Multivariate Interpolation and Approximation by Translates of a Basis Function

Robert Schaback

Abstract. This contribution will touch the following topics:

- Short introduction into the theory of multivariate interpolation and approximation by finitely many (irregular) translates of a (not necessarily radial) *basis function*, motivated by optimal recovery of functions from discrete samples.
- *Native* spaces of functions associated to conditionally positive definite functions, and relations between such spaces.
- *Error bounds* and *condition numbers* for interpolation of functions from native spaces.
- *Uncertainty Relation*: Why are good error bounds always tied to bad condition numbers?
- *Shift and Scale*: How to cope with the Uncertainty Relation?

§1. Introduction and Overview

This contribution contains the author's view of a certain area of multivariate interpolation and approximation. It is not intended to be a complete survey of a larger area of research, and it will not account for the history of the theory it deals with. Related surveys are [15, 21, 22, 27, 30, 47, 48, 58].

Section 2 will motivate why we are mainly interested in multivariate functions that are linear combinations

$$f(x) = \sum_{j=1}^{N} a_j \Phi(x - x_j),\ x \in \mathbb{R}^d,\ a_j \in \mathbb{R} \tag{1.1}$$

of *translates* of a fixed *basis* function $\Phi : \mathbb{R}^d \to \mathbb{R}$, where $X = \{x_1, \dots, x_N\}$ is a set of N scattered points in $\mathbb{R}^d$. The space of functions (1.1) often

Approximation Theory VIII
Charles K. Chui and Larry L. Schumaker (eds.), pp. 1–8.

ISBN 0-12-xxxxxx-x

carries a natural topology induced by an inner product, and its Hilbert space closure (the "native space" for Φ) will be studied in Section 3. Error bounds for interpolation of functions from the native space are intimately connected to upper bounds on $A_{X,\Phi}^{-1}$ for the N by N matrix

$$A_{X,\Phi} := (\Phi(x_j - x_k))_{1\le j,k\le N}. \tag{1.2}$$

This connection will be dealt with in Sections 4 and 5, where we see that basis functions Φ with good error bounds on their native space will necessarily have bad upper bounds for $\|A_{X,\Phi}^{-1}\|$. A consequence of this "uncertainty relation" is that one has to introduce scaled versions of Φ in order to cope with the bad condition of $A_{X,\Phi}$. Thus the final section contains refined error bounds for approximation by scales of basis functions.

§2. Optimal Recovery via Discrete Sampling

Sampling theory provides a good reason for considering functions of the form (1.1). The most common example is the reconstruction of a univariate bandlimited function f by its sinc series

$$f(x) = \sum_{j\in\mathbf{Z}} f(j)\,\mathrm{sinc}\,(x-j)$$

with

$$\mathrm{sinc}\,(x) = \frac{\sin \pi x}{\pi x}\,.$$

But we do not want to start with (1.1). Instead, we shall pose the reconstruction problem in a general way and find later that (1.1) is the natural solution.

Assume we want to recover a multivariate function $f : \mathbb{R}^d \to \mathbb{R}$ from a sample of values $y_1 = f(x_1), \ldots, y_N = f(x_N)$ on a discrete (and "scattered") set $X = \{x_1, \ldots, x_N\}$ in $\mathbb{R}^d$. Besides these values, we only have partial information on the spectrum of f, *i.e.*, on its Fourier transform. More precisely, if we know that

$$f(x) = (2\pi)^{-d}\int_{\mathbb{R}^d} \hat f(\omega) e^{ix\cdot\omega} d\omega,\ x\in\mathbb{R}^d \tag{2.1}$$

recovers f from its Fourier transform $\hat f$, we shall assume that $\hat f$ lies in a weighted L_2 space

$$L_2(\rho) := \left\{ g \in L_1(\mathbb{R}^d) : \int_{\mathbb{R}^d} |g(\omega)|^2 \rho(\omega) d\omega < \infty \right\},$$

where $\rho : \mathbb{R}^d \to \mathbb{R}_{>0} \cup \{+\infty\}$ with $\rho(-\cdot) = \rho(\cdot)$ is a Lebesgue–measurable function that is finite on at least an open subset of $\mathbb{R}^d$. These things can be generalized to a distributional setting, but we want to keep the presentation simple.

The weight ρ determines a norm $\| \ \|_{2,\rho}$ on the space

$$\mathcal{F} := \{f : \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\},\ \hat{f} \in L_2(\rho) \text{ and } (2.1)\} \tag{2.2}$$

via the bilinear form

$$(f, g)_{2,\rho} := (2\pi)^{-d} \int_{\mathbb{R}^d} \overline{\hat{f}(\omega)}\hat{g}(\omega)\rho(\omega)d\omega.$$

By symmetry of ρ and $\hat{f}(-\omega) = \overline{\hat{f}(\omega)}$ this form is real–valued. Then we can ask for *optimal recovery* in the sense of the problem

$$\text{Find } f \in \mathcal{F} \text{ with } \|f\|_{2,\rho} \text{ minimal}$$

$$\text{such that } f(x_j) = y_j,\ 1 \le j \le N.$$

The solution is given by the following theorem whose proof contains some standard arguments that we include to make this paper self–contained.

Theorem 2.1. *If* $\frac{1}{\rho} \in L_1(\mathbb{R}^d)$ *and* $\hat{\Phi} := \frac{1}{\rho}$, *then a solution of the above recovery problem exists uniquely and is necessarily of the form* (1.1). *The coefficients* $a = (a_1, \ldots, a_N)^T$ *can be found by solving the system*

$$A_{X,\Phi}a = y \tag{2.3}$$

where $y = (y_1, \ldots, y_N)^T$. *The matrix* $A_{X,\Phi}$ *comes from* (1.2) *and is positive definite and symmetric.*

Proof. Assume first that a solution $f \in \mathcal{F}$ to the recovery problem does exist. By the usual perturbation argument for characterization of best approximations in Euclidean spaces we then have

$$(f, g)_{2,\rho} = 0 \tag{2.4}$$

for all $g \in \mathcal{F}$ with $g(x_j) = 0,\ 1 \le j \le N$. Then there are $a_1, \ldots, a_N \in \mathbb{R}$ with

$$(f, g)_{2,\rho} = \sum_{j=1}^{N} a_j g(x_j)$$

for all $g \in \mathcal{F}$, and we introduce Fourier transforms on both sides to get

$$(2\pi)^{-d} \int_{\mathbb{R}^d} \overline{\hat{f}(\omega)}\hat{g}(\omega)\rho(\omega)d\omega = (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{g}(\omega) \sum_{j=1}^{N} a_j \cdot e^{i\omega\cdot x_j}\, d\omega$$

for all $g \in \mathcal{F}$. This is satisfied if and only if

$$\hat{f}(\omega) = \frac{1}{\rho(\omega)} \sum_{j=1}^{N} a_j e^{-i\omega\cdot x_j} =: \hat{\Phi}(\omega)\cdot\sigma_{a,X}(\omega)$$

$$\hat{\Phi} = \frac{1}{\rho}\ ,\ \sigma_{a,X} := \sum_{j=1}^{N} a_j e^{-i\omega\cdot x_j}$$

or iff f is of the form (1.1).

Our argument clearly shows how translates of a basis function come up very naturally, but so far it is no proof of the theorem. To prove existence, we show that $A_{X,\Phi}$ is positive definite. Indeed, for any $a \in \mathbb{R}^N$ we have

$$\begin{aligned} a^T A_{X,\Phi} a &= \sum_{j,k=1}^{N} a_j a_k \Phi(x_j - x_k) \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{1}{\rho(\omega)} \left|\sigma_{a,X}(\omega)\right|^2 d\omega \end{aligned} \tag{2.5}$$

and this nonnegative quantity is zero iff $\sigma_{a,X}(\omega)$ vanishes on the support of $1/\rho$. But this support contains an open set and $\sigma_{a,X}$ is analytic. Thus $\sigma_{a,X}(\omega)$ vanishes on $\mathbb{R}^d$ iff (2.5) vanishes. Since the functions $e^{-i\omega\cdot x_j}$ are linearly independent, the positive definiteness of $A_{X,\Phi}$ follows.

But then there is a unique f of the form (1.1) that satisfies (ff) and thus minimizes $\|f\|_{2,\rho}$ under all other interpolants from $\mathcal{F}$.

If f_1 and f_2 would both satisfy (2.4) and the interpolation conditions, then $\|f_1 - f_2\|_{2,\rho}^2 = 0$ follows from (2.4) and this proves unicity in general. ∎

Theorem 2.1 shows that translates of a single basis function naturally arise when solving an optimal recovery problem in spaces $\mathcal{F}$ of functions with Fourier transforms in weighted L_2 spaces. Note that shifts or translates do not directly occur in the problem setting, except that $\mathcal{F}$ is a space that is invariant under all shifts of translations in $\mathbb{R}^d$. The notion "shift invariant space", however, has been extensively used (see *e.g.*, [11, 12, 13]) for spaces invariant under shifts in $\mathbf{Z}^d$ only. We remark in passing that a function space $\mathcal{F}$ necessarily is of the form (2.2), provided that it is

(a) a Hilbert space,
(b) continuously imbedded in $L_2(\mathbb{R}^d)$,
(c) allowing continuous point evaluation functionals $\delta_x f = f(x)$ for $f \in \mathcal{F}$ and $x \in \mathbb{R}^d$, and
(d) separating points in $\mathbb{R}^d$, and

(e) invariant under all shifts in $\mathbb{R}^d$.

The weight function ρ then comes out to be $\rho = 1/\hat{\Phi}$ with $\Phi(x-y) := (\delta_x, \delta_y)_{\mathcal{F}'}$ for $x, y \in \mathbb{R}^d$. Details of this will be given in a forthcoming paper.

To give the reader an idea of the scope of this approach, we first recall an example with a very small space $\mathcal{F}$ of functions, *i.e.*, the *bandlimited* functions. Here, the weight function is

$$\rho_p(x) = \left\{ \begin{array}{cc} 1 & x \in B_p \\ \infty & x \notin B_p \end{array} \right\}$$

where B_p is the unit ball with respect to the L_p norm in $\mathbb{R}^d$. Then $1/\rho_p$ coincides with the characteristic function χ_{B_p} of B_p and $\hat{\Phi} = \chi_{B_p}$ holds. If $p = \infty$, then Φ is a d–fold tensor product of sinc functions. For $p = 2$ we get the jinc functions

$$\Phi(x) = \|x\|_2^{-d/2} J_{d/2}(\|x\|_2)$$

up to multiplicative constants. This is the simplest instance of a *radial* basis function Φ, *i.e.*,

$$\Phi(x) = \phi(\|x\|_2), \qquad \phi : \mathbb{R}_{\geq 0} \to \mathbb{R}.$$

Adopting the terminology of signal analysis, we note that this case yields very good localization in the frequency domain, because Fourier transforms are compactly supported. In the time domain we have algebraically decaying functions

$$\Phi(x) = \|x\|_2^{-d/2} J_{d/2}(\|x\|_2) = \mathcal{O}(\|x\|_2^{-(d+1)/2})$$

for $\|x\|_2 \to \infty$.

A more straightforward example with a small space $\mathcal{F}$ is given by the Gaussian $\Phi(x) = e^{-\|x\|^2/2}$ which is symmetric in time and frequency domain.

Note that the bandlimited case cannot be turned upside down by swapping frequency and time domain, because the jinc functions have sign changes and cannot be used as weight functions. To construct compactly supported functions with nonnegative Fourier transforms, some additional techniques are needed. For instance, convolution in the time domain always produces nonnegative Fourier transforms, but the problem then is to evaluate the result in the time domain. We cite a successful construction due to Wu [64] that proceeds as follows:

Table 1.
Time–Frequency Localization.

	Time ($\|x\|_2 \to \infty$)	Frequency ($\|\omega\|_2 \to \infty$)
Bandlimited	$\|x\|_2^{-(d+1)/2}$	comp. supported
$W_2^k(\mathbb{R}^d)$	$\|x\|_2^{k-(d+1)/2}e^{-\|x\|_2}$	$\|\omega\|_2^{-2k}$
Gaussian	$e^{-\|x\|_2^2/2}$	$e^{-\|\omega\|_2^2/2}$
inverse Multiquadrics	$-d < 2\beta$ $(c^2+\|x\|_2^2)^\beta$, $\beta < 0$	$\|\omega\|_2^{-(\beta+d+1)/2} \cdot e^{-\|\omega\|_2}$
$\Phi_{\ell,k}$ (Wu) in $\mathbb{R}^{2m+1}$, $0 \le m \le k \le \ell$	comp. supported	$\|\omega\|_2^{-2(\ell+1+m-k)}$

$$
\begin{aligned}
\varphi_\ell(r) &:= (1-r^2)_+^\ell\\
\phi_{\ell,0} &:= \varphi_\ell * \varphi_\ell\\
\phi_{\ell,k}(r) &:= \left(-\frac{1}{r}\ \frac{d}{dr}\right)\phi_{\ell,k-1}(r) \qquad 1 \le k \le \ell.
\end{aligned}
$$

It generates radial functions

$$\Phi_{\ell,k}(x) := \phi_{\ell,k}(\|x\|_2)$$

that are positive definite on $\mathbb{R}^d$ for $d \le 2k+1$. Details of this construction can be found in [64], and a toolbox for handling radial functions is in [59].

A rather large space $\mathcal{F}$ is generated by the weight

$$\rho_k(\omega) = (1+\|\omega\|^2)^k,\ \omega \in \mathbb{R}^d$$

for $2k > d$, and we get Sobolew space

$$\mathcal{F} = W_2^k(\mathbb{R}^d)$$

and the corresponding optimal basis function

$$\Phi(x) = \|x\|_2^{-(k-d/2)} K_{k-d/2}(\|x\|_2),$$

where K_ν is the Bessel function of third kind. Since the K_ν functions are nonnegative, this example can be turned upside down and yields "*inverse multiquadrics*"

$$\Phi(x) = (1+\|x\|_2^2)^{\beta/2}$$

for $0 > \beta > -d$, where now

$$\rho(\omega) = \|\omega\|_2^{(\beta+d)/2} K^{-1}_{(\beta+d)/2}(\|\omega\|_2)$$

acts as a weight function for the Fourier transforms. Note that this example and the Gaussian force Fourier transforms of functions in $\mathcal{F}$ to decay exponentially. Thus the space $\mathcal{F}$ consists of C^∞ functions.

We summarize the known prototypes in Table 1 and remark that something like the "up"–function still is missing: a case with compact support and C^∞ smoothness in the time domain while having exponential decay and positivity in the Fourier domain.

Historical Remarks

There is a vast literature on optimal recovery starting from Golomb/Weinberger [24] continuing via Sard [52], Micchelli/Rivlin [38], using results on abstract spline theory as recently summarized by Atteia [3] and ending up in Information–based Complexity as defined in Traub/Woźniakowski [63] and Bojanov/Woźniakowski [10]. Sampling theory also is a wide–ranging field (see *e.g.*, the review by Butzer/Stens [16]). We borrowed the ingredients of our presentation from the folklore of both subjects, ignoring (so far) the background of reproducing–kernel Hilbert spaces that will show up in the next section. Spaces of functions that have Fourier transforms in weighted L_2 spaces occur as examples for complex interpolation theory in the fundamental papers of Calderon [17] and Schechter [60]. See the review of Pisier [46] in this volume for details of complex interpolation. We touch upon another connection to interpolation theory of normed vector spaces at the end of Section 6. Another important link to classical results arises with the theory of Riesz and Bessel potentials, which correspond to thin–plate splines and multiquadrics. See, for instance, Aronszajn [1] and Calderon [18]. Our view on native spaces as "principal translation–invariant spaces" borrows from ideas of de Boor, de Vore and Ron [11, 12, 13] on shift–invariant spaces.

§3. Native Spaces

The preceding section showed that optimal reconstruction of functions in a quite general translation–invariant space $\mathcal{F}$ of functions naturally leads to the consideration of functions that are linear combinations of translates of a single basis function Φ. Thus the space $\mathcal{F}$ determined a function Φ.

We now go the opposite direction and start with a function Φ to generate a "native" space $\mathcal{F}$. In contrast to the first approach we do not work on all of $\mathbb{R}^d$ but rather on a subdomain $\Omega \subset \mathbb{R}^d$.

Definition 3.1 A function $\Phi : \Omega - \Omega \to \mathbb{R}$ with $\Phi(-\cdot) = \Phi(\cdot)$ is *positive definite* on Ω, if for all sets $X = \{x_1, \ldots, x_N\}$ of N pairwise distinct points in Ω the matrix $A_{X,\Phi}$ in (1.2) is positive definite.

Note that there is a slight difference in terminology with [39, 61, 62] and others, but it definitely is bad notation to call Φ positive definite when $A_{X,\Phi}$ is positive semidefinite or nonnegative definite for all X.

The space $\mathcal{D}(\Omega)$ of all finitely supported linear functionals on the space $\mathbb{R}^\Omega$ of all real–valued functions on Ω can now be equipped with an inner product. Indeed, if functionals

$$\lambda = \sum_{i=1}^{M} \lambda_i \delta_{x_i} \in \mathcal{D}(\Omega), \qquad \mu = \sum_{j=1}^{N} \mu_j \delta_{y_i} \in \mathcal{D}(\Omega) \tag{3.1}$$

with $\lambda_i, \mu_j \in \mathbb{R},\ x_i, y_j \in \Omega$ are given, then

$$(\lambda, \mu)_\Phi := \sum_{i=1}^{M} \sum_{j=1}^{N} \lambda_i \mu_j \Phi(x_i - y_j)$$

is an inner product on $\mathcal{D}(\Omega)$, and we can form the Hilbert space $\overline{\mathcal{D}(\Omega)}$ by taking the closure with respect to $(\cdot,\cdot)_\Phi$. Note that our first definition of a space concerned a space of functionals, but now we can go over to functions of the form (1.1) by

$$f_\lambda(x) := \sum_{i=1}^{M} \lambda_i \Phi(x - x_i) = (\delta_x, \lambda)_\Phi \tag{3.2}$$

for λ from (3.1) and $x \in \Omega$. Allowing all $\lambda \in \overline{\mathcal{D}(\Omega)}$ we define the native space for Φ to be

$$\mathcal{F}_{\Phi,\Omega} := \{f_\lambda(x) = (\delta_x, \lambda)_\Phi,\ x \in \Omega,\ \lambda \in \overline{\mathcal{D}(\Omega)}\} \tag{3.3}$$

and this space can easily shown to be isometrically isomorphic to $\overline{\mathcal{D}(\Omega)}$ because of

$$\mu(f_\lambda) = (\lambda, \mu)_\Phi$$

for all $\lambda, \mu \in \overline{\mathcal{D}(\Omega)}$, if we define

$$(f_\lambda, f_\mu)_\Phi := (\lambda, \mu)_\Phi$$

on $\mathcal{F}_{\Phi,\Omega}$. An equivalent definition, as given by Madych and Nelson [31, 32, 33] in case of $\Omega = \mathbb{R}^d$, is the space

$$F_{\Phi,\Omega} = \{f : \Omega \to \mathbb{R} : |\mu(f)| \le C(f)\|\mu\|_\Phi,\ \mu \in \mathcal{D}(\Omega)\} \tag{3.4}$$

of all functions on Ω that allow all elements of $\mathcal{D}(\Omega)$ as bounded linear functionals with respect to the topology induced by Φ on $\mathcal{D}(\Omega)$.

If we look at the dependence of $\mathcal{F}_{\Phi,\Omega}$ on Ω, we should consider a subdomain $\Omega_1 \subseteq \Omega$. Since clearly $\mathcal{D}(\Omega_1) \subseteq \mathcal{D}(\Omega)$, we find

$$\mathcal{F}_{\Phi,\Omega_1} \subseteq \mathcal{F}_{\Phi,\Omega\big|\Omega_1} := \{f \in \mathcal{F}_{\Phi,\Omega} \text{ restricted to } \Omega_1\}$$

if we use (3.3), but the equivalent definition (3.4) implies $\mathcal{F}_{\Phi,\Omega\big|\Omega_1} \subseteq \mathcal{F}_{\Phi,\Omega_1}$. This proves Iske's extension theorem

$$\mathcal{F}_{\Phi,\Omega\big|\Omega_1} = \mathcal{F}_{\Phi,\Omega_1} \qquad \text{for } \Omega_1 \subset \Omega,$$

if Φ is positive definite on Ω (see [26]). In case of $\Omega = \mathbb{R}^d$ it implies that any function in a native space $\mathcal{F}_{\Phi,\Omega_1}$ corresponding to a domain Ω_1 has a (non–unique) extension to all of $\mathbb{R}^d$ that preserves all the implicit smoothness assumptions that are hidden in the definition of $\mathcal{F}_{\Phi,\Omega}$. In case of the Gaussians or the inverse multiquadrics all functions in $\mathcal{F}_{\Phi,\Omega_1}$ are restrictions of functions in $\mathcal{F}_{\Phi,\Omega} \subset C^\infty(\mathbb{R}^d)$, *i.e.*, they have canonical $C^\infty(\mathbb{R}^d)$ extensions. The construction of an extension is clear from the definition of $\mathcal{F}_{\Phi,\Omega_1}$, since the function

$$f_\lambda(x) := (\delta_x, \lambda)_\Phi$$

for $\lambda \in \overline{\mathcal{D}(\Omega_1)} \subseteq \overline{\mathcal{D}(\Omega)}$ can not only be evaluated for $x \in \Omega_1$, but also for $x \in \Omega$, thus providing an extension. It is an interesting question to ask for the maximal domain Ω on which a given function Φ is positive definite; the existence of non–extensible positive definite functions on bounded domains is a hard problem already in the one–dimensional case (see [28, 61]), and there are results of Rudin [50, 51] and Krein [28] in the multivariate case. Another challenge is the full characterization of positive definite functions on $\mathbb{R}^d$ under the weakest possible assumptions (see *e.g.*, [19, 25, 62, 65] for this problem).

Historical Remarks

Our presentation mainly follows ideas of Madych-Nelson [31] but confines itself to *unconditionally* positive definite functions. The work of Madych-Nelson in turn builds mainly on results of the French school, expecially on Duchon's thin–plate splines [20] and earlier work on abstract spline theory as provided by Atteia [2], Laurent [29], and later also by Meinguet [37].

§4. Error Bounds and Pointwise Optimality

Consider quasi–interpolants of the form

$$f \mapsto s_{f,u,X} = \sum_{j=1}^{N} f(x_j) \cdot u_j \tag{4.1}$$

where $X = \{x_1, \ldots, x_N\} \subset \Omega \subseteq \mathbb{R}^d$ is a set of scattered points, and where $u_1, \ldots, u_N : \Omega \to \mathbb{R}$ are arbitrary functions. If $x \in \Omega$ is fixed, the error functional

$$\varepsilon_{x,u,X}(f) := f(x) - s_{f,u,X}(x)$$

is in $\mathcal{D}(\Omega)$, and thus there is an error bound of the form

$$|\varepsilon_{x,u,X}(f)| \le \|f\|_\Phi \cdot \|\varepsilon_{x,u,X}\|_\Phi \tag{4.2}$$

for functions f in the native space $\mathcal{F}_{\Phi,\Omega}$. The norm of the error functional is explicitly available via

$$\begin{aligned} \|\varepsilon_{x,u,X}\|_\Phi^2 = \Phi(0) - 2\sum_{j=1}^{N} u_j(x)\Phi(x - x_j) \\ + \sum_{k=1}^{N}\sum_{j=1}^{N} u_k(x)u_j(x)\Phi(x_k - x_j). \end{aligned} \tag{4.3}$$

This allows comparison of various quasi–interpolants, and one can ask for an optimal choice of $u_1(x), \ldots, u_N(x)$ that minimizes $\|\varepsilon_{x,u,X}\|_\Phi^2$ in the above representation as a nonnegative quadratic form. If $u^* := (u_1^*(x), \ldots, u_N^*(x))^T$ denotes the minimizer, then clearly

$$\sum_{j=1}^{N} \Phi(x_k - x_j)u_j^*(x) = \Phi(x - x_k), \qquad 1 \le k \le N. \tag{4.4}$$

Thus the functions $u_j^*(x)$ are in the span of the $\Phi(x - x_k)$, $1 \le k \le N$, and they necessarily satisfy the Lagrange–type interpolation conditions

$$u_j^*(x_k) = \delta_{jk}, \qquad 1 \le j, k \le N,$$

since the system (4.4) is uniquely solvable due to the positive definiteness of the coefficient matrix we already know from Theorem 2.1. This proves

Theorem 4.1. *The interpolant of the form* (1.1) *to scattered data is the unique minimizer of the error bound* (4.2) *under all quasi–interpolants* (4.1).

Note that this result indicates that general error bounds in native spaces cannot be improved by choosing the shifts occurring in (1.1) different from the data locations used in (4.1).

It is interesting to plot the pointwise norms $\|\varepsilon_{x,u,X}\|_\Phi$ of the error functionals as a function of x for $u = u^*$ or for different choices of u. This makes sense even in the classical univariate situation, and it would be worthwhile for applications to see a series of illustrative examples. Minimization with respect to u may lead to unexpected results in other than Hilbert space settings (for instance in $W^1_\infty[a,b]$). Another possibility is the variation of X while always using the optimal interpolant u^* that will depend on X. Then $\|\varepsilon_{x,u^*,X}\|^2_\Phi$ will be a smooth function that allows minimization with respect to X. This may lead to a future theory that generalizes perfect splines. Numerical results of A. Beyer [8] were encouraging.

At this point it is by no means evident that the interpolant $s_{f,u^*,X}$ minimizes $\|s\|^2_\Phi$ under $s \in \mathcal{F}_{\Phi,\Omega}$ with $s(x_j) = f(x_j)$ for $1 \le j \le N$, because there is no apparent link between $\| \ \|_\Phi$ of this section and $\| \ \|_{2,\rho}$ of Section 2. To bridge this gap in full generality seems to be an open problem which is related to the full characterization of (strictly) positive definite functions, as posed in the survey of E.W. Cheney [19] in this volume. If Φ can be recovered from its Fourier transform via

$$\Phi(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{\Phi}(\omega) e^{i\omega\cdot x} d\omega,$$

and if λ is a functional of the form (3.1), then

$$\begin{aligned} \|f_\lambda\|^2_\Phi &= \|\lambda\|^2_\Phi \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{\Phi}(\omega) \left| \sum_{j=1}^N \lambda_j e^{i\omega\cdot x_j} \right|^2 d\omega. \end{aligned} \tag{4.5}$$

Furthermore, the function f_λ of (3.2) will then have a Fourier transform

$$\hat{f}_\lambda(\omega) = \hat{\Phi}(\omega) \sum_{j=1}^N \lambda_j e^{-i\omega\cdot x_j}$$

such that we can formally write

$$\|f_\lambda\|^2_\Phi = (2\pi)^{-d} \int_{\mathbb{R}^d} |\hat{f}_\lambda(\omega)|^2 \cdot \frac{1}{\hat{\Phi}(\omega)} d\omega$$

to see the similarity between $\|\cdot\|_\Phi$ and $\|\cdot\|_{2,\rho}$ with $\rho = 1/\hat{\Phi}$.

The aforementioned characterization problem for (strictly) positive definite functions leads to some specific questions in this context:

1) Which conditions guarantee the existence of $\hat{\Phi} \in L_1(\mathbb{R}^d)$ needed for (4.5)?
2) What are the properties of $\hat{\Phi}$ that ensure (4.5) to be an inner product?

A partial answer to 1) was given by Iske in this dissertation. If Φ is locally absolutely integrable and of at most polynomial growth, then it has a generalized Fourier transform $\hat{\Phi}$. For Φ positive definite, Iske proved under mild additional assumptions that $\hat{\Phi}$ is in $L_1(\mathbb{R}^d)$. The second question is more difficult, because it involves the structure of the possible zeros of limits of exponential sums as occurring in (4.5). In the univariate case this requires tools from almost periodic functions (see Fang [23]).

The optimal error bound (4.2) is rather abstract and needs further elaboration in terms of the density of the data. A bounded domain $\Omega \subset \mathbb{R}^d$ usually is fixed, and for a positive $\tau \in \mathbb{R}$ one considers a local density

$$h_{X,\tau}(x) := \sup_{\|y-x\|_2 \leq \tau} \min_{1 \leq j \leq N} \|y - x_j\|_2, \; x \in \Omega \tag{4.6}$$

where for technical improvement one can replace the ball $\{y : \|y-x\|_2 \leq \tau\}$ by cones or cubes with vertex x. Then there are three somewhat different techniques for proving error bounds of the form

$$\|\varepsilon_{x,u^*,X}\|_\Phi^2 \leq F_\Phi(h_{X,\tau}(x)) \tag{4.7}$$

with $F : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ monotonic and $F(0) = 0$. If (4.7) holds, then (4.2) becomes

$$|f(x) - s_{f,u^*,X}(x)| \leq \|f\|_\Phi \cdot F_\Phi^{1/2}(h_{X,\tau}(x)) \tag{4.8}$$

for all $X = \{x_1, \ldots, x_N\} \subset \Omega$ and all $x \in \Omega$ with $h_{X,\tau}(x) \leq h_0$, where h_0 is a positive constant depending only on τ, Φ, and Ω. Table 2 shows the currently known functions F_Φ for various choices of Φ. Note, however, that Φ determines both F_Φ and $\|.\|_\Phi$, making error bounds hard to compare. In general, F_Φ gets smaller with improving smoothness of Φ, but at the same time $\|.\|_\Phi$ gets more restrictive, since $1/\hat{\Phi}$ acts as a penalty weight for Fourier transforms.

We finally give the reader some pointers to the three proof techniques of (4.7). The first is confined to radial functions $\Phi(x) = \phi(\|x\|_2)$, but does not involve Fourier transforms. It reduces the problem to quantitative polynomial approximation of ϕ on $[0, h]$ for $h \to 0$, and dates back to Madych-Nelson [32], being refined somewhat in [34]. The other two use Fourier transforms of non–radial functions and can be found in Madych-Nelson [33] and Wu/Schaback [66]. In all three cases it is crucial to have good bounds on Lebesgue constants of polynomial interpolation of perturbed regular data in $\mathbb{R}^d$. By using explicit geometric configurations for few data points, Powell [49] could get small and explicit bounds for thin–plate splines via knowledge of good Lebesgue constants. The problem of optimal bounds for Lebesgue constants of multivariate polynomial interpolation is still open in general and its solution would improve the known error bounds of type (4.8) a lot. To our knowledge, the best general bounds are in [34], and papers by Bloom [9] and Bos [14] show what is possible for regular data.

§5. Condition Numbers and Uncertainty Relation

Numerical experiments show that the condition number of $A_{X,\Phi}$ as in (1.2) is terribly large for smooth Φ like Gaussians or Multiquadrics when compared to non–smooth Φ like thin–plate splines. The spectral condition number of $A_{X,\Phi}$, being the quotient of largest and smallest eigenvalue, was observed to be boosted up mainly because of the smallest eigenvalue being extraordinarily small. To understand this phenomenon, a series of papers by Ball, Baxter, Narcowich, Sivakumar, and Ward [4, 5, 6, 7, 40, 41, 42, 43, 44, 45] investigated lower bounds

$$\lambda^T A_{X,\Phi}\lambda \geq c(X,\Phi)\|\lambda\|_2^2$$

of the quadratic form associated with $A_{X,\Phi}$. This is equivalent to bounding the smallest eigenvalue of $A_{X,\Phi}$ from below.

The basic trick in most of these papers is to construct a “minorant” Ψ such that

$$\lambda^T A_{X,\Phi}\lambda \geq \lambda^T A_{X,\Psi}\lambda \geq c(X,\Psi)\|\lambda\|_2^2,$$

where $A_{X,\Psi}$ is diagonally dominant and $c(X,\Psi)$ is readily available. A short and general account of the technique can be found in [56]. The results are always of the form

$$c(X,\Phi) \geq G_\Phi(q_X)$$

where $G_\Phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a monotonic function with $G(0) = 0$, and

$$q_X := \frac{1}{2} \min_{1\leq j<k\leq N} \|x_j - x_k\|_2$$

is the *separation distance* of points in X. Note that the bound does *not* depend on the number N of data points. Up to multiplicative constants, Table 2 shows the special cases known so far (see [56] for details).

There is a striking similarity between the functions F_Φ and G_Φ, and it seems to be impossible to find cases where the interpolation error, given by $F_\Phi^{1/2}$, is small and the condition, partly given by G_Φ^{-1}, is small as well.

This kind of "Uncertainty Relation" can be put on a solid basis by a simple argument from [56] that bridges the gap between upper bounds on errors for interpolants by X–translates of Φ and lower bounds on eigenvalues of $A_{X,\Phi}$. In fact, if we formally add $x_0 := x$ to the data set X and define

$$\begin{aligned} A_x &:= A_{X\cup\{x\},\Phi} \\ \lambda_x^* &:= (1, -u_1^*(x), \dots, -u_N^*(x)) \\ q_x &:= \min_{0\le j<k\le N} \|x_j - x_k\| \end{aligned}$$

then, when interpreted as an error bound,

$$\lambda_x^* A_x \lambda_x^* = \|\varepsilon_{x,u^*,X}\|_\Phi^2 \le F_\Phi(h_{X,\tau}(x)),$$

and when interpreted as a quadratic form,

$$\begin{aligned} \lambda_x^* A_x \lambda_x^* &\ge G_\Phi(q_x)\cdot \|\lambda_x^*\|_2^2 \\ &= G_\Phi(q_x)\left(1 + \sum_{j=1}^N |u_j^*(x)|^2\right). \end{aligned}$$

Thus

$$G_\Phi(q_x) \le G_\Phi(q_x)\left(1 + \sum_{j=1}^N |u_j^*(x)|^2\right) \le F_\Phi(h_{X,\tau}(x)) \tag{5.1}$$

proves that for comparable small arguments with

$$h_{X,\tau}(x) \approx h \approx q_x \tag{5.2}$$

one cannot have a small error bound $F_\Phi(h)$ without having a small lower bound $G_\Phi(h)$ on the smallest eigenvalue. And (5.2) clearly is possible for regular data sets X of spacing h and x placed at distance $\ge h/2$ from all points of X.

There are some further consequences of the Uncertainty Relation in the form (5.1). First, it suggests that optimal results are obtained in cases calculated near the limits of machine precision, may these be reached by huge amounts of densely distributed data points for non–smooth Φ, or may these be reached by moderate numbers of data points for smooth functions like *e.g.*, the Gaussians. When reconstructing surfaces from scattered data and varying certain parameters like c for multiquadrics $(c^2 + r^2)^{\beta/2}$ such that the condition of $A_{X,\Phi}$ tends towards the limits of machine precision, the nicest picture always is the one that immediately preceded the numerical breakdown.

Table 2.

All entries are modulo factors that are independent of r and h, but possibly dependent on parameters of Φ. Unreferenced cases for G are treated in [56].

$\Phi(x) = \phi(r),\ r = \\|x\\|_2$	$F_\Phi(h)$	$G_\Phi(h)$
$r^\beta,\ \beta \in \mathbb{R}_{>0} \setminus 2\mathbb{N}$ thin–plate splines	h^β [59]	h^β [4]: $d = \beta = 1$ [5], pg. 419: $\beta \in (0, 2)$ [42], § VI: $\beta = m - d/2$, d odd
$(-1)^{1+\beta/2} r^\beta \log r,\ \beta \in 2\mathbb{N}$ thin–plate splines	h^β [59]	h^β [42], § VI: $\beta = m - d/2$, d even
$(\gamma^2 + r^2)^{\beta/2},\ \beta \in \mathbb{R} \setminus 2\mathbb{N}_{\geq 0}$ Multiquadrics	$e^{-\frac{\delta}{h}}$ $\delta > 0$ [32]	$h^2 e^{-\frac{6}{h}}$ [5], pg. 90: $\gamma = 1 = \beta,\ d = 2$, $h e^{-\frac{2d}{h}}$ [5], pgs. 422–423: $\gamma = 1 = \beta$ $h^\beta \exp(-12.76\gamma d/h)$
$e^{-\beta r^2},\ \beta > 0$ Gaussians	$e^{-\frac{\delta}{h^2}}$ $\delta > 0$ [32]	$h^{-d} e^{-\frac{\gamma}{h^2}}$ [42], pg. 90: $\beta = 1$ $h^{-d} \exp(-40.71 d^2/(\beta h^2))$
$\frac{2\pi^{d/2}}{\Gamma(k)} K_{k-d/2}(r)(r/2)^{k-d/2}$ $2k > d$, Sobolev splines	h^{2k-d} as in [59],	h^{2k-d}

Second, one gets both new lower bounds for F_Φ, and upper bounds for G_Φ exhibiting the leeway for further optimization of both bounds. This improves earlier work by Ball/Sivakumar/Ward [5] and Schaback [53] on upper bounds for $G_\Phi(q_x)$, and it opens the race for closing the gap as much as possible by finding optimal constants.

We now want to compare several basis functions. Since we know that any basis function yields optimal error bounds with respect to its "native" space, the comparison must take place in "alien" spaces. We fix basis functions Φ_1 and Φ_2 with native spaces $\mathcal{F}_{\Phi_1}$ and $\mathcal{F}_{\Phi_2}$, and assume

$$\mathcal{F}_{\Phi_1} \subseteq \mathcal{F}_{\Phi_2},$$

so that now $\mathcal{F}_{\Phi_2}$ is alien to Φ_1 and vice versa. The data set X will be dropped in the notation for the rest of this section, and we denote by $s_{f,j}$ the interpolant to f on some X with respect to Φ_j.

We first consider interpolation of functions $f \in \mathcal{F}_{\Phi_2}$ by translates of Φ_1. These cannot reach the optimal error bounds for interpolation by Φ_2–translates, but they can be *quasi–optimal* in the sense

$$|f(x) - s_{f,1}(x)| \leq C \cdot \|f\|_{\Phi_2} \cdot F_{\Phi_2}^{1/2}(h_X(x)) \tag{5.3}$$

with a constant $C \geq 1$. Numerical results [53] suggest that quasi–optimality often holds, but so far there is only a proof for slightly perturbed interpolants [57] instead of $s_{f,1}$. The proof technique first involves an approximation f_ε of f up to some ε by chopping off the Fourier transform of f. Then $f_\varepsilon \in \mathcal{F}_{\Phi_1}$ holds and the bounds

$$\begin{aligned} |f(x) - f_\varepsilon(x)| &\leq \varepsilon \\ |f_\varepsilon(x) - s_{f_\varepsilon,1}(x)| &\leq \|f_\varepsilon\|_{\Phi_1} \cdot F_{\Phi_1}^{1/2}(h_X(x)) \end{aligned}$$

are used to prove (5.3) by choosing ε as a function of $\|f\|_{\Phi_2}$ and $F_{\Phi_2}^{1/2}(h_X(x))$ in a proper way, roughly by letting ε take the form of the right–hand side of (5.3). Unfortunately, this part of the proof does not work in general. So far, each of the traditional examples required a special analysis.

The fact that slightly perturbed interpolants often work better than exact interpolants is well known from other areas of Approximation Theory. It takes here a very specific form, and it was called "appproximate approximation" in recent papers [35, 36] by Maz'ya and Schmidt, used there in the somewhat different context of quasi–interpolation on gridded data with nonstationary Gaussians. In general, the notion of "approximate" approximation uses a two–parameter family $\{s_{h,\delta}\}$ of approximants (think of a shift parameter h and a scale parameter δ) such that there is no Weierstraß–type density result for $h \to 0$ and δ fixed, but where for each $\varepsilon > 0$ there is some $\delta > 0$ such that

$$\|f - s_{h,\delta}(f)\| \leq \varepsilon + K(f, h, \delta)$$

with $K(f, h, \delta) \to 0$ for $h \to 0$.

To give the reader an idea how this applies in our context, let δ be a parameter that controls chopping the Fourier transform of some general function $f \in L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$ at radius $1/\delta$, such that

$$\left| f(x) - (2\pi)^{-d} \int_{\|\omega\| \leq \delta^{-1}} \hat{f}(\omega) e^{i\omega \cdot x} d\omega \right|$$

$$\leq (2\pi)^{-d} \int_{\|\omega\|\geq \delta^{-1}} |\hat{f}(\omega)| d\omega =: E(f,\delta) \leq \varepsilon$$

for δ small enough. The function

$$\tilde{f}(x) := (2\pi)^{-d} \int_{\|\omega\|\leq \delta^{-1}} \hat{f}(\omega) e^{i\omega\cdot x} d\omega$$

now lies in the native space for any basis function Φ with $\hat{\Phi} \in L_1(\mathbb{R}^d)$ and $\hat{\Phi} > 0$ on $\mathbb{R}^d$, and has norm

$$\begin{aligned}
\|\tilde{f}\|_\Phi^2 &:= (2\pi)^{-d} \int_{\|\omega\|\leq \delta^{-1}} \frac{|\hat{f}(\omega)|^2}{\hat{\Phi}(\omega)}\, d\omega \\
&\leq \|f\|_2^2 \cdot (2\pi)^{-d} \int_{\|\omega\|\leq \delta^{-1}} \frac{1}{\hat{\Phi}(\omega)}\, d\omega \\
&=: \|f\|_2^2 \cdot L^2(\Phi,\delta)
\end{aligned}$$

where $L(\Phi,\delta)$ tends to infinity for $\delta \to 0$, depending on the decay properties of $\hat{\Phi}$ at infinity. Now we have

$$\begin{aligned}
|f(x) - \tilde{f}(x)| &\leq E(f,\delta) \\
|\tilde{f}(x) - s_{\tilde{f},X}(x)| &\leq \|\tilde{f}\|_\Phi \cdot F_\Phi^{1/2}(h_X(x)) \\
&\leq \|f\|_{L_2} \cdot L(\Phi,\delta) \cdot F_\Phi^{1/2}(h_X(x))
\end{aligned}$$

and by picking a suitably small $\delta(f,\varepsilon)$ we can get a bound

$$|f(x) - s_{\tilde{f},X}(x)| \leq \varepsilon + C(f,\varepsilon,\Phi) \cdot F_\Phi^{1/2}(h_X(x)).$$

Note that for $h_X(x) \to 0$ for sufficiently dense data sets X the second part of this bound behaves precisely like the bound of optimal interpolation in the native space, and this means exponential convergence in case of Gaussians or multiquadrics, for instance.

The above discussion of approximate approximation is related to K–functionals. Indeed, for any f from the space $BC(\overline{\Omega})$ of all bounded and continuous functions on the closure of a bounded domain $\Omega \subset \mathbb{R}^d$ we can write

$$\begin{aligned}
|f(x) - s_g(x)| &\leq |f(x) - g(x)| + |g(x) - s_g(x)| \\
&\leq |f(x) - g(x)| + \|g\|_\Phi F_\Phi^{1/2}(h_X(x))
\end{aligned}$$

if s_g is the interpolant to some intermediate function $g \in \mathcal{F}_\Phi$ and $X = \{x_1,\ldots,x_N\} \subset \Omega$. Thus

$$\begin{aligned}
\inf_{s\in S_{X,\Phi}} \|f-s\|_{\infty,\Omega} &\leq \inf_{g\in\mathcal{F}_\Phi} (\|f-g\|_{\infty,\Omega} + F_\Phi^{1/2}(h_X)\cdot \|g\|_\Phi) \\
&=: K(f, F_\Phi^{1/2}(h_X), BC(\Omega), \mathcal{F}_\Phi)
\end{aligned}$$

with $h_X := \|h_X(x)\|_{\infty,\Omega}$ and $S_{X,\Phi}$ standing for the space of functions of the form (1.1). It would be very interesting to see the classical machinery of K–functionals set to work towards Jackson–Bernstein theorems for approximation of functions from intermediate spaces between $BC(\Omega)$ and $\mathcal{F}_\Phi$.

§6. Shift and Scale

We now allow scaled versions

$$\Phi_\delta(\cdot) = \Phi\left(\frac{\cdot}{\delta}\right), \quad \hat{\Phi}_\delta(\cdot) = \delta^d \Phi(\cdot\delta) \quad \delta > 0$$

of a positive definite function Φ on $\mathbb{R}^d$ and consider a fixed bounded domain $\Omega \subset \mathbb{R}^d$. We first check the behavior of $\|\varepsilon_{x,u^*,X}\|^2_{\Phi_\delta}$ as a function of δ. Clearly, from the optimality of (4.4) we get

$$\begin{aligned}
(2\pi)^d \|\varepsilon_{x,u^*,X}\|^2_{\Phi_\delta} &= \min_u \int_{\mathbb{R}^d} \hat{\Phi}_\delta(\omega) \left| 1 - \sum_{j=1}^N u_j(x) e^{i\omega\cdot(x-x_j)} \right|^2 d\omega \\
&= \delta^d \min_u \int_{\mathbb{R}^d} \hat{\Phi}(\omega\delta) \left| 1 - \sum_{j=1}^N u_j(x) e^{i\omega\cdot(x-x_j)} \right|^2 d\omega \\
&= \min_u \int_{\mathbb{R}^d} \hat{\Phi}(\eta) \left| 1 - \sum_{j=1}^N u_j(x) e^{i\eta\cdot(x-x_j)/\delta} \right|^2 d\eta \\
&= (2\pi)^d \|\varepsilon_{x/\delta,u^*,X/\delta}\|^2_{\Phi}
\end{aligned}$$

as everybody would expect. In view of (4.6) in the bound (4.7) we have

$$\begin{aligned}
h_{X/\delta,\tau/\delta}(x/\delta) &= \sup_{\|y-x/\delta\|\leq\tau/\delta} \min_j \|y - x_j/\delta\| \\
&= \frac{1}{\delta} h_{X,\tau}(x) = \frac{1}{\delta} \sup_{\|y-x\|\leq\tau} \min_j \|y - x_j\|.
\end{aligned}$$

Thus (4.7) in the form

$$\|\varepsilon_{x,u^*,X}\|^2_{\Phi_\delta} = F_\Phi(h_{X/\delta,\tau/\delta}(x/\delta)) = F_\Phi\left(\frac{1}{\delta} h_{X,\tau}(x)\right) \tag{6.1}$$

shows that this part of the error depends on the *relative* scaling of Φ and the data, as expected. To make the first inequality of (6.1) fully valid we require a lower bound

$$\tau/\delta \geq \tau_0 > 0$$

for τ/δ, or an upper bound

$$\delta \leq \tau/\tau_0$$

for δ. We introduce

$$h := h_{X,\tau} := \|h_{X,\tau}(\cdot)\|_{\infty,\Omega}$$

to keep the notation somewhat simpler. The error bound (4.8) then is

$$|f(x) - s_{f,\Phi_\delta}(x)| \leq \|f\|_{\Phi_\delta} \cdot F_\Phi^{1/2}(h/\delta),$$

where s_{f,Φ_δ} is the interpolant to f with respect to Φ_δ. Note that the Uncertainty Relation does (at least for interpolation) not allow to make $F_\Phi(h/\delta)$ arbitrarily small in practical applications. Ignoring the additional factor $\|f\|_{\Phi_\delta}$, one would always be able to cope with large h by choosing a large δ to keep errors small. The *stationary* case of the literature takes δ proportional to h. We see that in this case convergence must come from the factor $\|f\|_{\Phi_\delta}$ for $\delta \to 0$, and this is the second reason why we now look at this quantity.

If we keep f fixed and check the condition $\|f\|_{\Phi_\delta} < \infty$, then

$$(2\pi)^d \|f\|^2_{\Phi_\delta} = \delta^{-d} \int_{\mathbb{R}^d} \frac{|\hat{f}(\omega)|^2}{\hat{\Phi}(\omega\delta)}\, d\omega.$$

The first and simplest case is thin–plate splines with

$$\hat{\Phi}(\omega) = \|\omega\|^{-d-\beta}$$

for some $\beta > 0$ and up to a constant factor. Then clearly $f \in \mathcal{F}_{\Phi_\delta}$ whenever $f \in \mathcal{F}_\Phi$, and

$$\|f\|^2_{\Phi_\delta} = \delta^\beta \|f\|^2_\Phi.$$

Since in this case $F_\Phi(h) = c^2(\beta, d) \cdot h^\beta$, we can write the overall error bound as

$$|f(x) - s_{f,\Phi_\delta}(x)| \leq c(\beta, d)\delta^{\beta/2}(h/\delta)^{\beta/2}\|f\|_\Phi$$

and get invariance with respect to scaling, as expected.

Now let us consider functions Φ with monotonic radial decay of $\hat{\Phi}$, *i.e.*,

$$\hat{\Phi}(\omega\delta) \geq \hat{\Phi}(\omega) > 0 \quad \text{for all } \delta \leq 1,\ \omega \in \mathbb{R}^d.$$

Then

$$\|f\|_{\Phi_\delta} \leq \delta^{-d/2}\|f\|_{\Phi_1}$$

for all $f \in \mathcal{F}_{\Phi_1},\ \delta \leq 1$. The error bound now is

$$|f(x) - s_{f,\Phi_\delta}(x)| \leq \|f\|_{\Phi_1} \cdot \delta^{-d/2} F_{\Phi_1}^{1/2}(h/\delta)$$

and one cannot use strictly stationary interpolation. Instead, one has to let h/δ tend to zero to let $F_{\Phi_1}^{1/2}(h/\delta)$ outweigh the factor $\delta^{-d/2}$. This is not very restrictive in case of exponential decay of $\hat{\Phi}$, *i.e.*, for Multiquadrics and Gaussians, because then it suffices to let h/δ decrease logarithmically with $h \to 0$, without spoiling exponential convergence for $h \to 0$.

More problematic are cases with algebraic decay of F_{Φ_1}, say, of type $F_{\Phi_1}(h) = \mathcal{O}(h^k)$. Then the error has the behavior

$$h^{k/2}\delta^{(-k-d)/2}.$$

To achieve an error of $\mathcal{O}(h^{\beta/2})$ with $0 < \beta \leq k$ (look at thin–plate splines for comparison) one has to take a scaling like

$$\delta = \mathcal{O}\left(h^{\frac{k-\beta}{k+d}}\right).$$

This allows to trade small errors for good condition by a suitable scaling, and this was the goal of this section. However, we have so far left out the most interesting cases, namely the compactly supported positive definite functions. Their Fourier transforms partially have zeros or are not known to be monotonic. Another extension concerns approximation instead of interpolation. In both cases one can use the techniques of the preceding section to generate intermediate functions $f_\delta \in \mathcal{F}_{\Phi_\delta}$ by chopping off the Fourier transform of f, and then interpolating f_δ by Φ_δ. We leave the details to a later presentation.

Acknowledgments. Help in proofreading was provided by A. Iske, and typing was done by P. Trapp.

References

1. Aronszajn, N., Potentiels Besseliens, Ann. Inst. Fourier Grenoble **15** (1965), 43–58.
2. Atteia, M., Analyse Numérique – Généralisation de la définition et des propriétés des spline functions, C.R. Acad. Sc., Paris, **260** (1965), 3550–3553.
3. Atteia, M., *Hilbertian Kernels and Spline Functions*, in: Studies in Computational Mathematics (4), North-Holland, Amsterdam, 1992.
4. Ball, K., Eigenvalues of Euclidean Distance Matrices, J. Approx. Theory **68** (1992), 74-82.
5. Ball, K., N. Sivakumar, and J.D. Ward, On the Sensitivity of Radial Basis Interpolation to Minimal Distance Separation, J. Approx. Theory **8** (1992), 401-426.

6. Baxter, B.J.C., Norm Estimates for Inverses of Distance Matrices, in *Mathematical Methods in Computer Aided Geometric Design*, T. Lyche and L.L. Schumaker (eds.), Academic Press, New York, 1989, 13–18.
7. Baxter, B.J.C., N. Sivakumar, J.D. Ward, Regarding the p–Norms of Radial Basis Interpolation Matrices, Techn. Report **281**, Department of Mathematics, Texas A & M University, 1992.
8. Beyer, A., Optimale Centerverteilung bei Interpolation mit radialen Basisfunktionen, Diplom Thesis, Universität Göttingen, 1994.
9. Bloom, T., The Lebesgue Constant for Lagrange Interpolation in the Simplex, J. Approx. Theory **54** (1988), 338–353.
10. Bojanov, B., and H. Woźniakowski, Optimal Recovery, Nova Science Publishers, New York, 1992.
11. Boor, C. de, R.A. DeVore, and A. Ron, Approximation from Shift–invariant Subspaces of $L_2(\mathbb{R}^d)$, Trans. Amer. Math. Soc. **341** (1994), 787–806.
12. Boor, C. de, R.A. DeVore, and A. Ron, The Structure of Finitely Generated Shift–invariant Spaces in $L_2(\mathbb{R}^d)$ J. Funct. Anal. **119** (1994), 37–78.
13. Boor, C. de, and A. Ron, Fourier Analysis of Approximation Orders from Principal Shift–invariant Spaces, Constr. Approx. **8** (1992), 427–462.
14. Bos, L.P., Bounding the Lebesgue Function for Lagrange Interpolation in a Simplex, J. Approx. Theory **38** (1983), 43–59.
15. Buhmann, M.D., New Developments in the Theory of Radial Basis Function Interpolation, in *Multivariate Approximations: From CAGD to Wavelets*, K. Jetter and F.I. Utreras (eds), World Scientific, London, 1993, 35–75.
16. Butzer, P.L., and R.L. Stens, Sampling Theory for Not Necessarily Band–limited Functions: A Historical Overview, SIAM Review **34** (1992), 40–53.
17. Calderón, A.P., Lebesgue Spaces of Functions and Distributions, Proc. Symp. Pure Math. **4**, AMS, Providence (1961), 33–49.
18. Calderón, A.P., Intermediate Spaces and Interpolation, the Complex Method, Studia Math. **24** (1964), 113–190.
19. Cheney, E.W., Survey Talk (this volume).
20. Duchon, J., Splines Minimizing Rotation–invariate Semi–norms in Sobolev Spaces, in *Constructive Theory of Functions of Several Variables*, Springer, Berlin–Heidelberg, W. Schempp and K. Zeller (eds.), 1979, 85–100.
21. Dyn, N., Interpolation of Scattered Data by Radial Functions, in *Topics in Multivariate Approximation*, C.K. Chui and L.L. Schumaker and Florencio Utreras (eds.), Academic Press, Boston, 1987, 47-61.

22. Dyn, N., Interpolation and Approximation by Radial and Related Functions, in *Approximation Theory VI, Vol. 1*, C.K. Chui and L.L. Schumaker and J.D. Ward (eds.) Academic Press, Boston, 1989, 211–234.
23. Fang, C. K., Strictly Positive Definite Functions, Preprint 1994.
24. Golomb, M., and H.F. Weinberger, Optimal Approximation and Error Bounds, in *On Numerical Approximation*, The University of Wisconsin Press, Madison, R.E. Langer (ed.), 1959.
25. Iske, A., Charakterisierung bedingt positiv definiter Funktionen für multivariate Interpolationsmethoden mit radialen Basisfunktionen, Ph. D. Dissertation, Göttingen, 1994.
26. Iske, A., Characterization of Function Spaces associated to Conditionally Positive Definite Functions, in *Mathematical Methods in Computer Aided Geometric Design III*, T. Lyche and L.L. Schumaker (ed.), Academic Press, New York, 1995.
27. Jetter, K., Multivariate Approximation from the Cardinal Interpolation Point of View, in *Approximation Theory VII*, E.W. Cheney and C.K. Chui and L.L. Schumaker (eds.), Academic Press, New York, 1992, 131–161.
28. Krein, M., Sur le problème du prolongement des fonctions hermitiennes positives et continues, Dokl. Akad. Nauk SSSR **26** (1940), 17–22.
29. Laurent, P.J., Construction of Spline Functions in a Convex Set, in *Approximation with special emphasis on Spline Functions*, I.J. Schoenberg (ed.), Academic Press, New York, 1969, 415–446.
30. Light, W., Using Radial Functions on Compact Domains, in *Wavelets, Images, and Surface Fitting*, P.–J. Laurent, A. Le Méhauté and L.L. Schumaker, (eds.), AKPeters, Boston, 1994, 351–370.
31. Madych, W.R., and S.A. Nelson, Multivariate Interpolation: a Variational Theory, Manuscript, 1983.
32. Madych, W.R., and S.A. Nelson, Multivariate Interpolation and Conditionally Positive Definite Functions, Approx. Theory Appl. **4** (1988), 77-89.
33. Madych, W.R., and S.A. Nelson, Multivariate Interpolation and Conditionally Positive Definite Functions II, Math. Comp. **54** (1990), 211–230.
34. Madych, W.R., and S.A. Nelson, Bounds on Multivariate Polynomials and Exponential Error Estimates for Multiquadric Interpolation, J. Approx. Theory **70** (1992), 94–114.
35. Maz'ya, V., and G. Schmidt, On Approximate Approximations, Preprint LITH–MAT–R–94–12, Linköping Univ., Dept. of Math., Linköping 1994.
36. Maz'ya, V., and G. Schmidt, On Approximate Approximations Using Gaussian Kernels, WIAS Preprint 111, Weierstraß–Institut Berlin,

1994.

37. Meinguet, J., An Intrinsic Approach to Multivariate Spline Interpolation at Arbitrary Points, in *Polynomial and Spline Approximation*, Badri N. Sohnrey (ed.), D. Reidel Publishing Company, 1979, 163–190.
38. Micchelli, C.A., and T.J. Rivlin, Lectures on Optimal Recovery, in *Numerical Analysis*, Lancaster 1984, P.R. Turner (ed.), Springer Verlag, Lecture Notes in Mathematics 1129, 1984, 12–93.
39. Micchelli, C.A., Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Definite Functions, Constr. Approx. **2** (1986), 11–22.
40. Narcowich, F.J., N. Sivakumar, and J.D. Ward, On Condition Numbers Associated with Radial–Function Interpolation, Techn. Report No. 287, Department of Mathematics, Texas A & M University, 1993.
41. Narcowich, F.J., and J.D. Ward, Norm Estimates for Inverses of Scattered–data Interpolation Matrices Associated with Completely Monotonic Radial Functions, Preprint.
42. Narcowich, F.J., N. Sivakumar, and J.D. Ward, Norm of Inverses and Condition Numbers for Matrices Associated with Scattered Data, J. Approx. Theory **64**, 1991, 69–94.
43. Narcowich, F.J., J.D. Ward, Norms of Inverses for Matrices Associated with Scattered Data, in *Curves and Surfaces*, P.J. Laurent and A. Le Méhauté and L.L. Schumaker (eds.) Academic Press, Boston, 1991, 341–348.
44. Narcowich, F.J., J.D. Ward, Generalized Hermite Interpolation Via Matrix–valued Conditionally Positive Definite Functions, Math. Comp. **63** (1994), 661–687.
45. Narcowich, F.J., J.D. Ward, Norm Estimates for the Inverses of a General Class of Scattered–Data Radial–Function Interpolation Matrices, J. Approx. Theory **69** (1992), 84–109.
46. Pisier, G., Survey Talk on Interpolation (this volume)
47. Powell, M. J. D., Radial Basis Functions for Multivariable Interpolation: a Review, in *Numerical Analysis*, D. F. Griffiths and G. A. Watson (eds.), Longman Scientific & Technical (Harlow), 1987, 223-241.
48. Powell, M.J.D., The Theory of Radial Basis Function Approximation in 1990, in *Advances in Numerical Analysis II: Wavelets, Subdivision Algorithms, and Radial Basis Functions*, W.A. Light (ed.), Oxford Univ. Press, Oxford, 1992, 105-210.
49. Powell, M.J.D., The Uniform Convergence of Thin–plate Spline Interpolation in Two Dimensions, DAMTP Report 1993/NA16, Cambridge, U.K.
50. Rudin, W., The Extension Problem for Positive Definite Functions, Illinois J. Math. **7**(1963), 532–539.

51. Rudin, W., An Extension Theorem for Positive Definite Functions, Duke Math. J. **37** (1970), 49–53.
52. Sard, A., Linear Approximation, Mathematical Surveys **9**, AMS, Providence (1963).
53. Schaback, R., Comparison of Radial Basis Function Interpolants, in *Multivariate Approximations: From CAGD to Wavelets*, K. Jetter and F. Utreras, (eds), World Scientific, London, 1993, 293–305.
54. Schaback, R., Lower Bounds for Norms of Inverses of Interpolation Matrices for Radial Basis Functions, J. Approx. Theory **79** (1994), 287–306.
55. Schaback, R., Reproduction of Polynomials by Radial Basis Functions, *Wavelets, Images, and Surface Fitting*, P.–J. Laurent, A. Le Méhauté and L.L. Schumaker (eds.) AKPeters, Boston (1994) 459–466.
56. Schaback, R., Error Estimates and Condition Numbers for Radial Basis Function Interpolation, Advances in Computational Mathematics **3** (1995), 251–264.
57. Schaback, R., Approximation by Radial Basis Functions with Finitely Many Centers, Constructive Approximation, to appear in *Constr. Approx.* 1995.
58. Schaback, R., Creating Surfaces from Scattered Data Using Radial Basis Functions, in *Mathematical Methods in Computer Aided Geometric Design III*, T. Lyche and L.L. Schumaker (ed.), Academic Press, New York, 1995.
59. Schaback, R., Z. Wu, Operators on Radial Functions, Preprint Göttingen 1994.
60. Schechter, M., Complex Interpolation, Compositio Math. **18** (1967), 117–147.
61. Stewart, J., Positive Definite Functions and Generalizations, an Historical Survey, Rocky Mountain J. Math. **6** (1976), 409–434.
62. Sun, X., Conditionally Positive Definite Functions and Their Application to Multivariate Interpolations, J. Approx. Theory **74** (1993), 159–180.
63. Traub, J.F., and H. Woźniakowski, *A General Theory of Optimal Algorithms*, Academic Press, New York, London, Toronto, Sydney, San Francisco, Elsevier, New York, 1980.
64. Wu, Z., Multivariate Compactly Supported Positive Definite Radial Functions, Preprint, Göttingen, 1994. To appear in AICM.
65. Wu, Z., Characterization of Positive Definite Radial Functions, Preprint, Göttingen, 1994.
66. Wu, Zong-min, R. Schaback, Local Error Estimates for Radial Basis Function Interpolation of Scattered Data, IMA J. Numer. Anal. **13** (1993), 13–27.

Robert Schaback
Universität Göttingen, Lotzestraße 16-18, D-37083 Göttingen, Germany
schaback@namu01.gwdg.de