Limit Problems for Interpolation by Analytic Radial Basis Functions

Robert Schaback [a]

[a] *Institut für Numerische und Angewandte Mathematik*
Lotzestraße 16–18
D–37083 Göttingen
Germany

Abstract

Interpolation by analytic radial basis functions like the Gaussian and inverse multiquadrics can degenerate in two ways: the radial basis functions can be scaled to become "increasingly flat", or the data points "coalesce" in the limit while the radial basis functions stays fixed. Both cases call for a careful regularization. If carried out explicitly, this yields a preconditioning technique for the degenerating linear systems behind such interpolation problems. This paper deals with both degeneration cases. For the "increasingly flat" limit, we recover results by Larsson and Fornberg together with Lee, Yoon, and Yoon concerning convergence of interpolants towards polynomials. With slight modifications, the same technique also allows to handle scenarios with coalescing data points for fixed radial basis functions. The results show that the degenerating local Lagrange interpolation problems converge towards certain Hermite-Birkhoff problems. This is an important prerequisite for dealing with approximation by radial basis functions adaptively, using freely varying data sites.

Key words: radial basis functions, moment conditions, preconditioning

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

1 Introduction

Since our analysis will have close connections to multivariate polynomial interpolation, we shall start with the latter. Then we turn to multivariate meshless kernel-based interpolation problems and focus on the "increasingly flat kernel" case, because it partially solves the "coalescing points" case also, as will turn out in section 8. Following (1), we shall explicitly precondition the degenerating interpolation problem in such a way that the limit of the preconditioned systems can be analyzed and calculated. For the "increasingly flat" case, this recovers results by (2; 3) concerning sufficient criteria for convergence of the interpolants to polynomials. Then we focus on the case of coalescing data to be interpolated by a fixed radial basis function. Though the limit will usually not be a polynomial, our technique applies with certain modifications, and it proves that the Lagrange problems on coalescing data points always converge towards certain Hermite interpolation problems.

2 Polynomial Interpolation

For multivariate polynomial interpolation on a set $X := \{x_1, \ldots, x_N\}$ in $I\!R^d$ there are a few important quantities to be defined a-priori. To this end, we use multi-indices $\alpha \in Z\!\!Z_0^d$ in the standard way, defining the monomials $x^\alpha \in I\!R^d$ for $x \in I\!R^d$ and the nonnegative integer $|\alpha| := \|\alpha\|_1$ as usual. A polynomial

$$p(x) := \sum_{\alpha \in Z\!\!Z_0^d} a_\alpha x^\alpha$$

with finitely many nonzero coefficients and degree ∂p is an interpolant to data $y_1, \ldots, y_N$ on X, if $p(x_k) = y_k,\ 1 \leq k \leq N$. Furthermore, the space $I\!P_k^d$ of polynomials of degree at most k in d variables has dimension $\binom{k+d}{d}$. Only in rare cases will the number N of given data be equal to one of these numbers. Anyway, there always is an integer $k_1 = k_1(X)$ with

$$\binom{k_1 - 1 + d}{d} < N \leq \binom{k_1 + d}{d} \tag{1}$$

which is a rough guess for the expectable degree of an interpolating polynomial, if the N data are well-situated in $I\!R^d$. However, even in case $N = \binom{k_1+d}{d}$ it is not at all clear whether the monomial basis $\{x^\alpha \ :\ |\alpha| \leq k_1\}$ is linearly independent on X. Therefore one has to look at *monomial* or *Vandermonde* matrices formed by entries x_j^α , where we let the row index be $j,\ 1 \leq j \leq N$ and the column index be the multi-index α. We order multi-indices $\alpha,\ \beta \in Z\!\!Z_0^d$ *polynomially* by defining $\alpha < \beta$ if either $|\alpha| < |\beta|$ or $|\alpha| = |\beta|$ with $\alpha < \beta$ lexicographically. This way, we can define the infinite monomial matrix

$$I\!M_{<\infty} = (x_j^\alpha)_{1 \leq j \leq N, \alpha \in Z\!\!Z_0^d}$$

with N rows, where the columns are formed by multi-indices in ascending polynomial order. Finite partial monomial matrices are, for instance,

$$\begin{aligned} I\!M_{\leq\beta} &= (x_j^\alpha)_{1\leq j\leq N, \alpha\in \mathbb{Z}_0^d, \alpha\leq\beta} \\ I\!M_{|.|\leq k} &= (x_j^\alpha)_{1\leq j\leq N, \alpha\in \mathbb{Z}_0^d, |\alpha|\leq k} \end{aligned}$$

with N rows also. Existence of an interpolant of degree k for arbitrary data on X is ensured if the monomial matrix $I\!M_{|.|\leq k}$ has rank N. Thus, a crucial number associated to polynomial interpolation on X is

$$k_2 := k_2(X) := \min\{k \ : \ rank(I\!M_{|.|\leq k}) = N\},$$

leading to existence of interpolating polynomials of degree at most k_2 for any data on X. We have $k_2 \leq N-1$, because the nonzero Lagrange-type polynomials

$$L_i(x) := \prod_{1\leq j\leq N,\, i\neq j} \frac{(x-x_j)^T(x_i-x_j)}{\|x_i-x_j\|_2^2}$$

have degree at most $N-1$ and are linearly independent on X. The matrix $I\!M_{|.|\leq k_2}$ has N linearly independent columns which cannot all occur already in $I\!M_{|.|<k_2}$. Note that the bound $k_2 \leq N-1$ is sharp in one-dimensional cases, and clearly $k_1 \leq k_2$ holds because of (1).

But uniqueness usually is more complicated and will not hold without further assumptions. Of course, one can always select N multi-indices $\alpha^1 < \ldots < \alpha^N$ with $|\alpha^i| \leq k_2$ such that the monomial matrix with entries $x_j^{\alpha^k}$, $1 \leq j,k \leq N$ is nonsingular. Then there is a unique interpolant in the span of x^{α^j}, $1\leq j\leq N$, but any other choice of multi-indices with the above property will lead to a different interpolant. The parameter describing uniqueness is

$$k_0 := k_0(X) := \max\{k \ : \ p \in I\!P_k^d,\ p(X) = \{0\} \Rightarrow p = 0\}$$

defined as the maximal k such that any polynomial from $I\!P_k^d$ vanishing on X must be identically zero. Equivalently, k_0 is the maximal polynomial degree for which interpolants, if they exist, are unique. The monomial matrix $I\!M_{|.|\leq k_0}$ must then have rank $\binom{k_0+d}{d} \leq N$, and we finally get

$$0 \leq k_0 \leq k_1 \leq k_2 \leq N-1 \tag{2}$$

as a fundamental relation between the problem parameters. Note that in case of N data on a line in $I\!R^d$ we have

$$0 = k_0 \leq k_1 \leq k_2 = N-1, \tag{3}$$

the intermediate k_1 being ridiculously dependent on the dimension d of the embedding space. This is why, in contrast to (2; 3), we consider k_1 as much less relevant for analysis than the other parameters, and ignore it from now on. Note that the classical geometric situation of data points in "general position" with respect to $I\!R^d$ is the case of maximal k_0, and this case can be described by $k_0 = k_1 = k_2$ in case $N = \binom{k_1+d}{d}$, while $k_0 = k_1 - 1 = k_2 - 1$ in case $N < \binom{k_1+d}{d}$.

The cited papers (2; 3) prove convergence of increasingly flat radial basis function interpolants towards polynomials if the condition

$$0 \leq k_2 - k_0 \leq 2$$

holds, the intermediate k_1 being irrelevant. As examples show, this inequality is sharp as a sufficient condition for convergence. The proofs of (2; 3) are done by an ingenious application of various linear equation systems connecting polynomial coefficients to moments. However, this paper uses techniques of (1) to arrive at the same result and provide additional information fopr the case of degeneration by coalescing data points.

In contrast to (2; 3) we use the concept of a *moment basis* as in (1), repeating part of the preconditioning technique used in the final section of that paper. There is a nonsingular lower triangular $N \times N$ *moment* matrix $M = (m_{ij}),\ 1 \leq i, j \leq N$, a unique set of integers

$$0 = t_0 \leq t_1 \leq \ldots \leq t_N = k_2 \tag{4}$$

and a set of ordered multi-indices

$$\alpha^1 < \alpha^2 < \ldots < \alpha^N \tag{5}$$

with the *moment conditions*

$$\begin{aligned} \sum_{j=1}^{i} m_{ij} x_j^{\alpha} &= 0 \text{ for all } \alpha < \alpha^i,\ 1 \leq i \leq N, \\ \sum_{j=1}^{i} m_{ij} x_j^{\alpha^i} &\neq 0,\ 1 \leq i \leq N, \\ |\alpha^i| &= t_i,\ 1 \leq i \leq N. \end{aligned} \tag{6}$$

Such a matrix can be generated by applying pivoted Gaussian elimination on the monomial matrix $I\!M_{|\cdot| \leq k_2}$, but we leave computational details to section 11.

3 Function Expansions

Following (4; 1; 2), we assume an analytic radial basis function

$$\phi(r) = f(r^2) = \sum_{n=0}^{\infty} f_{2n}(-1)^n r^{2n},\ 0 \le r < R \le \infty$$

with strictly positive $f_{2n},\ n \ge 0$ to be given and scale it into

$$\phi_\epsilon(r) := \phi(\epsilon r) = f(\epsilon^2 r^2) = \sum_{n=0}^{\infty} f_{2n}(-1)^n \epsilon^{2n} r^{2n},\ 0 \le r < R \le \infty.$$

The conditions on the f_{2n} are motivated from the standard assumption of complete monotonicity (5), but we insist on strictly positive constants here. This includes all standard analytic positive definite cases, e.g. the Gaussian and inverse multiquadrics.

If we insert $r := \|x-y\|_2$ into the expansion, we need an expansion of $\|x-y\|_2^{2n}$ into monomials x^β and y^α. To this end, we define two multi-indices $\alpha,\ \beta \in \mathbb{Z}_0^d$ to have *equal parity*, in short $(\alpha, \beta) \in \mathbb{Z}_P^{2d}$ or $EQP(\alpha, \beta)$ if all components α_j and β_j have equal parity for all $j,\ 1 \le j \le d$. For later use, the reader should be aware that the boolean-valued predicate EQP satisfies rules like

$$EQP(\alpha, \beta) = EQP(\alpha, \beta + 2\gamma) = EQP(\alpha + \gamma, \beta + \gamma)$$

for any choice of multi-indices $\alpha,\ \beta,\ \gamma \in \mathbb{Z}_0^d$, and likewise for plain integers.

We use Taylor's formula twice and the multinomial formula once to get

$$
\begin{aligned}
(-1)^n \|x-y\|_2^{2n} &= (-1)^n \sum_{\alpha \in \mathbb{Z}_0^d} \frac{(-y)^\alpha}{\alpha!} D^\alpha \|x\|_2^{2n} \\
&= (-1)^n \sum_{\alpha \in \mathbb{Z}_0^d} \sum_{\beta \in \mathbb{Z}_0^d} \frac{x^\beta}{\beta!} \frac{(-y)^\alpha}{\alpha!} D_{|0}^{\alpha+\beta} \|x\|_2^{2n} \\
&= \sum_{\substack{\alpha, \beta \in \mathbb{Z}_P^{2d} \\ |\alpha+\beta| = 2n}} \frac{x^\beta}{\beta!} \frac{y^\alpha}{\alpha!} (-1)^{n-|\alpha|} \frac{n!(\alpha+\beta)!}{\left(\frac{\alpha+\beta}{2}\right)!} \\
&= \sum_{\substack{(\alpha, \beta) \in \mathbb{Z}_P^{2d} \\ |\alpha+\beta| = 2n}} c(\alpha,\beta) x^\beta y^\alpha \\
&= \sum_{\substack{\alpha, \beta \in \mathbb{Z}_0^{d} \\ |\alpha+\beta| = 2n}} c(\alpha,\beta) x^\beta y^\alpha
\end{aligned}
$$

with the symmetric functions $c(\alpha, \beta)$ and $C(\alpha, \beta)$ on $\mathbb{Z}_0^d \times \mathbb{Z}_0^d$ defined by

$$
C(\alpha,\beta) := (-1)^{\frac{|\beta|-|\alpha|}{2}} c(\alpha,\beta) := \frac{\left|\frac{\alpha+\beta}{2}\right|!(\alpha+\beta)!}{\left(\frac{\alpha+\beta}{2}\right)!\alpha!\beta!} = \frac{D_{|0}^{\alpha+\beta}\|x\|_2^{|\alpha+\beta|}}{\alpha!\beta!} \tag{7}
$$

in case of $(\alpha, \beta) \in \mathbb{Z}_P^{2d}$ and zero else. By methods of (3), it will turn out in the following section that the function C above is a positive definite kernel on the set $\mathbb{Z}_0^d$. Note that power series with these coefficients have nice convergence properties, since Neumann's series yields

$$
\frac{1}{1+\|x-y\|_2^2} = \sum_{\alpha,\beta \in \mathbb{Z}_0^d} c(\alpha,\beta) x^\beta y^\alpha. \tag{8}
$$

This is why we do not have to worry about local convergence of series expansions occurring below. Inserting (7) into our expansion, we get

$$\begin{aligned}
\phi_\epsilon(\|x-y\|_2) &= \sum_{n=0}^{\infty} f_{2n}\epsilon^{2n}(-1)^n\|x-y\|_2^{2n} \\
&= \sum_{n=0}^{\infty} f_{2n}\epsilon^{2n} \sum_{\substack{(\alpha,\beta)\in \mathbb{Z}_P^{2d} \\ |\alpha+\beta|=2n}} c(\alpha,\beta)x^\beta y^\alpha \\
&= \sum_{(\alpha,\beta)\in \mathbb{Z}_P^{2d}} f_{|\alpha+\beta|}\epsilon^{|\alpha+\beta|}c(\alpha,\beta)x^\beta y^\alpha \\
&= \sum_{\alpha,\beta\in \mathbb{Z}_0^d} f_{|\alpha+\beta|}\epsilon^{|\alpha+\beta|}c(\alpha,\beta)x^\beta y^\alpha
\end{aligned}$$

where we define f_n to be zero for n odd.

4 Expansion Kernels

We now consider symmetric matrices having elements

$$f_{|\alpha+\beta|}c(\alpha,\beta) \tag{9}$$

for $\alpha,\ \beta \in I$ from any index set $I \subset \mathbb{Z}_0^d$. Fortunately, following (3), such matrices are nonsingular under mild assumptions.

Lemma 1 *Let $\Phi() := \phi(\|.\|_2)$ be a positive definite radial kernel which is inverse Fourier transformable on $\mathbb{R}^d$ from a generalized Fourier transform which is nonnegative everywhere and positive on a set of positive measure in $\mathbb{R}^d$. Then the kernel C of* (7) *is positive definite, and symmetric matrices formed by elements of the form* (9) *are nonsingular.*

Proof: We start with

$$
\begin{aligned}
\Phi(x-y) &= \sum_{n=0}^{\infty} f_{2n}(-1)^n \|x-y\|_2^{2n} \\
&= \sum_{n=0}^{\infty} f_{2n} \sum_{\substack{(\alpha,\beta)\in \mathbb{Z}_p^{2d} \\ |\alpha+\beta| = 2n}} f_{|\alpha+\beta|} c(\alpha,\beta) x^\beta y^\alpha \\
&= \sum_{(\alpha,\beta)\in \mathbb{Z}_p^{2d}} f_{|\alpha+\beta|} c(\alpha,\beta) x^\beta y^\alpha \\
&= \sum_{\alpha,\beta\in \mathbb{Z}_0^d} (-1)^{|\alpha|} \frac{D^{\alpha+\beta}\Phi(0)}{\alpha!\beta!} x^\beta y^\alpha
\end{aligned}
$$

for $x, y \in \mathbb{R}^d$, where the last equality is Taylor's formula. Since Φ is positive definite and inverse Fourier transformable, we look at a specific quadratic form with coefficients b_α for all $\alpha \in I \subset \mathbb{Z}_0^d$ and get

$$
\begin{aligned}
0 &\le \int_{\mathbb{R}^d} \hat{\Phi}(\omega) \left| \sum_{\alpha\in I} b_\alpha \omega^\alpha \right|^2 d\omega \\
&= \sum_{\alpha\in I} \sum_{\beta\in I} b_\alpha b_\beta \int_{\mathbb{R}^d} \hat{\Phi}(\omega)\omega^{\alpha+\beta} d\omega \\
&= \sum_{\alpha\in I} \sum_{\beta\in I} b_\alpha b_\beta (-iD)^{\alpha+\beta}\Phi(0) \\
&= \sum_{\substack{\alpha,\beta\in I \\ (\alpha,\beta)\in \mathbb{Z}_P^{2d}}} b_\alpha b_\beta (-1)^{|\alpha+\beta|/2} D^{\alpha+\beta}\Phi(0) \\
&= \sum_{\alpha\in I} \sum_{\beta\in I} \alpha!\, b_\alpha\, \beta!\, b_\beta\, f_{|\alpha+\beta|}\, (-1)^{\frac{|\alpha|-|\beta|}{2}}\, c(\alpha,\beta),
\end{aligned}
$$

where we have used that $c(\alpha, \beta)$ vanishes if α, β are not of equal parity. Therefore all matrices with entries $f_{|\alpha+\beta|}C(\alpha, \beta)$ based on arbitrary index sets I are positive semidefinite. But if the above expression is zero, and if we use our special assumption (which rules out the Bessel kernel), the polynomial in the first integrand must vanish on an open set, thus all coefficients must be zero. This proves positive definiteness. As a byproduct, we get positive definiteness of the C kernel itself, if we use the inverse quadric (8) with $f_{2n} = 1$ for all n. Furthermore, all symmetric matrices formed with elements $f_{|\alpha+\beta|}\ c(\alpha, \beta)$ will be nonsingular. □

Repeating the proof with complex coefficients b_α reveals that all matrices formed by elements $f_{|\alpha+\beta|}\ c(\alpha, \beta)$ are positive definite over $\mathbb{C}$. In fact, if one defines $g_\alpha := b_\alpha \alpha!(-1)^{|\alpha|/2}$, the sum above runs over $g_\alpha \overline{g_\beta} f_{|\alpha+\beta|}\ c(\alpha, \beta)$.

5 Expansions of Interpolants

If we solve an interpolation problem on $X := \{x_1, \dots, x_N\}$ using ϕ_ϵ and data $y_1, \dots, y_N$, the system

$$\sum_{j=1}^{N} a_j(\epsilon)\phi_\epsilon(\|x_j - x_\ell\|_2) = y_\ell,\ 1 \le \ell \le N \tag{10}$$

has a unique solution for all $\epsilon > 0$ which can be written as a quotient of determinants by Cramer's rule. The coefficients $a_j(\epsilon)$ come out as rational functions of ϵ with a leading term of the form ϵ^{-2k}. We start by connecting this k to relevant quantities for polynomial interpolation.

Theorem 1 *Under the assumptions of Lemma 1, the coefficients $a_j(\epsilon)$ have expansions starting with ϵ^{-2k_2}.*

Proof: We proceed very similarly to the final section of (1) on preconditioning. From now on, multi-indices $\alpha,\ \beta$ will always vary in $\mathbb{Z}_0^d$, and we only state additional conditions. Let $A(\epsilon)$ be the matrix arising in (10) and use the matrix M from (6) to form the matrix $MA(\epsilon)M^T$ with the (r,s)-entry

$$\begin{aligned}
& \sum_{j=1}^{r} m_{rj} \sum_{\ell=1}^{s} m_{s\ell}\phi_\epsilon(\|x_j - x_\ell\|_2) \\
= & \sum_{\alpha,\beta \in \mathbb{Z}_0^d} f_{|\alpha+\beta|}\epsilon^{|\alpha+\beta|} c(\alpha,\beta) \sum_{j=1}^{r} m_{rj}x_j^\beta \sum_{\ell=1}^{s} m_{s\ell}x_\ell^\alpha \\
= & \sum_{\substack{\beta \ge \alpha^r \\ \alpha \ge \alpha^s \\ EQP(\alpha,\beta)}} f_{|\alpha+\beta|}\epsilon^{|\alpha+\beta|} c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha)
\end{aligned}$$

with *moments*

$$\nu(r,\beta) := \sum_{j=1}^{r} m_{rj}x_j^\beta,\ 1 \le r \le N,\ \beta \in \mathbb{Z}_0^d \tag{11}$$

having the properties

$$\begin{aligned}
\nu(r,\beta) &= 0 \text{ for all } \beta < \alpha^r,\ 1 \le r \le N, \text{ in particular} \\
\nu(r,\alpha^s) &= 0 \text{ for all } 1 \le s < r \le N, \\
\nu(r,\alpha^r) &\ne 0,\ 1 \le r \le N
\end{aligned} \tag{12}$$

due to (6). We can collect the terms as

$$
\begin{aligned}
&\sum_{j=1}^{r} m_{rj} \sum_{\ell=1}^{s} m_{s\ell}\phi_\epsilon(\|x_j - x_\ell\|_2)\\
&= \epsilon^{t_r+t_s} \sum_{\substack{\beta \geq \alpha^r \\ \alpha \geq \alpha^s \\ EQP(\alpha,\beta)}} f_{|\alpha+\beta|}\epsilon^{|\alpha+\beta|-t_r-t_s} c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha)\\
&=: \epsilon^{t_r+t_s} B_{r,s}(\epsilon)
\end{aligned}
$$

to define a symmetric positive definite $N \times N$ matrix $B(\epsilon)$ which converges for $\epsilon \to 0$ to $B(0)$ with entries

$$
B_{rs}(0) = \sum_{\substack{|\beta| = t_r \\ |\alpha| = t_s \\ EQP(\alpha,\beta)}} f_{t_r+t_s} c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha)
$$

for $1 \leq r, s \leq N$ with equal parity of t_r and t_s, and zero else.

Lemma 2 *The matrix $B(0)$ is nonsingular.*

Proof: We take an arbitrary $u \in I\!R^N$, define the set

$$
I := \{\alpha \in Z\!\!Z_0^d \ : \ |\alpha| = t_r \text{ for some } r,\ 1 \leq r \leq N\}
$$

and a function R which associates to each $\beta \in I$ the set

$$
R(\beta) := \{j \ : \ |\beta| = t_j,\ 1 \leq j \leq N\}.
$$

Then we evaluate the quadratic form

$$
\begin{aligned}
&\sum_{r=1}^{N}\sum_{s=1}^{N} u_r u_s B_{r,s}(0)(-1)^{\frac{t_r-t_s}{2}} \\
=&\sum_{r=1}^{N}\sum_{s=1}^{N} u_r u_s (-1)^{\frac{t_r-t_s}{2}} \sum_{\substack{|\beta|=t_r\\|\alpha|=t_s\\\alpha,\beta\in I\\EQP(\alpha,\beta)}} f_{|\alpha+\beta|}c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha) \\
=&\sum_{\substack{\alpha,\beta\in I\\EQP(\alpha,\beta)}} f_{|\alpha+\beta|}c(\alpha,\beta)(-1)^{\frac{|\alpha|-|\beta|}{2}} \sum_{r\in R(\beta)} u_r\nu(r,\beta)\sum_{s\in R(\alpha)} u_s\nu(s,\alpha)
\end{aligned}
$$

which clearly is positive semidefinite due to Lemma 1 and because it is the limit of positive definite quadratic forms. It is positive definite, because from

$$\sum_{r\in R(\beta)} u_r\nu(r,\beta) = 0 \text{ for all } \beta\in I$$

we can conclude $u=0$ by inserting $\beta=\alpha^1,\dots,\alpha^N$ one after another, applying (12). This finishes the proof of the lemma. □

With an $N\times N$ diagonal matrix $D(\epsilon)$ with entries $\epsilon^{-t_k},\ 1\le k\le N$ the system (10) is rewritten as

$$
\begin{aligned}
y &= A(\epsilon)a(\epsilon)\\
D(\epsilon)My &= \underbrace{D(\epsilon)MA(\epsilon)M^TD(\epsilon)}_{=:B(\epsilon)}D^{-1}(\epsilon)(M^T)^{-1}a(\epsilon)\\
&= B(\epsilon)D^{-1}(\epsilon)(M^T)^{-1}a(\epsilon)
\end{aligned}
$$

to get the solution as a rational vector valued function

$$a(\epsilon) = M^TD(\epsilon)B^{-1}(\epsilon)D(\epsilon)My$$

for all positive ϵ with an asymptotic behavior which has at most $\epsilon^{2t_N}=\epsilon^{2k_2}$ in the denominator. □

We shall not use Theorem 1 directly, because it concerns the coefficients of interpolants in terms of the degenerating basis $\phi(\epsilon\|x-x_j\|_2),\ 1\le j\le N$. Naturally, these coefficients are much less stable than coefficients $u_j(x,\epsilon)$ of a Lagrange basis. This observation motivates the next section.

6 Expansions of Lagrange Bases

We write the standard linear system for Lagrange interpolating functions $u_j(x,\epsilon)$ satisfying $u_j(x_k,\epsilon) = \delta_{jk},\ 1 \leq j,k \leq N$ as $A(\epsilon)u(x,\epsilon) = \Phi_X(x,\epsilon)$ with

$$\Phi_X(x,\epsilon) := (\phi(\epsilon\|x - x_1\|_2), \dots, \phi(\epsilon\|x - x_N\|_2))^T$$

and transform it into

$$\underbrace{D(\epsilon)MA(\epsilon)M^T D(\epsilon)}_{=B(\epsilon)}\underbrace{D^{-1}(\epsilon)(M^{-1})^T u(x,\epsilon)}_{=:v(x,\epsilon)} = \underbrace{D(\epsilon)M\Phi_X(x,\epsilon)}_{=:w(x,\epsilon)} \tag{13}$$

to make it stably solvable, as we shall see, led by the last section of (1). We expand the elements of the $B(\epsilon)$ matrix as follows:

$$\begin{aligned} B_{rs}(\epsilon) &= \sum_{\substack{|\beta| \geq t_r \\ |\alpha| \geq t_s}} f_{|\alpha+\beta|}\epsilon^{|\alpha+\beta|-t_r-t_s} c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha) \\ &= \sum_{n=0}^{\infty} \epsilon^n \sum_{\substack{|\beta| \geq t_r \\ |\alpha| \geq t_s \\ |\alpha+\beta| = n+t_r+t_s}} f_{|\alpha+\beta|} c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha) \\ &=: \sum_{n=0}^{\infty} \epsilon^n B_{r,s,n} \end{aligned} \tag{14}$$

with coefficients

$$B_{r,s,n} = \sum_{\substack{|\beta| \geq t_r \\ |\alpha| \geq t_s \\ |\alpha+\beta| = n+t_r+t_s}} f_{|\alpha+\beta|} c(\alpha,\beta)\nu(r,\beta)\nu(s,\alpha)$$

which are zero unless $EQP(n, t_r + t_s)$ holds. The components of the right-hand side of (13) are

$$
\begin{aligned}
w_j(x,\epsilon) &:= \epsilon^{-t_j} \sum_{k=1}^{j} m_{jk} \phi(\epsilon \|x - x_k\|_2) \\
&= \epsilon^{-t_j} \sum_{k=1}^{j} m_{jk} \sum_{\alpha,\beta} f_{|\alpha+\beta|} \epsilon^{|\alpha+\beta|} c(\alpha,\beta) x^\beta x_k^\alpha \\
&= \sum_{\substack{\alpha,\beta \\ |\alpha| \geq t_j}} f_{|\alpha+\beta|} c(\alpha,\beta) \epsilon^{|\alpha|-t_j} (\epsilon x)^\beta \nu(j,\alpha) \\
&= \sum_{n=0}^{\infty} \epsilon^n \sum_{\substack{\alpha,\beta \\ |\alpha| = n+t_j}} f_{|\alpha+\beta|} c(\alpha,\beta) (\epsilon x)^\beta \nu(j,\alpha) \\
&= \sum_{n=0}^{\infty} \epsilon^n w_{j,n}(\epsilon x)
\end{aligned}
\tag{15}
$$

where we defined

$$
\begin{aligned}
w_{j,n}(y) &:= \sum_{\substack{\alpha,\beta \\ |\alpha| = n+t_j}} f_{|\alpha+\beta|} c(\alpha,\beta) y^\beta \nu(j,\alpha) \\
&=: \sum_{\beta} w_{j,n,\beta} y^\beta
\end{aligned}
$$

having coefficients

$$
w_{j,n,\beta} = \sum_{\substack{\alpha \\ |\alpha| = n+t_j}} f_{|\alpha+\beta|} c(\alpha,\beta) \nu(j,\alpha) \tag{16}
$$

which can be nonzero only if $EQP(|\beta|, n+t_j)$ holds. The special representation (15) of the right-hand side leads us to postulate a similar representation

$$
v_j(x,\epsilon) = \sum_{n=0}^{\infty} \epsilon^n v_{j,n}(x\epsilon) \tag{17}
$$

for the solution. If we plug this into the full system, we get

$$\begin{aligned}
&\sum_{n=0}^{\infty} \epsilon^n w_{r,n}(x\epsilon) \\
&= \sum_{n=0}^{\infty} \epsilon^n \sum_{s=1}^{N} B_{r,s,n} \sum_{m=0}^{\infty} \epsilon^m v_{s,m}(x\epsilon) \\
&= \sum_{k=0}^{\infty} \epsilon^k \sum_{m=0}^{k} \sum_{s=1}^{N} v_{s,m}(x\epsilon) B_{r,s,k-m}
\end{aligned}$$

and this is satisfied, because the nonsingularity of $B(0)$ proven in Lemma 2 allows to solve the recursive linear system

$$\begin{aligned}
w_{r,n}(y) &= \sum_{m=0}^{n} \sum_{s=1}^{N} v_{s,m}(y) B_{r,s,n-m} \\
&= \sum_{s=1}^{N} v_{s,n}(y) B_{r,s,0} + \sum_{m=0}^{n-1} \sum_{s=1}^{N} v_{s,m}(y) B_{r,s,n-m}
\end{aligned}$$

for all $n \geq 0, \ 1 \leq r \leq N$. This justifies (17) and allows a recursive componentwise calculation in the form

$$\begin{aligned}
\sum_{\beta \in \mathbb{Z}_0^d} w_{r,n,\beta} y^\beta &= \sum_{m=0}^{n} \sum_{s=1}^{N} \sum_{\beta \in \mathbb{Z}_0^d} v_{s,m,\beta} y^\beta B_{r,s,n-m} \\
&= \sum_{\beta \in \mathbb{Z}_0^d} y^\beta \left(\sum_{s=1}^{N} v_{s,n,\beta} B_{r,s,0} + \sum_{m=0}^{n-1} \sum_{s=1}^{N} v_{s,m,\beta} B_{r,s,n-m} \right) \qquad (18) \\
w_{r,n,\beta} &= \sum_{s=1}^{N} v_{s,n,\beta} B_{r,s,0} + \sum_{m=0}^{n-1} \sum_{s=1}^{N} v_{s,m,\beta} B_{r,s,n-m}
\end{aligned}$$

of the representation

$$v_{s,m}(y) =: \sum_{\beta \in \mathbb{Z}_0^d} v_{s,m,\beta} y^\beta.$$

Here, nonzero coefficients can only occur if $EQP(|\beta|, m + t_s)$ holds, as was the case for the $w_{r,n}$ expansion coefficients. To see this, we look inductively at (18) in case $EQP(|\beta|, n + t_r)$ fails. Then the left-hand side is zero, and so is the double sum, because it contains only terms with $EQP(|\beta|, m + t_s)$ and $EQP(n - m, t_r + t_s)$ which imply $EQP(|\beta|, n + t_r)$. Thus the solution is zero.

We now exploit $u(x,\epsilon) = M^T D(\epsilon) v(x,\epsilon)$ component-wise with (17) to get

$$\begin{aligned}
u_j(x,\epsilon) &= \sum_{m=0}^{\infty} \epsilon^m \sum_{k=j}^{N} m_{k,j} \epsilon^{-t_k} v_{k,m}(x\epsilon) \\
&= \sum_{m=0}^{\infty} \epsilon^m \sum_{k=j}^{N} m_{k,j} \epsilon^{-t_k} \sum_{\alpha \in \mathbb{Z}_0^d} v_{k,m,\alpha} x^\alpha \epsilon^{|\alpha|} \\
&= \epsilon^{-t_N} \sum_{n=0}^{\infty} \epsilon^n \sum_{k=j}^{N} m_{k,j} \sum_{\substack{\alpha \in \mathbb{Z}_0^d \\ |\alpha| \le n + t_k - t_N}} v_{k,n+t_k-t_N-|\alpha|,\alpha} x^\alpha \qquad (19) \\
&=: \epsilon^{-t_N} \sum_{n=0}^{\infty} \epsilon^n P_{j,n}(x)
\end{aligned}$$

with polynomials and coefficients

$$\begin{aligned}
P_{j,n}(x) &:= \sum_{\substack{\alpha \in \mathbb{Z}_0^d \\ |\alpha| \le n}} P_{j,n,\alpha} x^\alpha \\
P_{j,n,\alpha} &:= \sum_{\substack{k = j \\ |\alpha| \le n + t_k - t_N \le n}}^{N} v_{k,n+t_k-t_N-|\alpha|,\alpha} m_{k,j}.
\end{aligned} \qquad (20)$$

Note that the worst-case degeneration of Lagrange basis functions is only like $\epsilon^{-k_2} = \epsilon^{-t_N}$, while the solution of (10) can degenerate like ϵ^{-2k_2}.

7 Convergence Conditions

Now it is time to draw conclusions from the above expansions.

Lemma 3 *All polynomials $P_{j,n}$ are zero unless $EQP(n, k_2)$ holds.*

Proof: In fact, the equation for $P_{j,n,\alpha}$ contains only terms with

$$EQP(|\alpha|, n + t_k - t_N - |\alpha| + t_k) = EQP(0, n - t_N) = EQP(0, n - k_2).\square$$

As in the cited papers, the expansion (19) implies

$$P_{j,n}(x_k) = 0, \quad 1 \le j,k \le N, n \ge 0, \ n \ne k_2 = t_N$$
$$P_{j,k_2}(x_k) = \delta_{jk}, \ 1 \le j,k \le N.$$

Theorem 2 *For analytic positive definite radial basis functions with positive Fourier transforms on a set of positive measure,* increasingly flat *interpolants will converge to a polynomial, if $k_2 \le k_0 + 2$ holds.*

Proof: Assume non-convergence. Then there are j, n with $1 \le j \le N, .\ 0 \le n < k_2$ and $EQP(n, k_2)$ such that $P_{j,n}$ does not vanish. This polynomial then must have a degree larger than k_0, because it vanishes on X and is nonzero. This implies

$$k_0 < \deg P_{j,n} \le n \le k_2 - 2 < k_2 \text{ with } EQP(n, k_2).$$

□

8 Radial Coalescence

We now leave the "increasingly flat" scenario. From (15) we see that

$$w_j(x,\epsilon) = \sum_{n=0}^{\infty} \epsilon^n w_{j,n}(y) = \epsilon^{-t_j} \sum_{k=j}^{N} m_{jk}\phi(\|y - \epsilon x_k\|_2), \ y = \epsilon\, x$$

is the right-hand side of a Lagrange-type system of equations where ϕ is not scaled, but where the data points $\epsilon\, x_k$ coalesce radially into zero for $\epsilon \to 0$. This is a model case for what happens for fixed scaling of ϕ but for data points getting dense. The associated linear functionals

$$\lambda_{j,\epsilon}(f) := \epsilon^{-t_j} \sum_{k=1}^{j} m_{jk} f(\epsilon x_k), \ 1 \le j \le N, \tag{21}$$

when used for interpolation with the basis

$$w_j(x,\epsilon) = \lambda^y_{j,\epsilon}\phi(\|x - y\|_2)$$

generated by $\lambda_{j,\epsilon}$ acting with respect to y on $\phi(\|x - y\|_2)$, lead to the interpolation matrix with entries

$$\lambda^x_{r,\epsilon}\lambda^y_{s,\epsilon}\phi(\|x - y\|_2) = B_{r,s}(\epsilon) \tag{22}$$

we know already. The interpolation is carried out with the basis of functions $v_j(x,\epsilon)$ satisfying (13). Our scaling is such that in the dual of the native Hilbert space (6) for ϕ we have

$$\|\lambda_{r,\epsilon}\|_\phi^2 = B_{r,r}(\epsilon) \to B_{r,r}(0) > 0$$

for $\epsilon \to 0$. Consequently, we see that our use of M and $D(\epsilon)$ is the right way to precondition problems with this kind of coalescence. The finite-dimensional interpolation space which arises in the limit will now not consist of polynomials, but rather be spanned by the functions

$$w_{j,0}(y) = \sum_{\substack{\alpha,\beta \\ |\alpha| = t_j}} f_{|\alpha+\beta|} c(\alpha,\beta) y^\beta \nu(j,\alpha). \tag{23}$$

They span the same space as the functions $v_{j,0}$ we get when taking the limit of (13), because the $v_{j,0}$ are generated from the $w_{j,0}$ by application of $B(0)^{-1}$. The functions above are of the form $w_{j,0}(x,\epsilon) = \lambda_{j,0}^y \phi(\|x-y\|_2)$ for limit functionals

$$\lambda_{j,0}(x^\alpha) := \left\{ \begin{array}{cc} \nu(j,\alpha) & |\alpha| = t_j \\ 0 & \text{else} \end{array} \right\}$$

which act like t_j-fold derivatives at zero. They still are linearly independent because of

$$(\lambda_{r,0}, \lambda_{s,0})_\phi = B_{rs}(0),\ 1 \le r,s \le N.$$

Theorem 3 *Radially coalescent Lagrange interpolation problems converge towards Hermite interpolation problems with a maximal differentiation order k_2 of limit functionals.* $\square$

It should be remarked that (7) contains the basics of Hermite interpolation by radial basis functions.

9 Newton Interpolation

The foregoing sections contained some rather heavy machinery, but they followed a strategy which is well-known from univariate polynomial interpolation. In fact, the transition from Lagrange to Hermite interpolation via the Newton interpolation formula is precisely what happened above. To see this more clearly, we drop the parameter ϵ in this section.

First, in univariate situations, we make the transition from function values $f(x_1), \ldots, f(x_N)$ to the N divided differences

$$\lambda_j(f) := f[x_1, \ldots, x_j],\ 1 \le j \le N$$

in standard terminology, but written as N linear functionals which have the form (28) with a moment matrix that does not appear explicitly but satisfies (6). The connection between (28) and divided differences is based on the property

$$\lambda_j(x^\alpha) = 0,\ |\alpha| < t_j = |\alpha^j| = j - 1$$

in 1D, as assured by the moment matrix via (6). Then the Newton basis

$$v_j(x) := \prod_{i=1}^{j-1}(x - x_i),\ 1 \le j \le N$$

generates the Newton interpolant

$$p(x) := \sum_{j=1}^{N} \lambda_j(f) v_j(x), \tag{24}$$

and the interpolation process nicely converges for coalescing points into a Hermite interpolation problem, if written in Newton form. Note that the limit functionals are derivatives of order up to $N-1$ in 1D, but only up to order $k_2 \le N - 1$ in our multivariate theory.

While our functionals of (21) correspond nicely to divided differences, we still have to see how our basis corresponds to the Newton basis. The crucial fact is that the Newton basis satisfies

$$\lambda_j(v_k) = \delta_{jk},\ 1 \le j, k \le N, \tag{25}$$

as follows from (24). Note that the case $j < k$ relies on the fact that v_k vanishes on $x_1, \ldots, x_{k-1}$, while the case $j > k$ is standard for divided differences, because they annihilate lower-order polynomials.

In our technique, the system (13) can be written as

$$\sum_{j=1}^{N} (\lambda_j, \lambda_k)_\phi v_j(x) = w_k(x)$$

using (22), both for positive ϵ and in the limit $\epsilon \to 0$. But the definition of the w_k implies $\lambda_j(w_k) = (\lambda_j, \lambda_k)_\phi$, and thus we have (25). If we use the fact that the M matrix is lower triangular by construction, we immediately get something similar to the 1D case:

Lemma 4 *The functions $v_{j,\epsilon}$ of the transformed interpolation process satisfy*

$$v_{j,\epsilon}(x_k) = 0,\ 1 \le k < j \le N,\ \epsilon > 0. \qquad \square$$

10 General Coalescence

We now turn to the harder problem of N more or less *freely* coalescing points at zero. To this end, we assume that our data points $x_k(h)$ move along smooth curves for $h \to 0$ into $0 = x_k(0)$. For simplicity, we assume $\|x_k(h)\|_2 \le h$ throughout. The geometry now is h-dependent, and the characteristic multi-indices $\alpha^j(h)$ of (5) and the $t_j(h) := |\alpha^j(h)|$ of (4) will vary with h. But we shall focus on sequences $h_k \to 0$ where these discrete quantities do not vary any more. Thus we ignore their dependence on h again.

If we define points $y_k(h)$ by $x_k(h) = h\,y_k(h)$ such that the $y_k(h)$ still vary smoothly, the geometric quantities derived for the $y_k(h)$ are the same as those for $x_k(h)$, because the columns of the monomial matrices just get different scalar factors. We assume that higher-order monomials of the $y_k(h)$ can be stably calculated via

$$y_k^{\alpha}(h) = \sum_{j=1}^{N} d(j,h,\alpha) y_k^{\alpha^j}(h),\ 1 \le k \le N,\ |\alpha| > k_2 \tag{26}$$

from lower-order monomials, with uniformly bounded coefficients $d(j,h,\alpha)$. From the definition of k_2 this is clear if the y_k are constant, but we allow them to vary here, allowing a much more general but still somewhat regular coalescence of the $x_k(h)$.

The above identity describes how the column with multi-index α of the monomial matrix can be reconstructed from the N linear independent columns corresponding to the $\alpha^j,\ 1 \le j \le N$. In our coalescence scenario, the above identity, when rewritten in terms of the $x_k^{\alpha}(h)$, turns into

$$x_k^{\alpha}(h) = \sum_{j=1}^{N} d(j,h,\alpha) h^{|\alpha|-|\alpha^j|} x_k^{\alpha^j}(h),\ 1 \le k \le N,\ |\alpha| > k_2 \tag{27}$$

and describes in a natural way how the larger powers of the x_k vanish faster than the lower ones for $h \to 0$. This provides a good reason why (26) should be assumed.

We can always find h-dependent $N \times N$ moment matrices $M(h) = (m_{jk}(h))$

such that the linear functionals

$$\lambda_{j,h}(f) := \sum_{k=1}^{N} m_{jk}(h) f(x_k(h)),\ 1 \le j \le N, \tag{28}$$

are orthonormal in the native space of ϕ. This can be done by orthogonalizing in the span of the functionals $\delta_{x_k(h)},\ 1 \le k \le N$ in the native space, which is equivalent to orthogonalization of the $N \times N$ positive definite matrix with entries

$$(\delta_{x_j(h)}, \delta_{x_k(h)})_\phi = \phi(\|x_j(h) - x_k(h)\|_2),\ 1 \le j, k \le N,\ h > 0.$$

Due to their normalization, the functionals of (28) must be weak-$*$-convergent, and thus there are limit functionals $\lambda_{j,0}$ with norm one in the dual of the native space such that

$$\lambda_{j,0}(f) = \lim_{h \to 0} \lambda_{j,h}(f)$$

for suitable subsequences and all f in the native space of ϕ. The whole problem works in the span of the right-hand sides

$$w_{j,h}(y) := \lambda_{j,h}^x \phi(\|x - y\|_2)$$

which nicely converge in the native space towards

$$w_{j,0}(y) := \lambda_{j,0}^x \phi(\|x - y\|_2),\ 1 \le j \le N,$$

whatever these functions actually are, and the orthogonalization of our functionals imply the Lagrange property

$$\lambda_{k,h}^y w_{j,h}(y) = \lambda_{j,h}^x \lambda_{k,h}^x \phi(\|x - y\|_2) = \delta_{jk},\ 1 \le j, k \le N$$

for all positive h. Clearly, the limit functionals must be supported in zero only, but we want to figure out that they are necessarily derivatives at zero of order up to k_2. From (27) we get uniform convergence

$$\lambda_{m,h}(x^\alpha) = \sum_{j=1}^{N} d(j, h, \alpha) h^{|\alpha| - |\alpha^j|} \lambda_{m,h}(x^{\alpha^j}) \to 0$$

for $h \to 0$ and all $|\alpha| > k_2$. This proves

$$\lambda_{j,0}(x^\alpha) = 0 \text{ for all } |\alpha| > k_2,\ 1 \le j \le N.$$

But for general functions f the functionals act like

$$\begin{aligned}
\lambda_{j,0}(f) &= \sum_{\alpha \in \mathbb{Z}_0^d} \frac{D^\alpha f(0)}{\alpha!} \lambda_{j,0}(x^\alpha) \\
&= \sum_{|\alpha| \le k_2} \frac{D^\alpha f(0)}{\alpha!} \lambda_{j,0}(x^\alpha)
\end{aligned}$$

proving that they are derivatives of order at most k_2, as required. Now we can also check the limit of the orthogonality. Since convergence is not strong, we cannot directly conclude

$$\delta_{jk} = \lim_{h \to 0} (\lambda_{j,h}, \lambda_{k,h})_\phi \stackrel{?}{=} (\lambda_{j,0}, \lambda_{k,0})_\phi,$$

but we can consider the limit of

$$\begin{aligned}
\delta_{jk} &= (\lambda_{j,h}, \lambda_{k,h})_\phi \\
&= \sum_{\alpha,\beta \in \mathbb{Z}_0^d} c(\alpha,\beta) f_{|\alpha+\beta|} \lambda_{j,h}(x^\alpha) \lambda_{k,h}(y^\beta) \\
&\to \sum_{\substack{\alpha,\beta \in \mathbb{Z}_0^d \\ |\alpha|,|\beta| \le k_2}} c(\alpha,\beta) f_{|\alpha+\beta|} \lambda_{j,0}(x^\alpha) \lambda_{k,0}(y^\beta) \\
&= (\lambda_{j,0}, \lambda_{k,0})_\phi, \ 1 \le j,k \le N.
\end{aligned}$$

Theorem 4 *Lagrange interpolation problems based on coalescing data sites satisfying* (26) *converge to Hermite problems whose functions and functionals are defined by limit functionals being certain derivatives of order at most* k_2 *at the coalescence point.*

11 Computations

We add a few remarks concerning the actual calculation of all important terms arising in the above equations. A MAPLE© worksheet for 2D examples is available from the author.

First, for given N data points in $\mathbb{R}^d$ forming a set X, a sufficiently large monomial matrix $\mathbb{M}$ must be generated such that its rank is N. In worst possible, i.e. essentially univariate cases, this takes a degree up to $N-1$. In general, we can get away with the maximal degree $k_2 \le N-1$, but at this point we do not know k_2, forcing us to start with a monomial matrix that is "sufficiently large" to have rank N. A standard LU decomposition with row permutations then leads to a (usually non-square) U matrix with N rows, permutations just acting on the points of X. The first nonzero entry of U in row j, counted from left to right, corresponds to a unique column multi-index α^j, thus defining the sequence (5). This determines the $t_j := |\alpha^j|$ of (4) with $k_2 = t_N$, and the L^{-1} matrix of the decomposition yields the moment matrix M in (6). Looking at the largest row index i, $1 \le i \le N$ where the nonzero pivots still lie on the diagonal, we get k_0 as the largest k with $\binom{k+d}{d} \le i$. In short, if looking at the staircase-shaped positions of first nonzero elements in rows of U,

- k_0 is the degree after which the staircase leaves the diagonal to move to the right,
- k_1 is the degree necessary for forming the left $N \times N$ submatrix,
- k_2 is the degree at which the staircase hits the bottom row.

This illustrates (2). Up to here, there are no radial basis functions involved, and it is no problem to calculate a polynomial interpolant based on the monomials x^{α^i}, $1 \leq i \leq N$ for comparison with what comes later. This interpolant is somewhat special in that it uses the minimum possible number of monomials, and it only uses those with a certain minimality with respect to the ordering of the exponents. In many cases, it looks preferable to every other interpolating polynomial.

We now turn to increasingly flat radial basis function interpolants. Inserting an expansion of f as a sequence of numbers f_{2n}, we arrive at the problem to determine how far to calculate all of our expansions. Looking back at (20), it suffices to calculate the $P_{j,n,\alpha}$ for $1 \leq j \leq N$, $0 \leq n \leq k_2$ and $|\alpha| \leq k_2$. This requires $v_{j,n,\alpha}$ for the same range. From (18) we see that also the $w_{j,n,\alpha}$ share this range, and we need the $B_{r,s,k}$ for $1 \leq r, s \leq N$, $0 \leq k \leq k_2$. However, equations (16) and (14) imply that we need the c and ν values for multi-indices $|\alpha| \leq 2k_2$ to calculate those values. Altogether, this fixes finite data to work with, and it is quite straightforward to program all necessary linear algebra calculations.

Using this strategy, the 2D examples of (2) can easily be reproduced, the solutions given there being the polynomials P_{1,k_2} here. But since (2) deals with the irrelevant number k_1 and additional non-degeneracy conditions which are only implicitly related to k_0 and k_2, we add a table to supply the constants connecting these examples to the theory of this paper.

Example	k_0	k_1	k_2	N	Data
2.1	1	2	2	4	general
2.2	1	2	3	6	on parabola
2.3	0	2	5	6	on a line
2.4	2	2	2	6	general
2.5	1	2	3	6	on a circle

Clearly, by Theorem 2 only Example 2.3 can show degeneracy, and it does. If N data are given on a line, we have (3), and Theorem 2 implies that the first possibly degenerate case can be $k_2 = 3$ and $N = 4$. However, by MAPLE© one can show that the first degenerate situation for one-dimensional data embedded in 2D occurs for $N = 5$.

If data are on a non-degenerate conic like a circle or a parabola, we have $k_0 = 1$ as in Examples 2.2 and 2.5 of (2), because there is a nontrivial second-degree polynomial vanishing on all data sites, independent of the number N of these. In this case we need $k_2 \geq 4$ by Theorem 2 to find a degenerate case. This occurs, for instance, if two points are added to the 6 points on a parabola in Example 2.2 of (2). Then $k_2 = 4$ holds, and Theorem 2 turns out to be sharp.

Finally, we consider the coalescence case. There, the limit of Lagrange basis functions makes no sense. Likewise, the degenerating linear systems for coalescing points should not be solved at all. Thus we work our way backwards, constructing the Hermite limit interpolants first, and evaluating them at various sequences of coalescing points in order to show that the Hermite interpolants agree with the given function at coalescing points up to terms of order h.

Thus we pick a smooth function g and use fixed data locations $x_k,\ 1 \leq k \leq N$ to start with. The coalescing points will then be defined a-posteriori by the user to satisfy $z_k(h) := h\, x_k + O(h^2)$. Then we calculate either exact Hermite data $\lambda_{j,0}(g)$ or approximate data

$$\lambda_{j,h}(g) := h^{-t_j} \sum_{k=1}^{N} m_{jk} g(z_k(h))$$

near zero and form the functions

$$\begin{aligned} s(x) &:= \sum_{j=0}^{N} \lambda_{j,0}(g) v_{j,0}(x) \\ s_h(x) &:= \sum_{j=0}^{N} \lambda_{j,h}(g) v_{j,0}(x) \end{aligned}$$

using the fixed Hermite basis functions $v_{j,0}(x)$ which we calculate explicitly by multiplying the vector of functions $w_{j,0}$ of (23) by $B(0)^{-1}$. They satisfy $\lambda_{k,0}(v_{j,0}) = \delta_{jk},\ 1 \leq j,k \leq N$, as we pointed out in section 9. The functions s and s_h are exact or approximate solutions, respectively, to the Hermite problem at zero.

Then we can use MAPLE© to see symbolically that the error functions $g - s$ and $g - s_h$ behave like $O(h^2)$ for $h \to 0$ at the points $z_k(h)$. The degenerating interpolants defined via Lagrange interpolation at the $z_k(h)$ are not calculated at all. The fixed function s arises as the limit of all coalescing cases with points $z_k(h) := h\, x_k(h) + O(h^2)$, no matter how they are defined. If the functionals $\lambda_{j,0}$ are numerically approximated by $\lambda_{j,h}$, the same limit is attained via s_h for

$h \to 0$, but the calculation of $\lambda_{j,h}(g)$ is numerically unstable due to cancellation in (21). The numerical instability is confined to the approximate calculation of the Hermite functionals, while the linear system has fixed condition and does not degenerate. If the user calculates s instead of s_h, there are no numerical degenerations at all, at the expense of not solving the coalescing problems exactly.

12 Open Problems

The conditions given by Theorem 2 are sufficient to guarantee convergence for the "increasingly flat" case, and they are sharp as far as conditions are formulated using k_0 and k_2 only. However, convergence is equivalent to certain equations guaranteeing $P_n = 0$ for all $k_0 < n < k_2$ with $EQP(n, k_2)$, and these come up as complicated rational expressions involving the data set X and the expansion coefficients $f_{2\ell}$ of the radial basis function ϕ. Thus there may be special cases of X and ϕ where there is convergence outside the sufficient condition of Theorem 2. A particular case, conjectured by Driscoll and Fornberg (4) and proven in (1), surprisingly states that the Gaussian lets these conditions be satisfied in all cases, no matter what the geometry of X is. In other words: the Gaussian overcomes all possible geometric degenerations. The same property holds experimentally for the Bessel radial basis function $J_0(r)$, but this one fails to satisfy the assumptions of Lemma 1 and leads to singular matrices (3). The special role of the Gauss and Bessel kernels are still a mystery. Our MAPLE© procedures allow some explicit experimentation along these lines, but there are no theoretical results known so far.

For the "coalescence" scenario, our methods indicate how to cope with data points that come too close and thus spoil the condition of the linear system. We showed how to transform the Lagrange interpolation problem into a Hermite interpolation problem with a stable limit. This is a first case of a preconditioning technique for such systems, but it is still nor efficient enough. However, the analysis leads to a complete understanding of the degeneration process and the stable limit of the transformed process.

Coalescence situations will automatically occur, if adaptive methods calculate approximations of functions that are derivatives of the kernel at a fixed point. Investigations of such methods are under way, since they proved to be rather efficient (8; 9) in practice, even for solving partial differential equations by collocation (10; 11).

Acknowledgment

Special thanks go to Bengt Fornberg and in particular Elisabeth Larsson for various valuable comments and corrections.

References

[1] R. Schaback, Multivariate interpolation by polynomials and radial basis functions, Constructive Approximation 21 (2005) 293–317.

[2] E. Larsson, B. Fornberg, Theoretical and computational aspects of multivariate interpolation with increasingly flat radial basis functions, Comput. Math. Appl. 49 (1) (2005) 103–130.

[3] Y. J. Lee, G. J. Yoon, J. Yoon, Convergence property of increasingly flat radial basis function interpolation to polynomial interpolation, preprint, submitted, on `http://math.ewha.ac.kr/ yoon/approx/paper.html` (2005).

[4] T. Driscoll, B. Fornberg, Interpolation in the limit of increasingly flat radial basis functions, Comput. Math. Appl. 43 (2002) 413–422.

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

[6] H. Wendland, Scattered Data Approximation, Cambridge University Press, 2005

[7] Z. Wu, Hermite–Birkhoff interpolation of scattered data by radial basis functions, Approximation Theory and its Applications 8/2 (1992) 1–10.

[8] Y. Hon, R. Schaback, X. Zhou, An adaptive greedy algorithm for solving large rbf collocation problems, Numerical Algorithms 32 (2003) 13–25.

[9] R. Schaback, H. Wendland, Adaptive greedy techniques for approximate solution of large RBF systems, Numerical Algorithms 24 (2000) 239–254.

[10] C. Franke, R. Schaback, Solving partial differential equations by collocation using radial basis functions, Appl. Math. Comp. 93 (1998) 73–82.

[11] C. Franke, R. Schaback, Convergence order estimates of meshless collocation methods using radial basis functions, Advances in Computational Mathematics 8 (1998) 381–399.