Multivariate Interpolation by Polynomials and Radial Basis Functions

Robert Schaback

April 24, 2002

Abstract

In many cases, multivariate interpolation by smooth radial basis functions converges towards polynomial interpolants, when the basis functions are scaled to become "wide". In particular, examples show that interpolation by scaled Gaussians seems to converge towards the de Boor/Ron "least" polynomial interpolant. The paper starts by providing sufficient criteria for the convergence of radial interpolants, and the structure of the polynomial limit interpolation is investigated to some extent. The results lead to general questions about "radial polynomials" $\|x-y\|_2^{2\ell}$ and the properties of spaces spanned by linear combinations of their shifts. For their investigation a number of helpful results are collected. In particular, the new notion of a *discrete moment basis* turns out to be rather useful. With these tools, a variety of well–posed multivariate polynomial interpolation processes can be formulated, leading to interesting questions about their relationships. Part of them can be proven to be "least" in the sense of de Boor and Ron. Finally, the paper generalizes the de Boor/Ron interpolation process and shows that it occurs as the limit of interpolation by Gaussian radial basis functions. As a byproduct, we get a stable method for preconditioning the matrices arising with interpolation by smooth radial basis functions.

Keywords: radial basis functions, Gaussian, moment conditions.

Classification: 41A15, 41A25, 41A30, 41A63, 65D10

1 Introduction

Let $\phi : [0,\infty) \to I\!R$ be a smooth radial basis function that can be written as

$$\phi(r) = f(r^2) \text{ with a smooth function } f : I\!R \to I\!R,$$

and in particular we have in mind the Gaussians and inverse multiquadrics, i.e.

$$\phi(r) = \exp(-r^2) \text{ and } \phi(r) = (1+r^2)^{\beta/2},\ \beta < 0.$$

We scale ϕ in such a way that the functions get wider, i.e. we define

$$\phi_c(r) := \phi(r\sqrt{c}) = f(cr^2),\ c, r \geq 0 \tag{1}$$

and since we want to consider small c, we assume that f is analytic around zero.

We fix a set $X = \{x_1, \ldots, x_M\} \subset I\!R^d$ of scattered centers for interpolation, and consider the behaviour of the Lagrange interpolation basis for $c \to 0$. It is obtainable as the solution $(u_1^c(x), \ldots, u_M^c(x)) \in I\!R^M$ of the system

$$\sum_{j=1}^{M} \phi_c(\|x_j - x_k\|_2) u_j^c(x) = \phi_c(\|x - x_k\|_2) \text{ for all } 1 \leq k \leq M. \tag{2}$$

By a surprising observation of Driscoll/Fornberg [6] and Danzeglocke [5] there are many cases where the limits of the Lagrange basis functions $u_j^c(x)$ for $c \to 0$ exist and are multivariate polynomials in x. Our first goal is to prove this fact under certain assumptions on ϕ and X. From a recent paper by Fornberg, Wright and Larsson [7] it is known that convergence may depend critically on the geometry of X and certain properties of ϕ. We study these connections to some extent, and we want to characterize the final polynomial interpolant in a way that is independent of the limit process. This poses some interesting questions about multivariate polynomials and geometric properties of scattered data sets. The investigations about the limit polynomial interpolants are based on "radial" polynomials of the form $\|x-y\|_2^{2\ell}$ because the matrix entries in (2) have series expansions

$$\phi_c(\|x_j-x_k\|_2)=f(c\|x_j-x_k\|_2^2)=\sum_{\ell=0}^{\infty}\frac{f^{(\ell)}(0)}{\ell!}c^\ell\|x_j-x_k\|_2^{2\ell}$$

and the right–hand side contains

$$\phi_c(\|x-x_k\|_2)=f(c\|x-x_k\|_2^2)=\sum_{\ell=0}^{\infty}\frac{f^{(\ell)}(0)}{\ell!}c^\ell\|x-x_k\|_2^{2\ell}.$$

Thus this paper contains a somewhat nonstandard approach to multivariate interpolation, namely via linear combinations of "radial polynomials". In particular, we define two different classes of multivariate polynomial interpolation schemes that can be formulated without recurring to limits of radial basis functions. Examples show that the various methods are actually different. For their analysis, some useful theoretical notions are introduced, i.e. "discrete moment conditions" and "discrete moment bases". To establish the link from interpolation by scaled Gaussians to the de Boor/Ron "least" polynomial interpolation [2, 3, 4], we generalize the latter and, in particular, introduce a scaling and relate the theory to reproducing kernel Hilbert spaces. Using the new notion of a "discrete moment basis" we can prove that the algorithm of de Boor and Ron is the limit of radial basis function interpolation using the Gaussian kernel, if the kernel gets "wide". Finally, we prove that properly scaled discrete moment bases can be used for preconditioning the systems arising in radial basis function interpolation.

2 Limits of Radial Basis Functions

Because we shall be working with determinants, we fix the numbering of the points in X now. For a second ordered set $Y=\{y_1,\dots,y_M\}\subset I\!R^d$ with the same number M of data points we define the matrix

$$A_{c,X,Y}:=(\phi_c(\|x_j-y_k\|_2))_{1\le j,k\le M}=\left(f(c\|x_j-y_k\|_2^2)\right)_{1\le j,k\le M}.$$

Note that $A_{c,X,X}$ is symmetric and has a determinant that is independent of the order of the points in X. If ϕ is positive definite, the matrices $A_{c,X,X}$ are positive definite and have a positive determinant for all $c>0$.

Since f is analytic around the origin, the matrices $A_{c,X,Y}$ have a determinant with a convergent series expansion

$$\det A_{c,X,Y}=\sum_{k=0}^{\infty}c^k p_k(X,Y) \tag{3}$$

for small c, where the functions $p_k(X,Y)$ are polynomials in the points of X and Y. In particular they are sums of powers of terms of the form $\|x_j-y_k\|_2^2$. They can be determined by symbolic computation, and we shall give an explicit formula in section 3 and prove the upper bound $2k$ for their total degree in Lemma 5.

We define $X_j:=X\setminus\{x_j\}$, where x_j is deleted and the order of the remaining points is kept. Furthermore, in the sets $X_j(x):=(X\setminus\{x_j\})\cup\{x\},\ 1\le j\le M$ the point x_j is replaced by x, keeping the order.

The general structure of Lagrange basis functions is described by a standard technique:

Lemma 1 *[7] The Lagrange basis functions $u_j^c(x),\ 1\le j\le M$ for interpolation in $X=\{x_1,\dots,x_M\}$ by a scaled positive definite radial function ϕ_c have the form*

$$u_j^c(x):=\frac{\det A_{c,X,X_j(x)}}{\det A_{c,X,X}}=\frac{\sum_{k=0}^{\infty}c^k p_k(X,X_j(x))}{\sum_{k=0}^{\infty}c^k p_k(X,X)},\ 1\le j\le M. \tag{4}$$

Proof: The quotient of determinants is in the span of the functions $\phi_c(\|x-x_j\|_2),\ 1\le j\le M$, and it satisfies $u_j(x_k)=\delta_{jk},\ 1\le j,k\le M$. Since interpolation is unique, we are done. □

From (4) it is clear that the convergence behaviour of the Lagrange basis function $u_j^c(x)$ for $c\to 0$ crucially depends on the smallest values of k such that the real numbers $p_k(X,X)$ or $p_k(X,X_j(x))$ are nonzero. Examples show that this number in turn depends on the geometry of X, getting large when the set "degenerates" from "general position".

Definition 1 *Let $\kappa(d,M)$ be the minimal $k \geq 0$ such that the multivariate polynomial $X \mapsto p_k(X,X)$ is nonzero on the space R^{Md}. A set $X = \{x_1,\ldots,x_M\} \subset I\!R^d$ is* in general position with respect to *ϕ if $p_{\kappa(d,M)}(X,X) \neq 0$. A set $X = \{x_1,\ldots,x_M\} \subset I\!R^d$ has a* degeneration order *j with respect to ϕ if*

$$p_k(X,X) = 0 \text{ for all } 0 \leq k < \kappa(d,M) + j.$$

The maximal degeneration order *of a set $X = \{x_1,\ldots,x_M\} \subset I\!R^d$ will be denoted by $\delta(X)$. We then have*

$$\begin{array}{rclrl} p_k(X,X) & = & 0 & \text{for all } 0 \leq & k < \kappa(d,M) + \delta(X) \\ p_k(X,X) & \neq & 0 & & k = \kappa(d,M) + \delta(X). \end{array}$$

The degeneration order $\delta(X)$ is dependent on ϕ and the geometry of X. For convenience, we also use the notation

$$k_0(X) := \kappa(d,M) + \delta(X) \tag{5}$$

to describe the smallest $k \geq 0$ such that $p_k(X,X) \neq 0$. If ϕ is positive definite, we can conclude that $p_{k_0(X)}(X,X) > 0$ holds for all X. With this notion, the formula (4) immediately yields

Theorem 1 *[7] If $x \in I\!R^d$ and $j \in \{1,\ldots,M\}$ are such that*

$$p_k(X, X_j(x)) = 0 \text{ for all } k < k_0(X), \tag{6}$$

then the limit of $u_j^c(x)$ for $c \to 0$ is the value of the polynomial

$$\frac{p_{k_0(X)}(X, X_j(x))}{p_{k_0(X)}(X,X)}. \tag{7}$$

If (6) fails, then the limit is infinite. □

In the paper [7] of Fornberg et. al. there are cases where (6) fails for certain geometries, e.g. when ϕ is a multiquadric (inverse or not), when the set X consists of 5 points on a line in $I\!R^2$ and when the evaluation point x does not lie on that line. Strangely enough, the observations in [7] lead to the conjecture that the Gaussian is the only radial basis function where (6) never fails when data are on a line and evaluation takes place off that line. However, at the end of the paper we shall finish the proof of part of a related statement:

Theorem 2 *Interpolation with scaled Gaussians always converges to the de Boor/Ron polynomial interpolant when the Gaussian widths increase.*

The proof needs a rather special technique, and thus we postpone it to the penultimate section, proceeding now with our investigation of convergence in general. Unfortunately, condition (6) contains an unsymmetric term, and we want to replace it by

$$k_0(X_j(x)) \geq k_0(X), \text{ i.e. } \delta(X_j(x)) \geq \delta(X), \text{ i.e. } .p_k(X_j(x), X_j(x)) = 0 \text{ for all } k < k_0(X). \tag{8}$$

Then we can extend results by Fornberg et al. in [6, 7].

Theorem 3 *If the degeneration order $\delta(X)$ of X is not larger than the degeneration order $\delta(X_j(x))$ of $X_j(x)$, then the polynomial limit of the Lagrange basis function $u_j^c(x)$ for $c \to 0$ exists. In particular, convergence takes place when X is in general position with respect to ϕ.*

Proof: We assert boundedness of $u_j^c(x)$ for $c \to 0$ and then use Theorem 1. Let us denote the standard power function for interpolation on data X and evaluation at x by $P_X(x)$ and let us write $\|.\|_{\phi_c}$ for the norm in the native space of ϕ_c (see e.g. [9] for a short introduction). Then

Lemma 2 *The standard error bound of radial basis function interpolation yields the bound*

$$|u_j^c(x)| \leq P_{X_j}(x) \|u_j^c\|_{\phi_c} \tag{9}$$

for all $x \in I\!R^d$, all $c > 0$ and all j, $1 \leq j \leq M$.

Proof: Zero is the interpolant to u_j^c on $X_j = X \setminus \{x_j\}$. □

Lemma 3

$$\|u_j^c\|_{\phi_c}^2 = \frac{\det A_{c,X_j,X_j}}{\det A_{c,X,X}} \tag{10}$$

for $1 \leq j \leq M, c > 0$.

Proof: If α_j is the coefficient of $\phi(\|x-x_j\|_2)$ in the representation of u_j^c, we have

$$\|u_j^c\|_{\phi_c}^2 = \alpha_j,$$

because of the general fact that an interpolant s to data $f(x_k),\ 1 \leq k \leq M$ has the native space norm

$$\|s\|_{\phi_c}^2 = \left\| \sum_{j=1}^M \alpha_j \phi_c(\|x-x_j\|_2) \right\|_{\phi_c}^2 = \sum_{j=1}^M \alpha_j f(x_j).$$

Then (10) follows from Cramer's rule applied to the interpolation problem with Kronecker data $\delta_{jk},\ 1 \leq k \leq M$ solved by u_j^c. □

Lemma 4 *The power function has the representation*

$$P_X^2(x) = \frac{\det A_{c,X\cup\{x\},X\cup\{x\}}}{\det A_{c,X,X}}.$$

Proof: By expansion of the numerator, using (4) and the representation

$$P_X^2(x) = \phi_c(0) - \sum_{j=1}^M u_j^c(x)\phi_c(\|x-x_j\|_2).$$

This form is somewhat nonstandard. It follows from the optimality property of the power function, and it can be retrieved from [10], p. 92, (4.3.14). □

To finish the proof of Theorem 3, the above results yield

$$P_{X_j}^2(x_j) = \frac{\det A_{c,X,X}}{\det A_{c,X_j,X_j}}, \quad \|u_j^c\|_{\phi_c}^2 = \frac{1}{P_{X_j}^2(x_j)}$$

and

$$|u_j^c(x)| \leq P_{X_j}(x)\|u_j^c\|_{\phi_c} \leq \frac{P_{X_j}(x)}{P_{X_j}(x_j)}.$$

With the representation of the power function via determinants we get

$$(u_j^c(x))^2 \leq \frac{P_{X_j}^2(x)}{P_{X_j}^2(x_j)} = \frac{\det A_{c,X_j(x),X_j(x)}}{\det A_{c,X,X}} \tag{11}$$

The numerator and denominator of the right–hand side contain sets of M points each. If we assume (8), we arrive at

$$(u_j^c(x))^2 \leq \frac{\sum_{k=k_0(X_j(x))}^\infty c^k p_k(X_j(x), X_j(x))}{\sum_{k=k_0(X)}^\infty c^k p_k(X,X)} < \infty$$

which concludes the proof of Theorem 3. □

Remark. The first part of (11) is an interesting bound on Lagrange basis functions in radial basis function interpolation. If the set X is formed recursively by adding to X the point x_{M+1} where $P_X(x)$ is maximal (this adds the data location where the worst–case error occurs), one gets a sequence of Lagrange basis functions that is strictly bounded by 1 in absolute value. The implications on Lebesgue constants and stability of the interpolation process should be clear, but cannot be pursued here.

3 Basic Polynomial Determinants

To derive a formula for the polynomials p_k in (3) we need the expansion

$$f(z) = \sum_{k=0}^\infty f_k z^k \tag{12}$$

of f around the origin. If ϕ is positive definite, we know by the standard Bernstein-Widder representation (see [11] for a short summary) that all $(-1)^k f_k$ are positive. Furthermore, we use the standard notation for determinants

$$\det\, (b_{ij})_{1\leq i,j\leq M} = \sum_{\pi\in\mathcal{S}_M} (-1)^\pi \prod_{j=1}^M b_{j\pi(j)}$$

where π varies over all permutations in the symmetric group $\mathcal{S}_M$ and $(-1)^\pi$ is the number of inversions in π. Then

$$\begin{aligned}
\det A_{c,X,Y} &= \sum_{\pi\in\mathcal{S}_M}(-1)^\pi\prod_{j=1}^M f(c\|x_j-y_{\pi(j)}\|_2^2)\\
&= \sum_{\pi\in\mathcal{S}_M}(-1)^\pi\prod_{j=1}^M\sum_{m=0}^\infty f_mc^m\|x_j-y_{\pi(j)}\|_2^{2m}\\
&= \sum_{\pi\in\mathcal{S}_M}(-1)^\pi\sum_{\rho_1=0}^\infty\sum_{\rho_2=0}^\infty\cdots\sum_{\rho_M=0}^\infty\prod_{j=1}^M\left(f_{\rho_j}c^{\rho_j}\|x_j-y_{\pi(j)}\|_2^{2\rho_j}\right)\\
&= \sum_{\pi\in\mathcal{S}_M}(-1)^\pi\sum_{\rho\in I\!N_0^M}f_\rho c^{|\rho|}\prod_{j=1}^M\|x_j-y_{\pi(j)}\|_2^{2\rho_j}\\
&= \sum_{\rho\in I\!N_0^M}f_\rho c^{|\rho|}\sum_{\pi\in\mathcal{S}_M}(-1)^\pi\prod_{j=1}^M\|x_j-y_{\pi(j)}\|_2^{2\rho_j}\\
&= \sum_{k=0}^\infty c^k\sum_{\substack{\rho\in I\!N_0^M\\|\rho|=k}}f_\rho d_\rho(X,Y)
\end{aligned}$$

with multi–index notation

$$\begin{aligned}
f_\rho &:= \textstyle\prod_{j=1}^M f_{\rho_j}\\
d_\rho(X,Y) &:= \det\left(\|x_i-y_j\|_2^{2\rho_i}\right)_{1\le i,j\le M}\\
p_k(X,Y) &:= \sum_{\substack{\rho\in I\!N_0^M\\|\rho|=k}}f_\rho d_\rho(X,Y).
\end{aligned}$$

To see a bound on the degree, consider $|\rho|=k$ and conclude that

$$d_\rho(X,Y):=\det\left(\|x_i-y_j\|_2^{2\rho_i}\right)_{1\le i,j\le M}=\sum_{\pi\in\mathcal{S}_M}(-1)^\pi\prod_{j=1}^M\|x_j-y_{\pi(j)}\|_2^{2\rho_j}$$

has total degree at most $2|\rho|=2k$. Altogether we have

Lemma 5 *The polynomials $p_k(X,Y)$ have maximal degree $2k$ as polynomials in X and Y.* □

In Lemma 11 we shall provide a better result, but it requires more tools. We can also deduce that $k_0(X)$ for $|X|=M$ increases with M. In particular, we get

$$M\le\binom{2k_0(X)+d}{d}$$

from

Lemma 6 *If $p_k(X,Y)$ is nonzero for some special sets X,Y with $M=|X|=|Y|$, then $M\le\binom{2k+d}{d}$. Conversely, if $M>\binom{2k+d}{d}$, then $p_k(X,Y)=0$ for all sets X,Y with $M=|X|=|Y|$.*

Proof: If some $d_\rho(X,Y)$ is nonzero for $M=|X|=|Y|$, there are M linearly independent d–variate polynomials of degree at most $2|\rho|=2k$. This proves the first assertion, because these polynomials span a space of dimension $\binom{2k+d}{d}$. The second assertion is the contraposition of the first. □

Example 1 Let us look at some special cases that we prepared with MAPLE. We reproduce the results in [7], but we have a somewhat different background and notation. The 1D case with $M=2$ has in general $p_0(X,X)=0$, $p_1(X,X)=-2f(0)f'(0)(x_2-x_1)^2$. Thus $\kappa(1,2)=1$ and there is no degeneration except coalescence. The bound in Lemma 5 turns out to be sharp here. The case $M=3$ leads to $\kappa(1,3)=3$ with

$$p_3(X,X)=-2f'(0)(3f(0)f''(0)-f'(0)^2)(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2.$$

Geometrically, there is no degeneration except coalescence. The factor $3f(0)f''(0)-f'(0)^2$ could possibly lead to some discussion, but for positive definite ϕ it must be positive because we know that $f(0),-f'(0),f''(0)$ and $p_3(X,X)$ are positive. We find further $\kappa(1,4)=6$ with

$$\begin{aligned}
p_6(X,X) &= -\frac{4}{3}(3f''(0)^2-5f'(0)f'''(0))(3f(0)f''(0)-f'(0)^2)\\
&\quad(x_1-x_2)^2(x_1-x_3)^2(x_1-x_4)^2(x_2-x_3)^2(x_2-x_4)^2(x_3-x_4)^2.
\end{aligned}$$

The general situation seems to be $\kappa(1,M) = M(M-1)/2$ with $p_{\kappa(1,M)}$ being (up to a factor) the polynomial that consists of a product of all $(x_j - x_k)^2$ for $1 \le j < k \le M$, which is of degree $2\kappa(1,M) = M(M-1)$. Thus the maximal degree in Lemma 5 is actually attained again. Note that the 1D situation also carries over to the case when X and $X_j(x)$ lie on the same line in R^d.

Now let us look at 2D situations. The simplest nontrivial 2D case is for $M = 2$ when the evaluation is not on the line connecting the points of X. But from the 1D case we can infer

$$\kappa(2,2) = 1 = k_0(X) = k_0(X_j(x))$$

and do not run into problems, because we have Theorem 3. In particular, we find

$$p_1(X,X) = -2f(0)f'(0)((x_1 - x_2)^2 + (y_1 - y_2)^2).$$

Now we look at $M = 3$ in 2D. The general expansion yields $\kappa(2,3) = 2$ with

$$p_2(X,X) = 4f(0)f'(0)^2(\det B_X)^2$$

and B_X being the standard 3×3 matrix for calculation of barycentric coordinates based on X. Its determinant vanishes iff the points in X are collinear. Thus nondegeneracy of 3–point sets with respect to positive definite radial basis functions is equivalent to the standard notion of general position of 3 points in $I\!R^2$. To look for higher–order degeneration, we consider 3 collinear points now, and since everything is invariant under shifts and orthogonal transformations, we can assume that the data lie on the x–axis. This boils down to the 1D case, and we get $p_3(X,X) > 0$ with no further possibility of degeneration. But now we have to look into the first critical case, i.e. when X is collinear but $X_j(x)$ is not. This means that we evaluate the interpolant off the line defined by X. Theorem 3 does not help here. If we explicitly go back to (6), we still get convergence if we prove that $p_2(X, X_j(x)) = 0$ for all collinear point sets X and all $x \in I\!R^2$. Fortunately, MAPLE calculates

$$p_2(X, X_j(x)) = 4f(0)f'(0)^2(\det B_X)(\det B_{X_j(x)})$$

and thus there are no convergence problems. However, the ratio of the terms $p_3(X, X_j(x))$ and $p_3(X,X)$ now depends on ϕ.

Now we go for $M = 4$ in 2D and first find $\kappa(2,4) = 4$ from MAPLE, but it cannot factor the polynomial $p_4(X,X)$ properly or write it as a sum of squares. Taking special cases of 3 points not on a line, the polynomial $p_4(X,X)$ seems to be always positive except for coalescence. In particular, it does not vanish for 4 non–collinear points on a circle or a conic, as one would suspect. Taking cases of 3 points on a line, the polynomial $p_4(X,X)$ vanishes iff the fourth point also lies on that line. Thus there is some experience supporting the conjecture that nondegeneracy of 4 points in 2D with respect to positive definite functions just means that the points are not on a line. But if they are on a line, we find $k_0(X) = 6$ due to the 1D case, and thus $p_5(X,X)$ also vanishes. This is confirmed by MAPLE, and we now check the case where the points of X are on a line but those of $X_j(x)$ not. It turns out that then (6) holds for $k_0(X) = 6$, and the case does not show divergence.

The $M = 5$ situation in $I\!R^2$ has $\kappa(2,5) = 6$. The geometric interpretation of points in general position wrt. ϕ is unknown, because the zero set of $p_6(X,X)$ is hard to determine in general. If 4 points are fixed at the corners of the square $[0,1]^2$, and if the polynomial $\frac{2}{3}p_6(X,X)$ is evaluated for inverse multiquadrics with $\beta = -1$ as a function of the remaining point $x_5 = (\xi, \eta) \in I\!R^2$, we get the nonnegative polynomial

$$3\xi^2(1-\xi)^2 + 3\eta^2(1-\eta)^2 + (\xi(1-\xi) + \eta(1-\eta))^2$$

which vanishes only at the corners of the square. Thus it can be ruled out that degeneracy systematically occurs when 4 or 5 points are on a circle or three points are on a line. However, it turns out that $p_6(X,X)$ always vanishes if 4 points are on a line. The next coefficient $p_7(X,X)$, if calculated for 4 points on a line, vanishes either if the fifth point also lies on the line, or for $\beta = 0, 2, 3, 7$, or for coalescence. The final degeneration case thus occurs when all 5 points are on a line, and from 1D we then expect $k_0(X) = 10$.

Let us examine the divergence case described by Fornberg et. al. in [7]. It occurs when X consists of 5 points on a line, while evaluation takes place off that line. The 1D case teaches us that we should get $k_0(X) = 10$ for 5 collinear points, and MAPLE verifies this, at least for the fixed 5 collinear equidistant points on $[0,1] \times \{0\}$. However, we also find that

$$p_9(X, X_1(x)) = \frac{-9}{8388608}\eta^2\left(5f'(0)f'''(0) - 3f''(0)^2\right)\left(f(0)f''(0)f'''(0) + f'(0)f''(0)^2 - 2f'(0)^2f'''(0)\right)$$

for points $x = (\xi, \eta) \in \mathbb{R}^2$. If we put in multiquadrics, i.e. $f(t) = (1+t)^{\beta/2}$, we get the same result as in [7], which reads

$$p_9(X, X_1(x)) = \frac{-9}{268435456}\eta^2\beta^4(\beta-7)(\beta-2)^2$$

in our notation, proving that divergence occurs for multiquadrics except for the strange case $\beta = 7$. Another curiosity is that for multiquadrics the value $p_{10}(X, X)$ vanishes for the conditionally positive definite cases $\beta = 7$ and $\beta = 11.790$. As expected, this polynomial is positive for the positive definite cases, e.g. for negative β.

Checking the case where exactly 4 points of X are on a line, we find that (6) holds for $k_0(X) = 7$, and thus there is no convergence problem.

4 A Related Class of Polynomial Interpolation Methods

We can avoid all convergence problems if we boldly take (7) to define

$$u_j(x) := \frac{p_{k_0(X)}(X, X_j(x))}{p_{k_0(X)}(X, X)} \tag{13}$$

for all $1 \leq j \leq M$ and all $x \in \mathbb{R}^d$. The denominator will always be positive if we start with a positive definite function, and the discussion at the beginning of section 3 shows that the polynomials $p_k(X, Y)$ will always vanish if either X or Y have two or more coalescing points. Thus we get Lagrange interpolation polynomials for any kind of geometry. The result will be dependent on the function f and its Taylor expansion, and thus there is a full scale of polynomial interpolation methods which is available without any limit process. However, it is clear from (7) that polynomial limits of radial basis function interpolants, if they exist, will usually have the above form. It will be interesting to study how the technique of de Boor and Ron [2, 3, 4] relates to this. However, it uses a different truncation strategy.

Example 2 Let us check how the above technique overcomes the five–point degeneration case in Example 1. If we take the 5 equidistant points on $[0, 1] \times \{0\}$ and classical multiquadrics, the Lagrange basis function u_0 corresponding to the origin becomes

$$u_0(\xi, \eta) = \frac{1}{3}(\xi - 1)(4\xi - 3)(2\xi - 1)(4\xi - 1) + \frac{8}{21}\xi\eta^2(18\xi - 25),$$

and the second term is missing if we take the Gaussian. For $f(t) = \log(1 + t)$ the additional term is

$$\frac{-2}{3339}\eta^2(5195 + 15240\xi - 11424\xi^2 + 1008\eta^2).$$

There is dependence on f, but no degeneration. We simply ignore p_9 and focus on the quotient of values of p_{10}.

5 Point Sets, Polynomials, and Moments

Our results so far require knowledge and numerical availability of $k_0(X)$ and $p_{k_0(X)}(X, X_j(x))$. Section 3 gives a first idea for the evaluation of these quantities, but it still uses the limit process. It suggests that one looks at polynomials of the form $\|x - y\|_2^{2\ell}$, and we shall use this section to make a fresh start into multivariate polynomials and point sets. The relation to the earlier sections will turn up later.

Let $\mathbb{P}_m^d$ be the space of all d–variate real–valued polynomials of order (=degree+1) up to m, and let $X = \{x_1, \ldots, x_M\}$ be a fixed set of M points in $\mathbb{R}^d$. With the dimension $Q = \binom{m+d-1}{d}$ and a basis $p_1, \ldots, p_Q$ of $\mathbb{P}_m^d$ we can form the $Q \times M$ matrices P_m and the $M \times M$ matrices A_ℓ with

$$P_m := (p_i(x_j))_{1\leq i\leq Q,\ 1\leq j\leq M}, \quad A_\ell = ((-1)^\ell \|x_j - x_k\|_2^{2\ell})_{1\leq j,k\leq M}, \ \ell \geq 0 \tag{14}$$

to provide a very useful notion that is closely related to multivariate divided differences (see C. de Boor [1]):

Definition 2 *A vector $\alpha \in \mathbb{R}^M$ satisfies* discrete moment conditions of order m with respect to X *if $P_m\alpha = 0$ or*

$$\sum_{j=1}^{M} \alpha_j p(x_j) = 0 \textit{ for all } p \in \mathbb{P}_m^d$$

holds. These vectors form a linear subspace $MC_m(X) := \ker P_m$ of $\mathbb{R}^M$ for $M = |X|$.

Note that the definition involves all polynomials of order up to m, while the following involves radial polynomials of the form $\|x - x_j\|^{2\ell}$ for $0 \leq \ell < m$.

Theorem 4 *A vector $\alpha \in I\!R^M$ satisfies discrete moment conditions of order m with respect to X iff*

$$\alpha^T A_\ell \alpha = 0 \tag{15}$$

holds for all $0 \leq \ell < m$.

Note that the condition $A_\ell \alpha = 0$ would be more restrictive. It will come up later. The proof of Theorem 4 uses Micchelli's lemma from [8], which we restate here because we make frequent use of its proof technique later.

Lemma 7 *If $\alpha \in I\!R^M$ satisfies discrete moment conditions of order m, then the numbers $\alpha^T A_\ell \alpha$ vanish for all $\ell < m$ and $\alpha^T A_m \alpha$ is nonnegative. The latter quantity vanishes iff α satisfies discrete moment conditions of order $m + 1$.*

Proof: Let us take a vector $\alpha \in I\!R^M$ satisfying discrete moment conditions of order m, and pick any $\ell \leq m$ to form

$$\begin{aligned}
(-1)^\ell \alpha^T A_\ell \alpha &= \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k (\|x_j - x_k\|_2^2)^\ell \\
&= \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k \sum_{\ell_1+\ell_2+\ell_3=\ell} \|x_j\|_2^{2\ell_1} (-2(x_j, x_k))^{\ell_2} \|x_k\|_2^{2\ell_3} \\
&= \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k \sum_{\substack{\ell_1+\ell_2+\ell_3=\ell \\ \ell_2+2\ell_3 \geq m \\ \ell_2+2\ell_1 \geq m}} \|x_j\|_2^{2\ell_1} (-2(x_j, x_k))^{\ell_2} \|x_k\|_2^{2\ell_3}.
\end{aligned}$$

This value vanishes for $\ell < m$, and this also proves one direction of the second statement, if we formulate it for $m - 1$. For $\ell = m$ the two inequalities can only hold if $\ell_1 = \ell_3$. Thus $(-1)^\ell = (-1)^{\ell_2}$, and we can write in multi–index notation

$$\begin{aligned}
\alpha^T A_\ell \alpha &= \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k (-1)^\ell \|x_j - x_k\|_2^{2\ell} \\
&= \sum_{\substack{\ell_2=0 \\ \ell-\ell_2 \in 2Z\!\!Z}}^{\ell} 2^{\ell_2} \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k \|x_j\|^{\ell-\ell_2} \|x_k\|^{\ell-\ell_2} (x_j, x_k)_2^{\ell_2} \\
&= \sum_{\substack{\ell_2=0 \\ \ell-\ell_2 \in 2Z\!\!Z}}^{\ell} 2^{\ell_2} \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k \|x_j\|^{\ell-\ell_2} \|x_k\|^{\ell-\ell_2} \sum_{\substack{i \in I\!N_0^d \\ |i|=\ell_2}} x_j^i x_k^i \\
&= \sum_{\substack{\ell_2=0 \\ \ell-\ell_2 \in 2Z\!\!Z}}^{\ell} 2^{\ell_2} \sum_{\substack{i \in I\!N_0^d \\ |i|=\ell_2}} \sum_{j=1}^M \sum_{k=1}^M \alpha_j \alpha_k \|x_j\|^{\ell-\ell_2} \|x_k\|^{\ell-\ell_2} x_j^i x_k^i \\
&= \sum_{\substack{\ell_2=0 \\ \ell-\ell_2 \in 2Z\!\!Z}}^{\ell} 2^{\ell_2} \sum_{\substack{i \in I\!N_0^d \\ |i|=\ell_2}} \left(\sum_{j=1}^M \alpha_j \|x_j\|^{\ell-\ell_2} x_j^i \right)^2 \geq 0.
\end{aligned}$$

If this vanishes, all expressions

$$\sum_{j=1}^M \alpha_j \|x_j\|^{\ell-\ell_2} x_j^i$$

with $0 \leq \ell_2 \leq \ell$, $\ell - \ell_2 \in 2Z\!\!Z$, $i \in I\!N_0^d$, $|i| = \ell_2$ must vanish, and this implies that α satisfies discrete moment conditions of order $\ell + 1 = m + 1$. □

It is now easy to prove Theorem 4. If α satisfies discrete moment conditions up to order m, Micchelli's lemma proves that (15) holds. For the converse, assume that (15) is true for some $\alpha \in I\!R^M$ and proceed by induction. There is nothing to prove for order zero, and if we assume that we have the assertion up to order

$m-1 \geq 0$, then we use it to conclude that α satisfies discrete moment conditions of order $m-1$ because it satisfies (15) up to $\ell = m-1$. Then we apply Micchelli's lemma again on the level $m-1$, and since we have $\alpha^T A_{m-1}\alpha = 0$, we conclude that α satisfies discrete moment conditions of order m. □

There is another equivalent form of discrete moment conditions, taking the form of degree reduction of linear combinations of high–degree radial polynomials:

Lemma 8 *A vector $\alpha \in I\!R^M$ satisfies discrete moment conditions of order m, iff for all $2\ell \geq m$ the polynomials*

$$\sum_{k=1}^{M} \alpha_k \|x - x_k\|_2^{2\ell}$$

have degree at most $2\ell - m$.

Proof: Let us first assume that $\alpha \in I\!R^M$ satisfies discrete moment conditions of order m. We look at

$$\begin{array}{rl} & \displaystyle\sum_{k=1}^{M} \alpha_k \|x - x_k\|_2^{2\ell} \\ = & \displaystyle\sum_{k=1}^{M} \alpha_k \sum_{\ell_1+\ell_2+\ell_3=\ell} \|x\|_2^{2\ell_1}(-2(x,x_k))^{\ell_2}\|x_k\|_2^{2\ell_3} \\ = & \displaystyle\sum_{k=1}^{M} \alpha_k \sum_{\substack{\ell_1+\ell_2+\ell_3=\ell \\ \ell_2+2\ell_3 \geq m}} \|x\|_2^{2\ell_1}(-2(x,x_k))^{\ell_2}\|x_k\|_2^{2\ell_3} \end{array}$$

and this is of degree at most $2\ell_1 + \ell_2 = 2\ell - 2\ell_3 - \ell_2 \leq 2\ell - m$.

We now prove the converse and apply the same idea as in the proof of Micchelli's lemma to get

$$\begin{array}{rl} & \displaystyle\sum_{k=1}^{M} \alpha_k \|x - x_k\|_2^{2\ell} \\ = & \displaystyle\sum_{k=1}^{M} \sum_{\ell_1+\ell_2+\ell_3=\ell} \alpha_k \|x\|_2^{2\ell_1}(-2(x,x_k))^{\ell_2}\|x_k\|_2^{2\ell_3} \\ = & \displaystyle\sum_{\substack{\ell_2=0 \\ \ell-\ell_2=2\ell_1+2\ell_3 \in 2Z\!\!Z}}^{\ell} (-2)^{\ell_2} \sum_{\substack{i \in I\!N_0^d \\ |i|=\ell_2}} \|x\|_2^{2\ell_1} x^i \sum_{k=1}^{M} \alpha_k \|x_k\|_2^{2\ell_3} x_k^i \end{array}$$

for arbitrary vectors $\alpha \in I\!R^M$. If this is a polynomial of degree at most $2\ell - m$, then all sums

$$\sum_{k=1}^{M} \alpha_k \|x_k\|_2^{2\ell_3} x_k^i$$

with $2\ell_1 + \ell_2 > 2\ell - m$ or, equivalently, $\ell_2 + 2\ell_3 < m$ must vanish. Thus α satisfies moment conditions of order m. □

Note that in the above argument it suffices to pick just one ℓ with $2\ell \geq m$. Thus

Lemma 9 *A vector $\alpha \in I\!R^M$ satisfies discrete moment conditions of order m, iff for some ℓ with $2\ell \geq m$ the polynomial*

$$\sum_{k=1}^{M} \alpha_k \|x - x_k\|_2^{2\ell}$$

has degree at most $2\ell - m$. □

Now we use that the discrete moment spaces for a finite point set $X = \{x_1, \ldots, x_M\} \subset I\!R^d$ form a decreasing sequence

$$\cdots \subseteq MC_{m+1}(X) \subseteq MC_m(X) \subseteq \cdots \subseteq MC_0(X) = R^M. \tag{16}$$

This sequence must stop with some zero space at least at order M, because we can separate M points always by polynomials of degree $M-1$, using properly placed hyperplanes.

Definition 3 *For any finite point set $X = \{x_1, \dots, x_M\} \subset I\!R^d$ there is a unique largest natural number $\mu = \mu(X)$ such that $MC_\mu(X) \neq \{0\} = MC_{\mu+1}(X)$. We call $\mu(X)$ the* maximal discrete moment order *of X.*

With this notion, the sequence (16) can be written as

$$MC_{\mu+1}(X) = \{0\} \neq MC_\mu(X) \subseteq MC_{\mu-1}(X) \cdots \subseteq MC_0(X) = R^M. \tag{17}$$

There is a fundamental observation linked to the maximal discrete moment order.

Theorem 5 *If there is a polynomial interpolation process based on a set X, it cannot work exclusively with polynomials of degree less than $\mu(X)$.*

Proof: If we take a nonzero vector α from MC_μ, we see that it is in the kernel of all matrices P_m from (14) for all $m \leq \mu$. Thus these matrices can have full column rank M only if $m > \mu$. □.

The following notion is borrowed from papers of de Boor and Ron [2, 3, 4].

Definition 4 *We call a polynomial interpolation process for a point set X* least, *if it works with polynomials of degree at most $\mu(X)$.*

Remark: Below we shall see couple of least polynomial interpolation processes on X, including the one by de Boor and Ron.

We now go back to where we started from, and relate $\mu(X)$ with the quantity $k_0(X)$ defined in (5).

Lemma 10 *For all sets X and Y of M points in $I\!R^d$ we have*

$$p_k(X,Y) = 0 \text{ unless } 2k \geq \mu(Y)$$

and in particular $2k_0(X) \geq \mu(X)$.

Proof: Take a vector $\rho \in Z\!\!Z_0^M$ and the matrix

$$\left(\|x_i - y_j\|_2^{2\rho_i}\right)_{1 \leq i,j \leq M}.$$

We multiply by a nonzero vector $\alpha \in MC_\mu$ for $\mu := \mu(Y)$ and get

$$\begin{aligned}
& \sum_{j=1}^M \alpha_j \|x_i - y_j\|_2^{2\rho_i} \\
= & \sum_{j=1}^M \alpha_j \sum_{\ell_1+\ell_2+\ell_3=\rho_i} \|x_i\|^{2\ell_1}(-2(x_i^T y_j))^{\ell_2}\|y_j\|_2^{2\ell_3} \\
= & \sum_{\ell_1=0}^{\rho_i} \|x_i\|^{2\ell_1} \sum_{\substack{\ell_2+\ell_3=\rho_i-\ell_1 \\ \ell_2+2\ell_3\geq\mu}} \sum_{j=1}^M \alpha_j(-2(x_i^T y_j))^{\ell_2}\|y_j\|_2^{2\ell_3}
\end{aligned}$$

for all i, $1 \leq i \leq M$. Since $\ell_2 + 2\ell_3 \geq \mu$ means $2\rho_i - \mu \geq 2\ell_1 + \ell_2$, this vanishes for those i where $2\rho_i < \mu$. Thus the matrix can be nonsingular only if $2\rho_i \geq \mu$ for some i, and this implies $2\|\rho\|_\infty \geq \mu$. Since the polynomials $p_k(X,Y)$ are superpositions of determinants of such matrices with $2\|\rho\|_1 = 2k$, the assertion is proven. □

Lemma 11 *The Lagrange basis polynomials of* (7) *are of degree at most $2k_0(X) - \mu(X)$.*

Proof: We look at the above argument, but swap the meaning of X and Y there, replacing X by $X_j(x)$ and Y by X. The determinants vanish unless $2\|\rho\|_\infty \geq \mu(X)$, and the remaining terms are of degree at most

$$2\ell_1 + \ell_2 \leq 2\rho_i - \mu(X) \leq 2\|\rho\|_\infty - \mu(X) \leq 2\|\rho\|_1 - \mu(X) \leq 2k - \mu(X).$$

□

We note that there is a lot of leeway between the result of Lemma (11) and the actually observed degrees of the $p_{k_0(X)}(X, X_j(x))$. The latter seem to be bounded above by $\mu(X)$ instead of $2k_0(X) - \mu(X)$.

Theorem 6

$$k_0(X) = \sum_{j=1}^{\mu(X)} j(\dim MC_j - \dim MC_{j+1}) \geq \mu(X). \tag{18}$$

Proof: Let us take a nonzero vector $\alpha \in MC_\mu$ and evaluate the quadratic form

$$\begin{aligned} & \alpha^T A_{c,X,X}\alpha \\ = & \sum_{s=0}^{\infty} c^s f_s \alpha^T A_s \alpha \\ = & \sum_{s=\mu}^{\infty} c^s f_s \alpha^T A_s \alpha \\ = & c^\mu f_\mu \alpha^T A_\mu \alpha + \sum_{s=\mu+1}^{\infty} c^s f_s \alpha^T A_s \alpha. \end{aligned}$$

By Courant's minimum–maximum principle, this implies that $A_{c,X,X}$ has at least $\dim MC_\mu$ eigenvalues that decay at least as fast as c^μ to zero for $c \to 0$.

But there are no eigenvalues that decay faster than that. To see this, take for each $c > 0$ a normalized nonzero eigenvector α_c such that the unique smallest eigenvalue

$$\alpha_c^T A_{c,X,X}\alpha_c = \sum_{s=0}^{\infty} c^s f_s \alpha_c^T A_s \alpha_c =: \lambda_c$$

decays like c^μ or faster. The coefficients $\alpha_c^T A_s \alpha_c$ can increase with s only like $(\,\mathrm{diam}(X))^{2s}$, and thus we have a stable limit of the analytic function λ_c of c with respect to $c \to 0$. If we pick sequences of c's that converge to zero such that α_c converges to some nonzero normalized vector α, we see that necessary $\alpha \in MC_\mu$. But then λ_c cannot decay faster than c^μ for $c \to 0$. Going back to Courant's minimum–maximum principle, we now know that $A_{c,X,X}$ has precisely $\dim MC_\mu$ eigenvalues that decay exactly like c^μ to zero for $c \to 0$.

We can now repeat this argument on the subspace of $MC_{\mu-1}$ which is orthogonal to MC_μ. For each nonzero vector of this space, the quadratic form decays like $c^{\mu-1}$, and there are $\dim MC_{\mu-1} - \dim MC_\mu$ linear independent vectors with this property. Now we look for arbitrary vectors α_c that are orthogonal to the already determined $\dim MC_\mu$ eigenvectors of $A_{c,X,X}$ with eigenvalues of decay c^μ, and we assume that they provide eigenvalues with fastest possible decay. This decay cannot be of type c^μ or faster due to the assumed orthogonality, which allows passing to the limit. It must thus be of exact decay $c^{\mu-1}$. Induction now establishes the fact that for each j, $0 \le j \le \mu$ there are $\dim MC_j - \dim MC_{j+1}$ eigenvalues of $A_{c,X,X}$ with exact decay like c^j for $\to 0$. Thus the determinant decays exactly like the product of these, and this proves our assertion. □

Note that the above discussion fails to prove that the limiting polynomial interpolation process coming from a smooth radial basis function is least in all cases. We have to leave this problem open. Though $k_0(X)$ will exceed $\mu(X)$, for instance in 1D situations, there is plenty of cancellation in the polynomials $p_{k_0(X)}(X, X_j(x))$ that we have not accounted for, so far. On the other hand, we have not found any example where the polynomial limit of a radial basis function interpolation is not of least degree.

There is another interesting relation of μ to the spaces spanned by radial polynomials:

Lemma 12 *Define the M–vectors*

$$F_\ell(x) := \left((-1)^\ell \|x - x_k\|_2^{2\ell}\right)_{1\le k\le M}.$$

Then the $M \times M(s+1)$ matrix with columns $F_\ell(x_j)$, $1 \le j \le M$, $0 \le \ell \le s$ has full rank M if $s \ge \mu$, and μ is the smallest possible number with this property.

Proof: Assume that the matrix does not have full rank M for a fixed s. Then there is a nonzero vector $\alpha \in I\!R^M$ such that

$$A_\ell \alpha = 0 \text{ for all } 0 \le \ell \le s$$

and this implies discrete moment conditions of order $s+1$. Thus $s+1 \le \mu$. □

This teaches us that when aiming at interpolation by radial polynomials of the form $\|x - x_k\|_2^{2\ell}$ one has to go up to $\ell = \mu$ to get anywhere. But in view of Theorem 5 this reservoir of radial polynomials is way too large if we focus on the degree. We have to find a useful basis of an M–dimensional subspace of polynomials of degree at most μ, if we want a least interpolation method. The following notion will be very helpful for the rest of the paper.

Definition 5 *A discrete moment basis of R^M with respect to X is a basis*

$$\alpha^1, \dots, \alpha^M \text{ such that } \alpha^j \in MC_{t_j} \setminus MC_{t_j+1}$$

for the decomposition sequence (17) *and $t_1 = 0 \le \ldots \le t_M = \mu$.*

Remark. A discrete moment basis $\alpha^1, \dots, \alpha^M$ of $I\!R^M$ can be chosen to be orthonormal, when starting with α_M, spanning the spaces $MC_\mu \subseteq MC_{\mu-1} \subseteq \ldots \subseteq M_0 = R^M$ one after the other. But there are other normalizations that make sense, in particular the one that uses conjugation via A_ℓ on $MC_\ell \setminus MC_{\ell+1}$, because this matrix is positive definite there due to Micchelli's lemma. There is a hidden theoretical and numerical connection of discrete moment bases to properly pivoted LU factorizations of matrices as in (14) of values of polynomials (see also the papers [2, 3, 4], of de Boor and Ron), but we shall neither go into details nor require the reader to figure this out before we proceed.

We now consider the polynomials

$$v_j(x) := \sum_{i=1}^M \alpha_i^j \|x - x_i\|^{2t_j}, \ 1 \le j \le M \tag{19}$$

that are of degree at most $t_j \le \mu$ due to Lemma 9 and the definition of the discrete moment basis. They are low–degree linear combinations of radial polynomials, and their definition depends crucially on the geometry of X.

Lemma 13 *The $M \times M$ matrix with entries $v_j(x_k)$ is nonsingular.*

Proof: We multiply this matrix with the nonsingular $M \times M$ matrix containing the discrete moment basis $\alpha^1, \dots, \alpha^M$ and get a matrix with entries

$$\gamma_{jm} = \sum_{i=1}^M \sum_{k=1}^M \alpha_i^j \alpha_k^m \|x_i - x_k\|^{2t_j}$$

for $1 \le j, m \le M$. Consider $m > j$ and use $t_m \ge t_j$ to see that $\gamma_{jm} = 0$ as soon as $t_m > t_j$, because the entries can be written as values of a polynomial of degree $2t_j - t_j - t_m < 0$. Thus the matrix is block triangular, and the diagonal blocks consist of entries $\alpha_i^j \alpha_k^m \|x_i - x_k\|^{2t}$ with $t = t_j = t_m$. But these symmetric submatrices must be definite due to our construction of a discrete moment basis. We even could have chosen the basis such that the diagonal blocks are unit matrices, if we had used conjugation with respect to A_t. □

Theorem 7 *A least polynomial interpolation on X is possible using the functions v_j of Lemma 13. These are of degree at most $\mu = \mu(X)$, and thus*

$$\binom{\mu(X) + d - 1}{d} < M \le \binom{\mu(X) + d}{d} \le \binom{k_0(X) + d}{d}.$$

□

Example 3 Let us look at the special case with $M = 4, \ d = 2$ and points

$$x_1 := (0,0)^T, \ x_2 := (1,0)^T, \ x_3 := (0,1)^T, \ x_4 := (1/2,1)^T.$$

The discrete moment conditions on vectors $\alpha \in I\!R^4$ are

$$\sum_{j=1}^4 \alpha_j = 0 \text{ for all } \alpha \in MC_1,$$

$$\sum_{j=1}^4 \alpha_j = 0, \ \alpha_2 + \alpha_4/2 = 0, \ \alpha_3 + \alpha_4 = 0 \text{ for all } \alpha \in MC_2.$$

Furthermore we find $MC_3 = \{0\}, MC_2 = \text{ span } \{\alpha^4 := (1, -1, -2, 2)^T\}, \ MC_0 = I\!R^4$ and $MC_1 \setminus MC_2 = \text{ span } \{\alpha^2 := (1, -1, 0, 0)^T, \ \alpha^3 := (1, 0, -1, 0)^T\}$, such that a discrete moment basis of $I\!R^4$ can be formed by $\alpha^1 := (1, 0, 0, 0)^T$ with $\alpha^2, \ \alpha^3$, and α^4 together with $t_0 = 0 < t_1 = t_2 = 1 < t_3 = 2 = \mu$. From Theorem 6 we conclude that $k_0(X) = 1 \cdot 2 + 2 \cdot 1 = 4$. MAPLE confirms this, and the Lagrange basis of the form (13) comes out to be quadratic for all f that one could start with, but the result depends on f. For example, the Lagrange basis function for the origin is

$$\begin{array}{rll} 1 - \frac{2}{9}x^2 + \frac{8}{9}xy - \frac{7}{9}x - y & \text{for } f(t) = e^{-t} & \phi = \text{ Gaussian} \\ \frac{1}{37}\left(37 + 32xy - 35y - 10x^2 - 2y^2 - 27x\right) & \text{for } f(t) = 1/(1+t) & \phi = \text{ inverse multiquadric.} \end{array}$$

The Gaussian case coincides with the de Boor/Ron solution from section 6 of [3]. The method based on (19) yields the basis function

$$\frac{1}{19}\left(19-13x-17y-6x^2-2y^2+16xy\right).$$

Thus we have different methods, but we note that the de Boor/Ron interpolation method coincides with the limit of interpolation with shifted and scaled Gaussians. We shall prove this in the penultimate section.

6 Polynomial Reproduction

By (4), the Lagrange basis functions, if they exist, have analytic expansions

$$u_j^c(x)=\sum_{m=0}^{\infty} b_{jm}(x)c^m$$

for small c with certain multivariate polynomials b_{jm} depending on X and ϕ. We put this and the expansion (12) into the defining equations (2) to get

$$\begin{aligned}
\sum_{j=1}^{M}\sum_{m=0}^{\infty} b_{jm}(x)c^m \sum_{\ell=0}^{\infty} f_\ell c^\ell \|x_j-x_k\|_2^{2\ell} &= \sum_{s=0}^{\infty} f_s c^s \|x-x_k\|_2^{2s},\ 1\le k\le M\\
\sum_{s=0}^{\infty} c^s \sum_{\ell=0}^{s} f_\ell \sum_{j=1}^{M} b_{j,s-\ell}(x)\|x_j-x_k\|_2^{2\ell} &= \sum_{s=0}^{\infty} f_s c^s \|x-x_k\|_2^{2s},\ 1\le k\le M
\end{aligned}$$

and by comparison of coefficients we arrive at

$$\sum_{\ell=0}^{s} f_\ell \sum_{j=1}^{M} b_{j,s-\ell}(x)\|x_j-x_k\|_2^{2\ell} = f_s\|x-x_k\|_2^{2s},\ 1\le k\le M,\ s\ge 0. \tag{20}$$

In view of Lemma 12 it seems to be the upshot of (20) that the reproduction of all vectors $F_s(x)$ of radial polynomials (where for this section we introduce the factors f_ℓ instead of $(-1)^\ell$ into the A_ℓ matrices) is possible from the data of all $F_\ell(x_j),\ 1\le j\le M,\ 0\le \ell\le s$. This works if $s\ge\mu$ by that Lemma, but the reconstruction is not unique, and equations (20) describe a special selection of a reconstruction. It is strange that the reconstruction via the above equations always works if convergence of the radial basis function interpolation takes place, also for small $s<\mu$, and that it can be chosen to have a convolution–type structure.

We are interested in the special Lagrange basis polynomials

$$u_j^0(x)=b_{j0}(x),\ 1\le j\le M \text{ with } b_{j0}(x_k)=\delta_{jk},\ 1\le j,k\le M$$

that arise in (20). We want to derive properties and defining equations. To this end, define the set

$$D_{X,\phi} := \left\{(\alpha,s)\in I\!R^M\times I\!N_0\ :\ f_\ell\sum_{k=1}^{M}\alpha_k\|x_j-x_k\|^{2\ell}=0 \text{ for all } 0\le\ell<s,\ 1\le j\le M\right\}$$

and polynomials

$$p_{\alpha,s}(x) := f_s\sum_{k=1}^{M}\alpha_k\|x-x_k\|^{2s} \text{ for all } (\alpha,s)\in D_{X,\phi}. \tag{21}$$

Applying such a linear combination to (20) with respect to the points x_k yields

$$\sum_{j=1}^{M} b_{j,0}(x)p_{\alpha,s}(x_j)=p_{\alpha,s}(x),\ (\alpha,s)\in D_{X,\phi}. \tag{22}$$

Theorem 8 *The limiting polynomial interpolation process reproduces all polynomials of the form* (21). □

Lemma 14 *If $f_0\neq 0$, the limiting polynomial interpolation process reproduces constants.*

Proof: Consider elements $(\alpha,0)\in D_{X,\phi}$ in (21). □

Lemma 15 *If $f_0,f_1\neq 0$, the limiting polynomial interpolation process reproduces all linear functions in the linear span of the data.*

Proof: Consider elements $(\alpha, 1) \in D_{X,\phi}$ in (21). The restriction is $\sum_k \alpha_k = 0$, and due to Lemma 14 we have reproduction of constants. Thus we can reproduce any polynomial of the form

$$p(x) = \|x - x_k\|_2^2 - \|x - x_\ell\|_2^2 - \|x_k\|_2^2 + \|x_\ell\|_2^2 = x^T(x_\ell - x_k)$$

for $1 \leq k < \ell \leq M$. $\square$

Let us generalize this. For all $s \geq 0$, define the spaces

$$D_s := \left\{\alpha \in I\!R^M \ : \ (\alpha, s) \in D_{X,\phi}\right\}, \quad R_s := \text{ span } \{p_{\alpha,s} \ : \ \alpha \in D_s\} \tag{23}$$

and the maps $T_s \ : D_s \to R_s$ with

$$T_s(\alpha) := p_{\alpha,s} \text{ for all } \alpha \in D_s.$$

We have $T_s(D_s) = R_s$ by definition, and we find that $\ker T_s = D_{s+1}$. Note further that $D_0 = I\!R^M$. This yields a decomposition sequence of $I\!R^M$ into a sequence of spaces isomorphic to $R_0, R_1, \ldots$ and we have to see whether this sequence exhausts all of $I\!R^M$. The decomposition can only fail if $D_s = D_t$ for some s and all $t \geq s$, while D_s still is nonzero. But this cannot happen, because then there is some nonzero $\alpha \in I\!R^M$ which is in all D_s for all $s \geq 0$. But then

$$0 = \sum_{j,k=1}^{M} \alpha_j \alpha_k \phi(\|x_j - x_k\|_2)$$

and this cannot hold for positive definite functions ϕ. Thus we know that there is a minimal finite decomposition

$$D_{\sigma+1} = \{0\} \neq D_\sigma \subseteq \ldots \subseteq D_1 \subseteq D_0 = R^M \tag{24}$$

with factor spaces $R_s = T_s(D_s) = D_s/D_{s+1}$ for $0 \leq s \leq \sigma$. Unfortunately, the spaces $R_0, R_1, \ldots$ turn out to be overlapping, and we only know that

$$R := \sum_{s=0}^{\sigma} R_s$$

is a polynomial space of dimension at most M that is reproduced by our interpolation.

The relation to discrete moments is that the D_ℓ spaces are the intersection of kernels of matrices $A_0, \ldots, A_{\ell-1}$. This is a different type of discrete moment condition, and in general $\sigma \leq \mu$. See Example 4 below.

If (22) were sufficient to determine the functions b_{j0} completely, one could cancel the factors f_s for positive definite functions and get a construction technique that is even independent of f and ϕ. But by looking at cases like in Example 3 for positive definite functions one can see that the results actually depend on f and ϕ. Thus one has to go back to (20). Here, one could write the equations for $1 \leq k \leq M$ and $0 \leq s \leq S$ for some large positive $S \geq \mu(X)$, and then by Lemma 12 one has a solvable system of $M(S+1)$ equations for $M(S+1)$ unknowns. However, it turns out that there is some serious rank loss, while for large S the particular functions b_{j0} come out uniquely when running a symbolic MAPLE program, leaving many $b_{j\ell}$ for larger ℓ undetermined. An example follows below. It is a challenging open problem to find a good technique to determine the $b_{j\ell}$ for all $1 \leq j \leq M$ and small $\ell \geq 0$ in a finite and stable way.

We finally want to collect more information on the polynomials that are reproduced. Define, by splitting (20), the polynomials

$$z_{k,s}(x) := f_s\|x - x_k\|_2^{2s} - \sum_{\ell=0}^{s-1} f_\ell \sum_{j=1}^{M} b_{j,s-\ell}(x)\|x_j - x_k\|_2^{2\ell} = f_s \sum_{j=1}^{M} b_{j,0}(x)\|x_j - x_k\|_2^{2s} \tag{25}$$

for all $1 \leq k \leq M,\ s \geq 0$. Then $z_{k,s}(x_j) = f_s\|x_j - x_k\|_2^{2s},\ 1 \leq j,k \leq M$ and (20) implies

$$\begin{aligned}
& \sum_{j=1}^{M} b_{j,0}(x) z_{k,s}(x_j) \\
= \quad & \sum_{j=1}^{M} b_{j,0}(x) f_s \|x_j - x_k\|_2^{2s} \\
= \quad & f_s\|x - x_k\|_2^{2s} - \sum_{\ell=0}^{s-1} f_\ell \sum_{j=1}^{M} b_{j,s-\ell}(x)\|x_j - x_k\|_2^{2\ell} \\
= \quad & z_{k,s}(x),
\end{aligned}$$

proving that all the $z_{k,s}(x)$ are reproduced. There are infinitely many of them, but they span a space of dimension at most M, and thus their degree cannot increase indefinitely with s. So far, all examples indicate that the limiting polynomial interpolants have degrees not exceeding the maximal discrete moment order μ, but a proof is still missing.

Example 4 Let us continue our previous case in Example 3 of 4 points in 2D in view of the above terminology. Looking at σ instead of μ, we find that $D_2 = \{0\},\ D_1 = MC_1,\ D_0 = I\!R^4 = MC_0$, and thus $\mu = 2 > \sigma = 1$. We now look at (20) and (22) to determine the polynomials $b_{j0}(x),\ 1 \leq j \leq 4$. Here, we assume all f_ℓ to be nonzero. The system (22) leads to no more than 3 equations, and we can use $(\alpha^1, 0),\ (\alpha^2, 1),\ (\alpha^3, 1) \in D_{X,\phi}$ as pairs (α, s) there. The result consists of the three equations that describe reproduction of linear polynomials:

$$\begin{array}{ccccccccc} b_{10}(x,y) & + & b_{20}(x,y) & + & b_{30}(x,y) & + & b_{40}(x,y) & = & 1 \\ & & b_{20}(x,y) & & & + & b_{40}(x,y)/2 & = & x \\ & & & & b_{30}(x,y) & + & b_{40}(x,y) & = & y. \end{array} \tag{26}$$

We thus turn to (20) in general, and discuss the cases $s = 0, 1, 2, \ldots$ one after the other. To start with, the case $s = 0$ yields the first equation of (26). Assuming that this condition is already satisfied, we consider $s = 1$ of (20) for $\alpha \in MC_1$ and find the two other equations of (26). Taking these for granted also, we consider a general α there and get a new equation

$$f_1\|(x,y)\|_2^2 = f_1 \sum_{j=1}^{4} b_{j0}(x,y)\|x_j\|_2^2 + f_0 \sum_{j=1}^{4} b_{j1}(x,y)$$

that contains the new polynomial $b_1(x,y) := \sum_{j=1}^4 b_{j1}(x,y)$ and has the explicit form

$$f_1(x^2+y^2) = f_1\left(b_{20}(x,y) + b_{30}(x,y) + \frac{5}{4}b_{40}(x,y)\right) + f_0 b_1(x,y).$$

We now have a new equation, but also a new variable. Thus the cases $s = 0$ and $s = 1$ do not yet lead to a unique determination of the b_{j0}. We have to turn to $s = 2$ in (20), but we do not want to introduce more variables than absolutely necessary. We find

$$\begin{array}{rcl} f_2 \displaystyle\sum_{k=1}^{4} \alpha_k \|(x,y) - x_k\|_2^4 & = & f_2 \displaystyle\sum_{j=1}^{4} b_{j0}(x,y) \sum_{k=1}^{4} \alpha_k \|x_j - x_k\|_2^4 \\ & + & f_1 \displaystyle\sum_{j=1}^{4} b_{j1}(x,y) \sum_{k=1}^{4} \alpha_k \|x_j - x_k\|_2^2 \\ & + & f_0 \displaystyle\sum_{j=1}^{4} b_{j2}(x,y) \sum_{k=1}^{4} \alpha_k \end{array}$$

and we would like to find a nonzero vector $\alpha \in I\!R^4$ such that

$$\sum_{k=1}^{4} \alpha_k = 0,\ \sum_{k=1}^{4} \alpha_k \|x_j - x_k\|_2^2 = c \text{ for all } j,\ 1 \leq j \leq 4$$

with a constant c, because then we would have no new variable. Fortunately, the vector $\alpha^4 = (1, -1, -2, 2)^T$ spanning MC_2 does the job with $c = -1/2$, and with our new equation

$$\begin{array}{rcl} f_2 \displaystyle\sum_{k=1}^{4} \alpha_k^4 \|(x,y) - x_k\|_2^4 & = & f_2 \displaystyle\sum_{j=1}^{4} b_{j0}(x,y) \sum_{k=1}^{4} \alpha_k^4 \|x_j - x_k\|_2^4 \\ & - & f_1 \dfrac{1}{2} b_1(x,y) \end{array}$$

we now have five equations for five variables. This is the point where the solution of (20) leads to unique determination of the b_{j0} and the sum over the b_{j1}. Note that we are still far from determining all the $b_{j\ell}$, but we are interested in the b_{j0} only and succeeded to get away with small s in (20). This principle worked in all cases that we tested with MAPLE, and it directly yielded the Lagrange polynomial bases via symbolic computation.

7 The method of de Boor and Ron revisited

The goal of this section is to prove Theorem 2. For this we require at least a scaled version of the de Boor/Ron technique, and we take the opportunity to rephrase with slightly increased generality.

For all $\alpha \in \mathbb{Z}_0^d$ let w_α be a positive real number, and consider the inner product

$$(p,q)_w := \sum_{\alpha\in\mathbb{Z}_0^d} \frac{1}{w_\alpha}(D^\alpha p)(0)(D^\alpha q)(0)$$

on the space $\mathbb{P}_\infty^d$ of all d–variate polynomials. The de Boor/Ron interpolant arises in the special case $w_\alpha = \alpha!$. Now we want to link the theory to radial basis function techniques. If we assume

$$\sum_{\alpha\in\mathbb{Z}_0^d} \frac{w_\alpha}{\alpha!^2} < \infty$$

and define the kernel

$$K_w(x,y) := \sum_{\alpha\in\mathbb{Z}_0^d} w_\alpha \frac{x^\alpha}{\alpha!}\frac{y^\alpha}{\alpha!}, \tag{27}$$

all polynomials $p \in \mathbb{P}_\infty^d$ are reproduced on $[-1,1]^d$ via

$$p(x) = (p, K_w(x,\cdot))_w \tag{28}$$

and this identity carries over to the Hilbert space completion

$$\mathcal{H}_w := \left\{ g \in C^\infty(\mathbb{R}^d) \;:\; g(x) = \sum_{\alpha\in\mathbb{Z}_0^d} (D^\alpha g)(0)\frac{x^\alpha}{\alpha!},\; \sum_{\alpha\in\mathbb{Z}_0^d} \frac{1}{w_\alpha}(D^\alpha g)^2(0) < \infty \right\}$$

of the polynomials under the above inner product. The kernel K_w is positive definite on $[-1,1]^d$, and larger domains can be treated by scaling. Since polynomials separate points, it is clear that for all finite sets $X = \{x_1,\ldots,x_M\} \subset [-1,1]^d$ we have linear independence of the functions $K_w(\cdot,x_j)$, and interpolation in X by the span of these functions is uniquely possible.

So far, we used standard arguments of radial basis function theory. In the papers of de Boor and Ron, transition to a polynomial interpolation process is done via truncation, not via passing to the limit of a scaling. For all functions g from $\mathcal{H}_w$ the notation

$$g^{[k]}(x) := \sum_{\substack{\alpha\in\mathbb{Z}_0^d\\ |\alpha|=k}} (D^\alpha g)(0)\frac{x^\alpha}{\alpha!}$$

is introduced, while $g\!\downarrow$ stands for the nonzero function $g^{[k]}$ with minimal k. For a finite set $X = \{x_1,\ldots,x_M\} \subset [-1,1]^d$ the spaces

$$E_{w,X} := \text{ span }\{K_w(x,\cdot) \;:\; x \in X\},\; P_{w,X} := \text{ span }\{g\!\downarrow \;:\; g \in E_{w,X}\}$$

are introduced, and $P_{w,X}$ is a space of polynomials.

Theorem 9 *Interpolation on X by functions in $P_{w,X}$ is uniquely possible.*

Proof: Assume that there is some nonzero $p \in P_{w,X}$ such that $p = g_p\!\downarrow$ and $p_{|X} = 0$. Then (28) yields orthogonality $(p,g)_w = 0$ for all $g \in E_{w,X}$, and we have the contradiction

$$0 = (p,g_p)_w = (g_p\!\downarrow, g_p)_w = (g_p\!\downarrow, g_p\!\downarrow)_w.$$

So far, we have followed the proof of de Boor and Ron, but now we want to use a discrete moment basis $\alpha^1,\ldots,\alpha^M$ to link the process with what we have done in previous sections. We define functions

$$v_r(y) := \sum_{j=1}^M \alpha_j^r K_w(x_j,y) = \sum_{|\alpha|\geq t_r} w_\alpha \frac{y^\alpha}{\alpha!}\sum_{j=1}^M \alpha_j^r \frac{x_j^\alpha}{\alpha!},\; g_r := v_r\!\downarrow,\; 1 \leq r \leq M. \tag{29}$$

Due to the property of a discrete moment basis we see that not all of the quantities

$$c_{\alpha,r} := \sum_{j=1}^{M} \alpha_j^r \frac{x_j^\alpha}{\alpha!}$$

for $|\alpha| = t_r$ can vanish, because otherwise $\alpha^r \in MC_{t_r+1}$. Thus we have the homogeneous representation

$$g_r(y) = v_r\!\downarrow\, = \sum_{|\alpha|=t_r} w_\alpha c_{\alpha,r} \frac{y^\alpha}{\alpha!} \tag{30}$$

and $(g_r, g_s)_w = 0$ for $t_r \neq t_s$. The matrix formed by the $(g_r, g_s)_w$ is a positive semidefinite block–diagonal Gramian. To prove its definiteness, we can focus on a single diagonal block with $t = t_r = t_s$. Collecting the indices r with $t_r = t$ into a set I_t, we assert linear independence of the functions g_r for $r \in I_t$. For a vanishing linear combination

$$\begin{aligned}
0 &= \sum_{r\in I_t} \gamma_r g_r(y) \\
&= \sum_{r\in I_t} \gamma_r \sum_{|\alpha|=t} w_\alpha c_{\alpha,r} \frac{y^\alpha}{\alpha!} \\
&= \sum_{|\alpha|=t} \frac{y^\alpha}{\alpha!} w_\alpha \sum_{r\in I_t} \gamma_r c_{\alpha,r} \\
&= \sum_{|\alpha|=t} \frac{y^\alpha}{\alpha!} w_\alpha \sum_{r\in I_t} \gamma_r \sum_{j=1}^{M} \alpha_j^r \frac{x_j^\alpha}{\alpha!} \\
&= \sum_{|\alpha|=t} \frac{y^\alpha}{\alpha!} w_\alpha \sum_{j=1}^{M} \left(\sum_{r\in I_t} \gamma_r \alpha_j^r \right) \frac{x_j^\alpha}{\alpha!}
\end{aligned}$$

we conclude that $\sum_{r\in I_t} \gamma_r \alpha^r$ is a vector in MC_{t+1}, and this can hold only if the coefficients are zero. Thus the space $P_{w,X}$ contains the M linearly independent homogeneous polynomials $g_1, \ldots, g_M$ of increasing degrees $0 = t_1 \leq \cdots \leq t_M = \mu$, and the theorem is proven. Due to Theorem 5, the degree is "least", as known from the de Boor/Ron papers. □

We now proceed towards proving that the limit of interpolants by Gaussians is equal to the de Boor/Ron polynomial interpolant. We need something that links kernels of the form (27) to radial kernels.

Lemma 16 *If ϕ is a positive definite analytic radial basis function that can be written via an analytic function f satisfying* (1) *and* (12), *then*

$$f(x^T y) = \sum_{\alpha\in\mathbb{Z}_0^d} \frac{f^{|\alpha|}(0)}{\alpha!} x^\alpha y^\alpha = \sum_{\alpha\in\mathbb{Z}_0^d} f_{|\alpha|} x^\alpha y^\alpha \tag{31}$$

for all $x, y \in \mathbb{R}^d$.

Proof: We use the Bernstein–Widder representation

$$f(r) = \int_0^\infty e^{-rt} d\mu(t),\ r \geq 0$$

to get

$$(-1)^j f^{(j)}(0) = (-1)^j j! f_j = \int_0^\infty t^j d\mu(t) \in (0,\infty) \text{ for all } j \geq 0$$

and similarly, factoring the exponential in the integral,

$$D_x^\alpha f(x^T y) = y^\alpha \int_0^\infty (-1)^{|\alpha|} t^{|\alpha|} e^{-tx^T y} d\mu(t)$$

and

$$D_x^\alpha f(x^T y)_{|x=0} = y^\alpha f^{|\alpha|}(0) = y^\alpha \alpha! f_{|\alpha|}$$

where D_x^α takes derivatives of order α with respect to x. The assertion follows from putting the result into the power series expansion at zero. □

At first sight, the above result is disappointing, because one cannot easily use (31) in (27), since the coefficients in (31) are alternating. However, a closer look reveals that the major part of the de Boor/Ron theory does not rely on the signs of the coefficients. It is the link to radial basis functions as reproducing kernels that does not work without further arguments. This has a positive consequence: the generalized de Boor/Ron approach as given at the start of the section will yield many new cases of positive definite non–radial interpolants with polynomial truncations that furnish least polynomial interpolants. On the downside, we cannot expect to find a direct link between interpolation by general positive definite radial basis functions and the generalized de Boor/Ron method.

But for Gaussians, we can add some more work, factoring

$$\exp(-c\|x_j - x_k\|_2^2) = \exp(-c\|x_j|_2^2) \cdot \exp(2cx_j^T x_k) \cdot \exp(-c\|x_k\|_2^2).$$

We then rewrite the Lagrange system

$$\sum_{j=1}^{M} u_j^c(x) \exp(-c\|x_j - x_k\|_2^2) = \exp(-c\|x - x_k\|_2^2),\ 1 \le k \le M$$

in the form

$$\sum_{j=1}^{M} z_j^c(x) \exp(2cx_j^T x_k) = \sum_{j=1}^{M} u_j^c(x) \exp(-c(\|x_j\|^2 - \|x\|^2)) \exp(2cx_j^T x_k) = \exp(2cx^T x_k),\ 1 \le k \le M$$

with $z_j^c(x) := u_j^c(x) \exp(-c(\|x_j\|^2 - \|x\|^2)),\ 1 \le j \le M$. Now we can use the technique of the previous lemma directly, putting in the expansions for the exponential and working with the kernel $K_c(x, y) := \exp(2cx^t y)$ in our presentation of the de Boor/Ron technique, with a slight abuse of notation. This implies that the functions $z_j^c(x)$ are Lagrangian interpolants in the span of the $K_c(x_j, x) = \exp(2cx_j^t x)$. Any interpolation by scaled Gaussians can be converted by the above transformation to and from an interpolation using the kernel K_c.

We now look at what happens if the de Boor/Ron truncation process is carried out on interpolants defined via K_c. The functions in (30) come out as

$$\begin{aligned} g_r^c(y) &= \sum_{|\alpha|=t_r} w_\alpha^c c_{\alpha,r} \frac{y^\alpha}{\alpha!} \\ &= \sum_{|\alpha|=t_r} \frac{(2c)^{|\alpha|}}{\alpha!} c_{\alpha,r} \frac{y^\alpha}{\alpha!} \\ &= (2c)^{t_r} \sum_{|\alpha|=t_r} \frac{1}{\alpha!} c_{\alpha,r} \frac{y^\alpha}{\alpha!} \\ &= (2c)^{t_r} g_r^{dBR}(y) \end{aligned}$$

i.e. they are just scalar multiples of the functions g_r^{dBR} of the de Boor/Ron process. Thus the polynomial space spanned by truncation of the K_c is independent of c and coincides with the de Boor/Ron polynomial interpolation space.

We have successfully moved from interpolation by Gaussian radial basis functions to interpolation by scaled exponentials, and we have seen that the truncation of the latter is the de Boor/Ron polynomial space. But we now have to investigate the limit of the interpolants spanned by the scaled exponentials $K_c(x_j, \cdot)$ for $c \to 0$ to see whether they converge towards the de Boor/Ron truncation.

We go back to (29) to define functions v_r^c as

$$\begin{aligned} v_r^c(y) &:= \sum_{j=1}^{M} \alpha_j^r K_c(x_j, y) \\ &= \sum_{|\alpha|\ge t_r} \frac{(2c)^{t_r}}{\alpha!} \frac{y^\alpha}{\alpha!} c_{\alpha,r} \\ &= \sum_{s\ge t_r} (2c)^s \sum_{|\alpha|=s} \frac{1}{\alpha!} \frac{y^\alpha}{\alpha!} c_{\alpha,r} \end{aligned}$$

such that

$$\lim_{c\to 0} \frac{v_r^c(y)}{(2c)^{t_r}} = g_r^{dBR}(y),\ 1 \le r \le M.$$

This means that the space spanned by the $K_c(x_j, \cdot)$ contains a basis that converges towards a basis of the de Boor/Ron polynomial space for $c \to \infty$. If the Lagrange basis for interpolation in the span of the $K_c(x_j, \cdot)$ is written in terms of this basis, it converges towards a polynomial limit. This ends the proof of Theorem 2. □

8 Preconditioning

The transition from the Gaussian system to the scaled basis $\frac{v_r^c(y)}{(2c)^{t_r}}$ should be useful as a preconditioning technique. In general, we show in this section how to use a discrete moment basis for preconditioning badly conditioned matrices arising from interpolation by smooth radial basis functions.

We go back to the beginning of the paper and precondition the matrix $A_{c,X,X}$ arising in (2) by a scaled discrete moment basis

$$c^{-t_1/2}\alpha^1, \ldots, c^{-t_M/2}\alpha^M \tag{32}$$

in the following way. If we put the discrete moment basis into an $M \times M$ matrix B_c and form the positive definite symmetric matrix $Z(c) := B_c A_{c,X,X} B_c^T$, the matrix entries will be

$$\begin{aligned}
z_{rs}(c) &:= \sum_{j,k=1}^{M} \alpha_j^r c^{-t_r/2} \alpha_k^s c^{-t_s/2} \phi_c(\|x_j - x_k\|_2) \qquad\qquad 1 \le r, s \le M\\
&= \sum_{n=0}^{\infty} f_n c^{n-t_r/2-t_s/2} \sum_{j,k=1}^{M} \alpha_j^r \alpha_k^s \|x_j - x_k\|_2^{2n}\\
&= \sum_{\substack{n=0\\ 2n \ge t_r+t_s}}^{\infty} f_n c^{n-t_r/2-t_s/2} \sum_{j,k=1}^{M} \alpha_j^r \alpha_k^s \|x_j - x_k\|_2^{2n}
\end{aligned}$$

with well–defined limits

$$z_{rs}(0) = \left\{ \begin{array}{cc} f_{(t_r+t_s)/2} \sum_{j,k=1}^{M} \alpha_j^r \alpha_k^s \|x_j - x_k\|_2^{t_r+t_s} & t_r + t_s \text{ even} \\ 0 & \text{else} \end{array} \right\}$$

for $c \to 0$. The matrix $Z(0)$ is positive semidefinite by construction, and we assert

Theorem 10 *The matrix $Z(0)$ is positive definite.*

Proof: We use the proof technique of Theorem 6. The product of all eigenvalues of $A_{c,X,X}$ decays with exponent $k_0(X)$ as in (18), while the maximum eigenvalue stays bounded above independent of c. But our matrix transformation performs a multiplication of the spectral range by $c^{-k_0(X)}$, because $k_0(X) = \sum_{j=1}^{M} t_j$ is just another way to write (18). Thus the smallest eigenvalue of the product must stay away from zero when $c \to 0$. But since the matrix $Z(0)$ is well–defined, the maximal eigenvalue of the product $Z(c)$ must be bounded, and $Z(0)$ altogether has a strictly positive spectrum. $\square$

Example 5 If we go back to the four points in $I\!R^2$ of Examples 3 and 4 and scale the discrete moment basis as in (32) via

$$B_c := \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{\sqrt{c}} & -\frac{1}{\sqrt{c}} & 0 & 0 \\ \frac{1}{\sqrt{c}} & 0 & -\frac{1}{\sqrt{c}} & 0 \\ \frac{1}{c} & -\frac{1}{c} & -\frac{2}{c} & \frac{2}{c} \end{pmatrix},$$

MAPLE produces a limit matrix

$$Z(0) := \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{19}{4} \end{pmatrix}.$$

for Gaussians and

$$Z(0) := \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{4} \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -\frac{1}{4} & 0 & 0 & -\frac{19}{16} \end{pmatrix}$$

for (negative definite) inverse multiquadrics with $\beta = -1$. If we take four equidistant points on the line $[0, 1] \times \{0\}$, we find

$$Z(0) := \begin{pmatrix} 1 & 0 & -\frac{2}{9} & 0 \\ 0 & \frac{2}{9} & 0 & -\frac{4}{27} \\ -\frac{2}{9} & 0 & \frac{4}{27} & 0 \\ 0 & -\frac{4}{27} & 0 & \frac{40}{243} \end{pmatrix} \text{ for } B_c := \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{\sqrt{c}} & -\frac{1}{\sqrt{c}} & 0 & 0 \\ \frac{1}{c} & -\frac{2}{c} & \frac{1}{c} & 0 \\ \frac{1}{c\sqrt{c}} & -\frac{3}{c\sqrt{c}} & \frac{3}{c\sqrt{c}} & -\frac{1}{c\sqrt{c}} \end{pmatrix}.$$

in case of Gaussians. The discrete moment basis now contains divided differences, and the zero structure is different from the previous case, because we have $t_j = j - 1,\ 1 \le j \le 4$ here.

Acknowledgements: There were some very helpful e–mail exchanges with Bengt Fornberg.

References

[1] C. de Boor. A multivariate divided difference. In C. K. Chui and L. L. Schumaker, editors, *Approximation Theory VIII, Vol. 1: Approximation and Interpolation*, pages 87–96. World Scientific Publishing Co, 1995.

[2] C. de Boor and A. Ron. On multivariate polynomial interpolation. *Constructive Approximation*, 6:287–302, 1990.

[3] C. de Boor and A. Ron. Computational aspects of polynomial interpolation in several variables. *Mathematics of Computation*, 8:705–727, 1992.

[4] C. de Boor and A. Ron. The least solution for the polynomial interpolation problem. *Mathematische Zeitschrift*, 210:347–378, 1992.

[5] P. Danzeglocke. Personal communication. January 2002.

[6] T.A. Driscoll and B. Fornberg. Interpolation in the limit of increasingly flat radial basis functions. *Comput. Math. Appl.*, to appear, 2002.

[7] B. Fornberg, G. Wright, and E. Larsson. Some observations regarding interpolants in the limit of flat radial basis functions. Preprint Univ. of Colorado, Boulder, 2002.

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

[9] R. Schaback. Approximation by radial basis functions with finitely many centers. *Constructive Approximation*, 12:331–340, 1996.

[10] R. Schaback. Reconstruction of multivariate functions from scattered data. Manuscript, available via http://www.num.math.uni-goettingen.de/schaback/research/group.html, 1997.

[11] R. Schaback and H. Wendland. Characterization and construction of radial basis functions. In N. Dyn, D. Leviatan, D. Levin, and A. Pinkus, editors, *Multivariate Approximation and Applications*, pages 1–24. Cambridge University Press, 2001.

Author's address:
Prof. Dr. R. Schaback
Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestraße 16–18
D–37-83 Göttingen
e-mail: `schaback@math.uni-goettingen.de`
`http://www.num.math.uni-goettingen.de/schaback`