolation, if parameter values are assigned in some way or other. The calculation of *implicit* representations from scattered data will be treated at the end of this section. In both cases, the CAGD user will convert the objects in a second computation stage to some other representation that the underlying system readily accepts.

Unfortunately, interpolants of the form (2) do not satisfy the convex hull property with respect to their control points. This can be remedied by going over to *rational* representations of the form

$$s(x) = \frac{\sum_{j=1}^{N} \omega_j b_j \phi(\|x - x_j\|_2)}{\sum_{j=1}^{N} \omega_j \phi(\|x - x_j\|_2)} \tag{11}$$

with positive weights $\omega_1, \ldots, \omega_N \in \mathbb{R}_{>0}$. One can use (11) for construction of free–form objects, and (11) itself can be called a **N**on–**U**niform **R**adial **R**ational **B**–**S**pline.

To handle L_2–approximation problems one must be able to compute inner products

$$(\phi(\|\cdot - x_j\|_2),\ \phi(\|\cdot - x_k\|_2))_{L^2(\mathbb{R}^d)} \tag{12}$$

efficiently. Due to compact supports, the Gramian with entries (12) will be sparse again, but it is necessary to circumvent the complicated d–variate integral. Fortunately, the inner product (12), when written as a value

$$(\phi *_d \phi)(\|x_j - x_k\|_2)$$

of the d–variate convolution of $\phi(\|.\|_2)$ with itself, can be easily evaluated if $\phi *_d \phi$ is explicitly available as a univariate function. Due to a theorem in [31] it is possible to boil a d–variate convolution recursively down to a univariate convolution. More precisely, if D is the differential operator occuring in (10) with inverse

$$(IF)(r) = \int_r^\infty sf(s)ds,$$

then

$$f *_{d+2} g = D(If *_d Ig)$$

holds whenever both sides of this identity make sense for certain univariate functions $f, g : \mathbb{R}_{\geq 0} \to \mathbb{R}$. A similar but more complicated formula holds for recursion from $\mathbb{R}^{d+1}$ to $\mathbb{R}^d$, but we skip over details here and refer the reader to [31]. As a simple example we provide the trivariate convolution

$$\psi_{2,1} = \phi_{2,1} *_3 \phi_{2,1} = D(\phi_{2,0} *_1 \phi_{2,0})$$

of Wu's function $\phi_{2,1}$, which is a two-piece univariate rational spline

$$\begin{aligned}
s(x) &= (85 - 720x^2 + 3528x^4 - 3780x^5 - 960x^6 + \\
&\quad + 2160x^7 + 192x^9)/85, \quad x \in [0, 1/2] \\
s(x) &= 4(x-1)^6(16x^4 + 96x^3 + 156x^2 + 56x - 9)/(85x), \quad x \in [1/2, 1].
\end{aligned}$$

Multivariate interpolation can also serve to generate *implicitly* defined geometric objects that fit through a set of scattered data $y_1, \ldots, y_N \in \mathbb{R}^D$. If one takes a radial function ϕ of compact support and interpolates 1 at $y_1, \ldots, y_N$ by solving the system

$$\sum_{j=1}^N b_j \phi(\|y_k - y_j\|_2) = 1,\ 1 \leq k \leq N \tag{13}$$

for $b_j \in \mathbb{R}^1$, one can form the object

$$\mathcal{O} = \left\{ y \in \mathbb{R}^D : \sum_{j=1}^{N} b_j \phi(\|y - y_j\|_2) = 1 \right\}$$

which definitely contains the points $y_1, \ldots, y_N$ and is compact. Note that some clipping will be required, because the object will be a closed planar curve ($D = 2$) or a closed $3D$ surface ($D = 3$), possibly with several connected components. The interior of $\mathcal{O}$ is

$$\left\{ y \in \mathbb{R}^d : \sum_{j=1}^{N} b_j \phi(\|y - y_j\|_2) > 1 \right\}$$

and one can fix certain additional interpolation points y_k as being interior or exterior points by incorporating them into the interpolation using a value > 1 or < 1 instead of 1 in (13).

We include some preliminary tests for low numbers of data points in the plane. The four points (1,0), (0,1), (-1,0), (0,-1) are picked up by a closed curve which is the contour line of

$$s(y) = \sum_{j=1}^{N} b_j \phi(\|y - y_j\|_2)$$

at level 1 (see Figures 11, 12, and 13).

Again, the implicit curves are much smoother than expected, since the deficiency of smoothness occurs only at the data locations.

If we place 5 points on each edge of the square $[-1, 1]^2$ we get Figures 14, 15, and 16.

I would be interesting to see how the method behaves in $3D$ for large numbers of data points. However, it should only be used as an intermediate stage to produce input data for other representations of the same object. For this purpose it will be sufficient to do local evaluation, clipping will not be required.

Another remarkable property of positive definite radial functions $\phi \in C^2 \cap \mathbf{PD}_d$ is

$$-\Delta_d \phi \in C^0 \cap \mathbf{PD}_d$$

for the d–variate Laplacian Δ_d applied to ϕ as $\phi(\|\cdot\|_2)$, resulting again in a univariate positive definite radial function $-\Delta_d \phi$. Furthermore, this operation preserves compact supports and generalizes to other elliptic operators when applying Fourier transform techniques. This allows collocation for elliptic boundary value problems of the form

$$\begin{aligned} -\Delta_d s &= f \text{ on } G \subset \mathbb{R}^d \\ s &= \varphi \text{ on } \partial G \end{aligned}$$

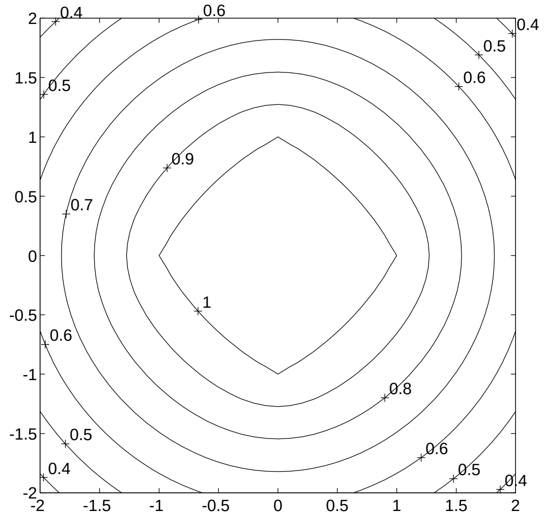


Fig. 11. Contour lines for Wu's C^0 function $\phi_{1,1}$.

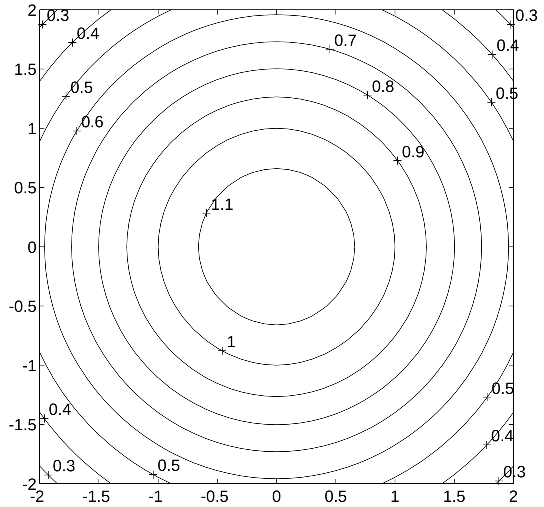


Fig. 12. Contour lines for Wu's C^2 function $\phi_{2,1}$.

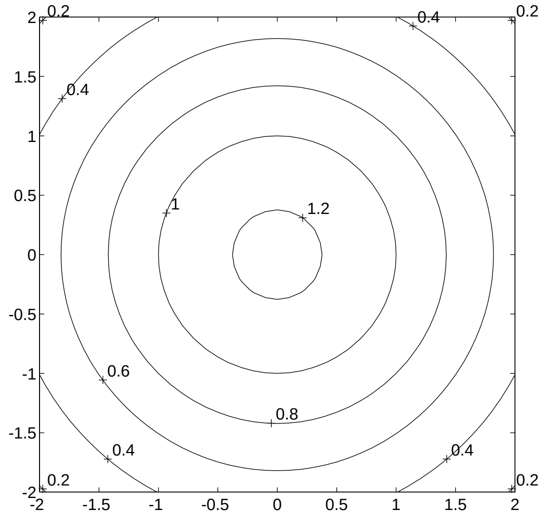


Fig. 13. Contour lines for Wu's C^4 function $\phi_{3,1}$.

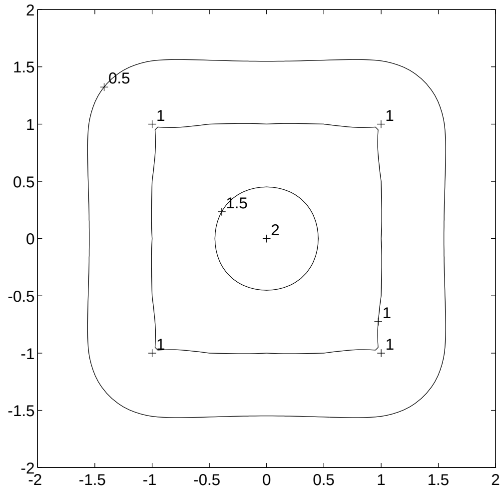


Fig. 14. Contour lines for Wu's C^0 function $\phi_{1,1}$.

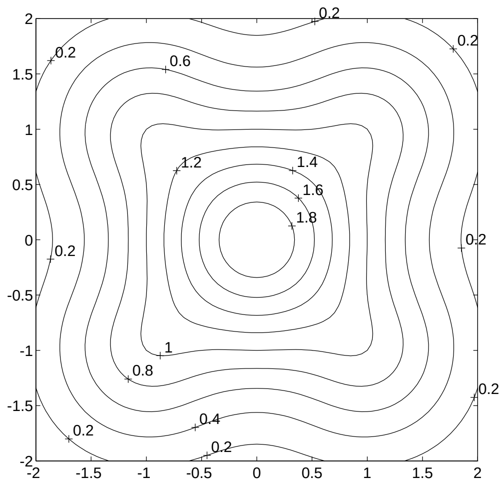


Fig. 15. Contour lines for Wu's C^2 function $\phi_{2,1}$.

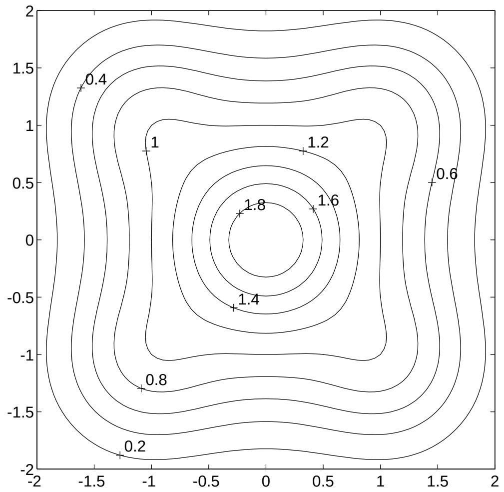


Fig. 16. Contour lines for Wu's C^4 function $\phi_{3,1}$.

using scalar–valued representations of the form

$$s(x)=\sum_{j=1}^{N} b_j(-\Delta_d)\phi(\|x-x_j\|_2)+\sum_{m=1}^{M} c_m\phi(\|x-y_m\|_2)$$

for collocation points $x_1,\dots,x_N\in G$ and $y_1,\dots,y_M\in\partial G$. The underlying theory can be found in [36]. So far there were no practical tests carried out with this method. This collocation technique is simpler to implement than finite elements, but a thorough error analysis still is missing.

§5. Efficiency Considerations

If a representation (2) uses a radial function $\Phi(\cdot)=\phi(\|\cdot\|_2)$ with ϕ compactly supported on $[0,\rho]\subset\mathbb{R}_{\geq 0}$, the evaluation of $s(x)$ for a given $x\in\mathbb{R}^d$ requires quick access to all x_j with $\|x-x_j\|_2\leq\rho$. This is a variation of the "k nearest neighbors" problem of computational geometry [25]. If evaluation of $s(x)$ for a very large number of points x of a compact set $\Omega\subset\mathbb{R}^d$ is required, a preprocessing step using a space decomposition technique should be implemented. Such things are common practice in Computer Graphics software, for instance to speed up ray tracing algorithms [10]. For completeness, we add a sketch of the basic technique and relate it to the computational efficiency of solving the underlying sparse linear system.

Assume Ω to be contained in a large L_∞ box $B_0\subset\mathbb{R}^d$ which in turn is split into smaller L_∞ boxes $B_1,\dots,B_M$. For each B_m our preprocessing step should produce a list L_m satisfying

$$L_m\supseteq\{j:1\leq j\leq N,\ K_\rho(x_j)\cap B_m\neq\emptyset\},$$

where we used $K_\rho(x):=\{y\in\mathbb{R}^d:\|y-x\|_2\leq\rho\}$. If the lists $L_1,\dots,L_M$ are available, the evaluation of $s(x)$ for any given $x\in\Omega\subseteq B_0\subset\mathbb{R}^d$ consists of the steps

1. Find some m with $x\in B_m$.
2. Evaluate (2) with j running over L_m.

Note that step 1 can be done very efficiently for regular gridded structures laid over Ω or for hierarchical space decompositions by median splits, for instance. Note that we do allow the B_m to overlap and to be of different size.

It remains to indicate how to build up the lists $L_1,\dots,L_M$ by preprocessing. This can be done by the following algorithm:

Let all L_m be empty.
for $1\leq j\leq\mathbb{N}$ do
  find all m with $B_m\cap K_\rho(x_j)\neq\emptyset$ and
  add j to L_m for these m.
end do.

For regular splittings of B_0 into boxes $B_1,\dots,B_M$ the *find* part of this algorithm can be implemented very effectively. For hierarchical splittings we recommend more sophisticated techniques that recursively construct the boxes together with the lists by splitting boxes B that contain "too many" entries in the list

$$L(B) \supseteq \{j : 1 \le j \le N,\ B \cap K_\rho(x_j) \neq \emptyset\}.$$

This can be started with $L(B_0) = \{1, \ldots, N\}$ and carried out by median splits [10]. We leave details to the experienced graphics programmers among the readers, but we add some hints to the proper choice of the parameters ρ and M for large N.

This is a rather delicate problem because small values of ρ will have disastrous effects on the reproduction quality of geometric objects while leading to very efficient evaluation algorithms and stable solutions of the equations (4). Details on this intrinsic relation between interpolation errors and condition numbers can be found in [29].

After some practical experience we found the following guidelines to be useful:

a) First fix a function ϕ by consideration of smoothness requirements for s in (2). This is a crucial step, because from a theoretical point of view (see [19], [26]) each ϕ is an optimal choice for a specific space of smooth objects.

b) If N is small, say $N \le 200$ for a standard 1994 workstation, then don't care about compact supports. Depending on smoothness requirements, try thin–plate splines or multiquadrics, but for the latter one should make sure that the constant c in $\phi(r) = (c^2 + r^2)^{\beta/2}$ is roughly proportional to the minimum distance of interpolation points in the sample. Otherwise the condition of A in (3) causes problems.

c) If N is large and if the data set is very unevenly distributed, try a hierarchical approach or different supports for ϕ around each x_j. Details of this will be given at the end of this section.

d) If N is large and if the data set is well distributed, one can fix the support radius ρ in such a way that each $K_\rho(x_i)$ contains roughly the same number, say n, of data points from $X = \{x_1, \ldots, x_N\}$. Small values of n and ρ will increase efficiency and decrease reproduction quality. Furthermore, the proper choice of n and ρ will depend on ϕ, because smooth functions ϕ will have a rapid decrease towards the boundary of their support. Thus they practically behave like functions with a much smaller support. Anyway, preprocessing should be done such that each list L_m contains $\mathcal{O}(n)$ points. This will imply that a large number $M = \mathcal{O}(N)$ of boxes is required.

An evaluation of $s(x)$ will then require $\mathcal{O}(n)$ instead of $\mathcal{O}(N)$ operations, if preprocessing has been done. Preprocessing itself needs $\mathcal{O}(N \cdot n)$ operations for regular splits. Furthermore, there is an $\mathcal{O}(M \cdot n) = \mathcal{O}(N \cdot n)$ storage requirement for the lists, and special precautions should be made to guarantee data locality when doing very many evaluations of s at nearby points $x \in \mathbb{R}^d$. In these considerations, we regard N, M, and n as large when compared to d and D.

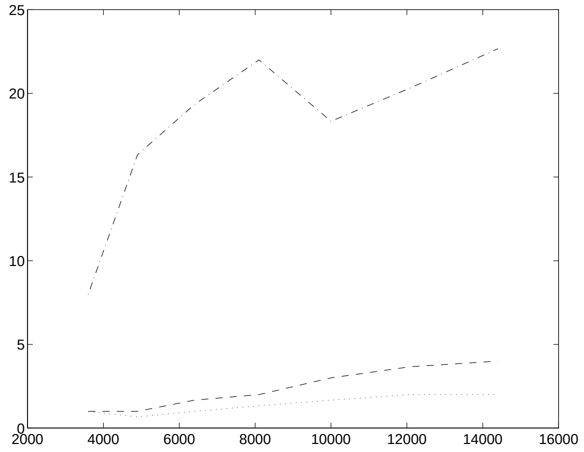


Fig. 17. System time in seconds versus N.

The solution of the system induced by (4) should make use of sparsity. Each matrix–vector multiplication requires $\mathcal{O}(N \cdot n)$ operations. If a conjugate gradient method is used, the number of such multiplications to reach a fixed precision will be dependent on the condition number κ of the matrix. This condition number κ in turn depends on n, not on N, if ρ is scaled with N such that there are $\mathcal{O}(n)$ elements in each $K_\rho(x_j)$. In particular, $\kappa(n)$ behaves like a power n^γ of n if ϕ is of limited smoothness, and the exponent γ increases with ϕ's smoothness. Details of the analysis of condition numbers for matrices arising from radial basis functions can be found, for instance in [3, 4, 13, 20, 28].

For practical purposes it will suffice to keep n bounded by, say, a number between 5 and 50 for standard applications, even if N is very large. Recall that the examples of section 3 were carried out with $n = 9$. Then the conjugate gradient method for solving (4) up to a fixed precision will need $\mathcal{O}(N)$ operations, keeping the whole complexity at $\mathcal{O}(N)$ operations. See figure 17 for mean values of 3 test runs with n fixed and N increasing, where the results for preprocessing (dashdotted line) varied wildly because of frequent heap access, but clearly stayed at $\mathcal{O}(N)$ for large N. The dashed line is the time for the conjugate gradient method, while the dotted line represents the computation of N values of $s(x)$ at random locations. Note, however, that these efficiency considerations hold only for fixed precision requirements and varying sample sizes. It is much more involved to vary the required precision and to calculate the necessary values of N, n, and ρ in terms of the precision, provided that a fixed compact domain $\Omega \subset \mathbb{R}^d$ is gradually filled up with data points.

For cases with very unevenly distributed data points one can try to replace (2) by

$$s(x) = \sum_{j=1}^{N} b_j \Phi_{\rho_j}(x - x_j)$$

where $\Phi_\rho(x) = \phi(\rho \cdot \|x\|)$ allows to scale each radial basis function by some positive quantity ρ_j that should be roughly proportional to the distance of x_j to its nearest n neighbors. If ϕ is compactly supported, this approach will again lead to an N by N sparse matrix with $\mathcal{O}(n)$ entries per row or column, but so far there is no general proof of the nonsingularity of the matrix.

Another possibility for unevenly distributed points will be a split of $X = \{x_1, \ldots, x_N\}$ into an evenly distributed $X^{(1)}$ and the rest $X^{(2)} = X \setminus X^{(1)}$. One can then interpolate f on $X^{(1)}$ by some $s^{(1)}$ using a function $\phi^{(1)}$ and the standard technique. For interpolation of $f - s^{(1)}$ on X one can now choose a $\phi^{(2)}$ with smaller support and make use of the the fact that $f - s^{(1)}$ already vanishes at points of $X^{(1)}$. The final result will then be $s^{(1)} + s^{(2)}$. If the data are dense in certain areas and well distributed elsewhere, this hierarchical approach will be easily applicable. Similar techniques are common practice in multivariate data analysis [2]. Large–scale "trends" are handled first and "details" are added on by rapidly varying local functions. The overall analysis of such an hierarchical approach still is missing.

Acknowledgements. The author gratefully appreciates help in proofreading by A. Iske, and help in typesetting by P. Trapp.

References

1. Abramowitz, M., and Stegun, I.A., *Handbook of Mathematical Functions*, Dover, New York, 1970.
2. Agterberg, F. P *Geomathematics. Mathematical Background and Geo–Sciences Applications*, Elsevier, New York, 1974.
3. Ball, K., Eigenvalues of Euclidean Distance Matrices, Journal of Approximation Theory, **68** (1992), 74–82.
4. Baxter, B.J.C., Norm estimates for inverses of Distance Matrices, *Mathematical Methods in Computer Aided Geometric Design*, Academic Press, New York, T. Lyche and L.L. Schumaker (eds.), 1989, 13–18.
5. Buhmann, M.D., New Developments in the Theory of radial basis function interpolation, *Multivariate Approximations: From CAGD to Wavelets*, K. Jetter and F.I. Utreras (eds), World Scientific, London, 1993, 35–75.
6. Buhmann, M., and Ron, A., Radial basis functions L^p–approximation orders with scattered centres, *Wavelets, Images, and Surface Fitting*, P.–J. Laurent and A. Le Méhauté and L.L. Schumaker (eds.), AKPeters, Boston, 1994, 93–112.
7. Dyn, N., Interpolation of scattered data by radial functions, *Topics in Multivariate Approximation*, C.K. Chui and L.L. Schumaker and Florencio Utreras (eds.), Academic Press, Boston, 1987, 47-61.
8. Dyn, N., Interpolation and Approximation by Radial and Related Functions, *Approximation Theory VI*, C.K. Chui and L.L. Schumaker and J.D. Ward (eds.) Academic Press, Boston, 1989.

9. Dyn, N., Levin, D., and Rippa, S., Surface interpolation and smoothing by "thin plate" splines, *Approximation Theory VI*, Spline Functions and Applications, S.P.Singh (ed.) NATO ASI Series, 1983, 445–449.
10. Foley, J. D., Dam, A. van, Feiner, S. K., and Hughes, J. F., *Computer Graphics, Principles and Practice*, The Systems Programming Series, Addison-Wesley, Reading, Mass., 1990.
11. Franke, R., *Scattered Data Interpolation. Test of Some methods*, American Mathematical Society, 1982, 181–200.
12. Iske, A., Charakterisierung bedingt positiv definiter Funktionen für multivariate Interpolationsmethoden mit radialen Basisfunktionen, Ph. D. Dissertation, Göttingen, 1994.
13. Jetter, K., Riesz Bounds in Scattered Data Interpolation and L_2 Approximation, *Multivariate Approximations: From CAGD to Wavelets*, K. Jetter and F. Utreras (eds), World Scientific, London, 1993, 167–177.
14. Jetter, K., Multivariate Approximation from the cardinal interpolation point of view, *Approximation Theory VII*, E.W. Cheney and C.K. Chui and L.L. Schumaker (eds.), Academic Press, New York, 1992, 131–161.
15. Jetter, K., and Stöckler, J., A Generalization of de Boor's Stability Result and Symmetric Preconditioning, preprint.
16. W. A. Light, Some aspects of Radial Basis Function Approximation, *Approximation Theory, Spline Functions and Applications*, S.P.Singh (ed.) NATO ASI Series **356** (1992), 163–190.
17. Light, W., Using radial functions on compact domains, *Wavelets, Images, and Surface Fitting*, P.–J. Laurent, A. Le Méhauté and L.L. Schumaker, AKPeters, Boston, 1994, 351–370.
18. G. Matheron, Splines and Kriging: their formal equivalence, *Down to the earth statistics*, D.F. Merriam (ed.), Academic Press, New York, 1981, Geology Contributions, **8**, 77–95.
19. Madych, W.R. and Nelson, S.A., Multivariate Interpolation and conditionally positive definite functions, Approx. Th. and its Applications, **4**, 1988, 77–89.
20. Narcowich, F.J. and Ward, J.D, Norm Estimates for the Inverses of a General Class of Scattered–Data Radial–Function Interpolation Matrices, Journal of Approximation Theory, **69** (1992), 84–109.
21. Powell, M. J. D., *Radial basis functions for multivariable interpolation: a review in Numerical Analysis*, D. F. Griffiths and G. A. Watson (eds.), Longman Scientific & Technical (Harlow), 1987, 223-241.
22. Powell, M. J. D., Tabulation of thin plate splines on a very fine two–dimensional grid, *Numerical Methods of Approx. Th.*, D. Braess and L.L. Schumaker (eds.), Birkhäuser, Basel, 1992, 221–244.
23. Powell, M.J.D., Univariate Multiquadric Interpolation: Some Recent Results, *Curves and Surfaces*, P.J. Laurent and A. Le Méhauté and L.L. Schumaker (eds.), Academic Press, 1991, 371–382.
24. Powell, M.J.D., The theory of radial basis function approximation in 1990, Department of Applied Mathematics and Theoretica Physics, University of Cambridge, 1992, DAMTP 1990/NA11.

25. Preparata, F. P., and Shamos, M. I., *Computational Geometry*, Texts and Monographs in Computer Science, D. Gries (ed.), Springer, New York, 1985.
26. Schaback, R., Comparison of Radial Basis Function Interpolants, *Multivariate Approximations: From CAGD to Wavelets*, K. Jetter and F. Utreras, (eds), World Scientific, London, 1993, 293–305.
27. Schaback, R., Reproduction of Polynomials by Radial Basis Functions, *Wavelets, Images, and Surface Fitting*, P.–J. Laurent, A. Le Méhauté and L.L. Schumaker, AKPeters, Boston, 1994, 459–466.
28. Schaback, R., Lower Bounds for Norms of Inverses of Interpolation Matrices for Radial Basis Functions, to appear in *Journal of Approximation Theory*, 1994.
29. Schaback, R., Error Estimates and Condition Numbers for Radial Basis Function Interpolation, to appear in *AICM*, 1994.
30. Schaback, R., Wendland, H., Special Cases of Compactly Supported Radial Basis Functions, preprint, Göttingen, 1994.
31. Schaback, R., Wu, Zong-min, Operators on Radial Functions, in preparation, 1994.
32. Stein, E. M., and Weiss, G,. *Introduction to Fourier Analysis on Euclidean Spaces.*, Princeton University Press, 1971.
33. Stewart, J., Positive definite functions and generalizations, an historical survey, Rocky Mountain J. Math., **6** (1976), 409–434.
34. Wendland, Holger, Ein Beitrag zur Interpolation mit radialen Basisfunktionen, Diplomarbeit, Universität Göttingen, 1994.
35. Wu, Z., Die Kriging–Methode zur Lösung mehrdimensionaler Interpolationsprobleme, Dissertation, Universität Göttingen, 1986.
36. Wu, Z., Hermite–Birkhoff Interpolation of scattered data by radial basis functions, J. Approx. Th and its Appl., **8:2** (1992), 1–10.
37. Wu, Z., Multivariate compactly supported positive definite radial functions, preprint, 1994.
38. Wu, Z., Characterization of Positive Definite Radial Functions, Übersichtsvortrag Ulvik 1994, erscheint im zugehörigen Tagungsband, 1994.
39. Wu, Zong-min, R. Schaback, Local error estimates for radial basis function interpolation of scattered data, *Ima J. Numer. Anal.*, **13** (1993), 13–27.

Robert Schaback
Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestraße 16-18
D-37083 Göttingen
Germany
schaback@namu01.gwdg.de