Chapter 2
The Spherical Basis Function Method

2.1 Introduction

Since the early 1990s the radial basis function (RBF) method has become a well established tool for reconstructing functions and for solving partial differential equations based on data prescribed at scattered locations throughout a domain in $\mathbb{R}^d$. The method is extremely flexible, it has good approximation properties and does not become more elaborate as the space dimension increases. In the past decade or so approximation on spheres has become an area of growing interest with applications to physical geodesy, potential theory, geophysics, oceanography and meteorology. As more satellites are launched into space, the acquisition of global data is becoming more widespread and the demand for spherical data processing and solving problems of a global nature is increasing. In this chapter we will introduce the spherical basis function (SBF) method; the spherical analogue of the famous RBF method. In particular, we will use tools from Chap. 1 to construct a theoretical framework within which we can analyse the accuracy of the method. Specifically, we present two point-wise error bounds which both rely on the remarkable fact that, provided the data locations fill up the sphere sufficiently well, then it is possible to annihilate spherical harmonics of a certain order by using only a linear combination of point evaluations. The first error bound we encounter uses a global annihilation, i.e., every data location is used in the linear combination of point evaluations. In this case the relationship between the density of the data locations and the order of spherical harmonics to be annihilated is explicit and this is crucial to the error analysis that follows. The second error bound delivers a result of the same strength but uses a local annihilation of spherical harmonics and from this perspective one may consider this to be a stronger result. A drawback of this local approach is that, unlike the global case, there is no explicit connection between the density of the points and the degree of spherical harmonics to be annihilated and, as a result, much more additional work is required to establish the final result. The payoff for putting in this extra effort will be realized in the next chapter where the local error bound plays a key rôle in making significant improvements in the error bounds.

S. Hubbert et al., *Spherical Radial Basis Functions, Theory and Applications*,
SpringerBriefs in Mathematics, DOI 10.1007/978-3-319-17939-1_2

2.2 A Brief History of the RBF Method

There already exists several excellent textbooks on the subject of radial basis functions [Fass07, Wen05, Buh03]. In view of this, the aim of this section is simply to establish the idea behind the method and also to highlight some of the crucial ideas and pioneering discoveries. This will set the scene for the rest of the chapter where we will demonstrate how these ideas can be recast onto the spherical setting.

The problem of interpolating data measured at scattered locations in Euclidean space $\mathbb{R}^d$ $(d > 1)$ arises in many areas of science and engineering. The importance of this problem is reflected in the literature, where a large number of different methods for its solution have been proposed. The problem itself is stated as follows.

Problem 1 Given a set $X = \{\mathbf{x}_i\}_{i=1}^N$ of distinct data points in $\mathbb{R}^d$ and a target function $f : \mathbb{R}^d \to \mathbb{R}$, find a function $s : \mathbb{R}^d \to \mathbb{R}$ that satisfies the interpolation conditions

$$s(\mathbf{x}_i) = f(\mathbf{x}_i), \quad 1 \le i \le N. \tag{2.1}$$

The *radial basis function* (RBF) approach proposes a solution of the form

$$s(\mathbf{x}) = \sum_{j=1}^{N} \lambda_j \phi(d(\mathbf{x}, \mathbf{x}_j)), \quad \text{for} \ \ \lambda_j \in \mathbb{R}, \ \ 1 \le j \le N, \tag{2.2}$$

where $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ is usually the Euclidean metric and $\phi : [0, \infty) \to \mathbb{R}$ is the RBF. Applying the interpolation conditions (2.1) provides the linear system

$$A_\phi \boldsymbol{\lambda} = f, \quad \text{where} \ \ A_\phi \in \mathbb{R}^{N\times N} : \ \ A_{\phi,ij} = \phi(d(\mathbf{x}_i, \mathbf{x}_j)), \quad 1 \le i, j \le N. \tag{2.3}$$

Thus a unique RBF interpolant exists if and only if the interpolation matrix A_ϕ is non-singular. One of the most attractive features of the RBF method is the fact that a unique interpolant is often guaranteed under rather mild conditions on the data points. In particular, if we choose our basis function to be any one of the following

$$\left.\begin{aligned} \phi(r) &= e^{-cr^2} && \text{(Gaussian)},\\ \phi(r) &= (r^2 + c^2)^{-\frac{1}{2}} && \text{(inverse multiquadric)},\\ \phi(r) &= (r^2 + c^2)^{\frac{1}{2}} && \text{(multiquadric)},\\ \phi(r) &= r && \text{(linear)}, \end{aligned}\right\} \quad r \ge 0, \quad c > 0, \tag{2.4}$$

then uniqueness is guaranteed provided that the N data points $(N \ge 2)$ are distinct, which is as simple a condition as one could wish for.

Two other commonly used RBFs are

$$\begin{aligned} \phi(r) &= r^2 \log r \quad \text{(thin plate spline)}, \\ \phi(r) &= r^3 \qquad\qquad\qquad \text{(cubic)}, \end{aligned} \quad r \geq 0. \tag{2.5}$$

Now, in contrast to the previous candidates (2.4), there is no guarantee that the resulting interpolation matrix A_ϕ will be non-singular. Indeed, for the thin plate spline we can choose $\mathbf{x}_2, \ldots, \mathbf{x}_N$ to be any distinct points on the unit sphere centred at $\mathbf{x}_1$, in which case the first row and column of A_ϕ consists entirely of zeros, and hence is singular. In such cases, it is usual to add to s a polynomial of degree $k \geq 1$, and so consider an interpolant of the form

$$s(\mathbf{x}) = \sum_{j=1}^{N} \lambda_j \phi(d(\mathbf{x}, \mathbf{x}_j)) + \sum_{j=1}^{M} \mu_j p_j(\mathbf{x}), \tag{2.6}$$

where $M = \dim \Pi_k(\mathbb{R}^d)$ and $\{p_1, \ldots, p_M\}$ is a basis for $\Pi_k(\mathbb{R}^d)$. If we then impose the usual interpolation conditions

$$\sum_{j=1}^{N} \lambda_j \phi(d(\mathbf{x}_i, \mathbf{x}_j)) + \sum_{j=1}^{M} \mu_j p_j(\mathbf{x}_i) = f(\mathbf{x}_i), \quad 1 \leq i \leq N, \tag{2.7}$$

we observe that the addition of the polynomial introduces an extra M degrees of freedom. These are usually taken up by insisting that the RBF coefficients satisfy the following moment conditions:

$$\sum_{j=1}^{N} \lambda_j p_i(\mathbf{x}_j) = 0, \quad 1 \leq i \leq M, \tag{2.8}$$

or equivalently, we have the linear system

$$\begin{pmatrix} A_\phi & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{0} \end{pmatrix} \tag{2.9}$$

where A_ϕ is as in (2.3) and where $P \in \mathbb{R}^{N \times M}$ is given by

$$P_{ij} = p_j(\mathbf{x}_i), \quad \text{for } 1 \leq i \leq N \text{ and } 1 \leq j \leq M. \tag{2.10}$$

Thus a unique augmented RBF interpolant exists if and only if the augmented interpolation matrix (2.9) is non-singular.

One of the earliest examples of an RBF interpolant dates back to the late 1960s when the cubic spline method was developed for interpolating a univariate function $f : [x_1, x_N] \to \mathbb{R}$ at distinct data points $x_1 < \cdots < x_N$. The resulting interpolant s is composed of cubic polynomial pieces, that are joined so that the second derivative of s is continuous. A good account of the cubic spline method is given in [Pow81] where it is shown that in order to guarantee the uniqueness of the interpolant, it is necessary to impose suitable end conditions at x_1 and x_N. One useful condition is to set

$$s''(x_1) = s''(x_N) = 0, \tag{2.11}$$

in which case s is called the "natural" cubic-spline and has the form

$$s(x) = \sum_{j=1}^{N} \lambda_j |x - x_j|^3 + a + bx, \quad x \in \mathbb{R}. \tag{2.12}$$

We remark that the end conditions (2.11) are equivalent to the moment conditions $\sum \lambda_j = \sum \lambda_j x_j = 0$ which appear in (2.8), and thus we may regard the natural cubic-spline method as a special case of univariate RBF interpolation with $\phi(r) = r^3$. The natural cubic spline is a good starting point because the method itself has several interesting theoretical properties. In particular, if we consider the following function space

$$H = \{f \in L_2(\mathbb{R}) : |f| := \left(\int_{\mathbb{R}} |f''(x)|^2 dx\right)^{\frac{1}{2}} < \infty\}, \tag{2.13}$$

then $|\cdot|$ is a semi-norm with null space $\Pi_1(\mathbb{R})$ making H a semi-Hilbert space. Further, it is well known that (2.12) is the unique solution to the following variational problem

$$\text{minimise}\, \{|s| : s \in H \;\text{ and }\; s(x_i) = f(x_i) \;\; 1 \le i \le N\}. \tag{2.14}$$

For a detailed account of the many aspects of spline interpolation see [dB78].

The next appearance of an RBF came in 1971 when Rolland Hardy, a geoscientist, first suggested the use of the multiquadric basis function (see (2.4)) to interpolate scattered data in the plane. The discovery of this function arose from a purely heuristic approach to a problem in topography and the success of the resulting interpolation scheme, for solving 2D contour and mapping problems, is reported in [Ha71]. This appears to be the first application of the RBF method beyond the univariate setting.

The next landmark discovery occurred in 1977 when Jean Duchon [Du77] approached the data fitting problem from the variational perspective. Duchon was one of the first mathematicians to generalise the notion of a natural cubic spline to higher dimensions and, to illustrate his contribution, consider a non-negative integer $m > d/2$ which indexes the following space of functions

$$H_{m,d} = \{f \in L_2(\mathbb{R}^d) : D^{\boldsymbol{\alpha}} f \in L_2(\mathbb{R}^d), \;\text{ for all } |\boldsymbol{\alpha}| = m\}. \tag{2.15}$$

This space is then equipped with the semi-norm

$$|f|_{m,d} = \left(\sum_{|\alpha|=m} \frac{m!}{\alpha!} \int_{\mathbb{R}^d} |(D^\alpha f)(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2}, \tag{2.16}$$

with null space $\Pi_{k-1}(\mathbb{R}^d)$.

We note that, for $d = 1$ and $m = 2$, this space is the same as H above. More generally, $H_{m,d}$ is closely related to the Sobolev space $W^m(\mathbb{R}^d)$ and shares many of its properties. In particular, since $m > d/2$, we know from the Sobolev embedding theorem that $H_{m,d}$ is a semi-Hilbert space of continuous functions. Thus, following the cubic spline approach, for any $f \in H_{m,d}$ the following variational problem was considered:

$$\text{minimise}\{|s|_{m,d} : s \in H_{m,d} \text{ and } s(\mathbf{x}_i) = f(\mathbf{x}_i) \ 1 \le i \le N\}. \tag{2.17}$$

Using sophisticated techniques from distribution theory Duchon showed that the solution to (2.17), which he termed the $D^m(\mathbb{R}^d)-$spline, has the form of an augmented RBF (2.6), where

$$\phi(r) := \begin{cases} (-1)^n r^{2m-d} \log r \ \ (\text{with } n = m - \frac{d-2}{2}), & \text{if } d \text{ is even}, \\ (-1)^n r^{2m-d} \quad (\text{with } n = m - \frac{d-1}{2}), & \text{if } d \text{ is odd}, \end{cases} \tag{2.18}$$

and where the augmented polynomial is of degree $n-1$. We note again that for $d = 1$ and $m = 2$ we recover the natural cubic spline.

A year later Duchon [Du78] presented a study of the accuracy of $D^m(\mathbb{R}^d)-$spline interpolation. To set the scene, it is assumed that we wish to interpolate a function $f \in H_{m,d}$ over a set of distinct data points $X = \{\mathbf{x}_i\}_{i=1}^N$, located in a smooth, bounded domain $\Omega \subset \mathbb{R}^d$. The density of the set $X \subset \Omega$ is measured by using the mesh-norm

$$h := h(X, \Omega) := \sup_{\mathbf{y}\in\Omega} \min_{\mathbf{x}_i \in X} \|\mathbf{y} - \mathbf{x}_i\|, \tag{2.19}$$

and our aim is to investigate how the $D^m(\mathbb{R}^d)-$spline interpolant s_f approximates f as the data points become dense in Ω, that is, as $h \to 0$.

Definition 2.1 (*The Duchon strategy*) The Duchon strategy for delivering error bounds for RBF interpolation consists of the following steps:

1. construct a scalable quasi-uniform mesh for the domain, that is a collection of points on Ω so that it can be covered by a union of small open balls B_i (centred at each of the mesh points) that have uniformly bounded overlap:
2. estimate the local interpolation error using data prescribed on each B_i;
3. by way of a suitable extension operator create the *gluing result* which combines the local error estimates to provide a final estimate for Ω.

The justification for each of the three steps above is provided in [Du78], where geometric arguments and techniques from Sobolev space theory play a prominent role. Moreover, employing the strategy yields error bounds of the form

$$\|s_f - f\|_{L_p(\Omega)} = O(h^{m-\frac{d}{2}+\frac{d}{p}}), \quad p \in [2, \infty]. \tag{2.20}$$

In 1982, Richard Franke [F82] published the results of his survey on scattered data interpolation methods for $\mathbb{R}^2$. In his report, over 30 different methods are tested including Hardy's multiquadric and Duchon's thin plate spline. Each method was assessed over a range of criterion including accuracy, ease of implementation and visual smoothness. Of all the methods tested Hardy's multiquadric scheme performed the best and Duchon's thin plate spline was also highly rated. These findings were particularly intriguing since, at the time, there was no mathematical basis to justify the use of multiquadric interpolation. In view of this Franke proposed the conjecture that the interpolation matrix (2.3) corresponding to the multiquadric basis function is non-singular.

The invertibility of the interpolation matrices associated with the common RBFs was proven in two stages. First, Schoenberg [Sch38] in 1938 proved the unique solvability of (2.3) for a small class of RBFs. Then, in 1986, Micchelli [Mi86] extended Schoenberg's result and established a larger class of RBFs for which (2.9) is uniquely solvable. In particular, Micchelli showed how this extension could be used to settle Franke's conjecture on the multiquadric basis function. For the convenience of the reader, we present a brief account of the Schoenberg-Micchelli theory.

Definition 2.2 (*Positive Definite Functions*) A continuous function $\phi : [0, \infty) \to \mathbb{R}$ is said to be positive definite ($\phi \in PD$) if, for any $d \geq 1$ and any set $X = \{\mathbf{x}_i\}_{i=1}^N$ of distinct points in $\mathbb{R}^d$, the quadratic form

$$\boldsymbol{\lambda}^T A_\phi \boldsymbol{\lambda} = \sum_{j=1}^{N} \sum_{k=1}^{N} \lambda_j \lambda_k \phi(d(\mathbf{x}_j, \mathbf{x}_k)) \tag{2.21}$$

is non-negative for all $\boldsymbol{\lambda} \in \mathbb{R}^N$. Furthermore, if (2.21) is positive for all $\boldsymbol{\lambda} \in \mathbb{R}^N \backslash \{\mathbf{0}\}$ then we say that ϕ is strictly positive definite ($\phi \in SPD$).

In addition, we consider the following interesting class of functions first studied by Bernstein in the early 1930s.

Definition 2.3 (*Completely Monotone Functions*) A continuous function $f : [0, \infty) \to \mathbb{R}$ is said to be completely monotone on $(0, \infty)$ if $f \in C^\infty(0, \infty)$ and

$$(-1)^l f^{(l)}(r) \geq 0, \quad \text{for all } r > 0 \text{ and } l = 0, 1, 2 \ldots. \tag{2.22}$$

In [Sch38] Schoenberg provided the following important theorem

Theorem 2.1 (Schoenberg) *A continuous function* $\phi : [0, \infty) \to \mathbb{R}$, *belongs to PD if and only if the function* $f = \phi(\sqrt{\cdot})$ *is completely monotone on* $(0, \infty)$. *Moreover, if, in addition,* f *is not a constant then* ϕ *belongs to SPD*.

Using Schoenberg's theorem we can immediately deduce that the Gaussian and inverse multiquadric basis functions belong to *SPD*. This observation establishes the solvability of the RBF method for these cases.

Following on from Schoenberg's work we consider the following, more general, class of functions.

Definition 2.4 (*Conditionally Positive Definite Functions*) A continuous function $\phi : [0, \infty) \to \mathbb{R}$ is said to be conditionally positive definite of order $m \in \mathbb{N}$ ($\phi \in CPD(m)$) if, for any $d \geq 1$ and any set $X = \{\mathbf{x}_i\}_{i=1}^N$ of distinct points in $\mathbb{R}^d$, the quadratic form (2.21) is non-negative on the subspace

$$V_{m-1} = \{\boldsymbol{\lambda} \in \mathbb{R}^N : \sum_{i=1}^{N} \lambda_i p(\mathbf{x}_i) = 0 \ \text{for all} \ p \in \Pi_{m-1}(\mathbb{R}^d)\}. \tag{2.23}$$

Furthermore, if (2.21) is positive for all $\boldsymbol{\lambda} \in V_{m-1} \backslash \{\mathbf{0}\}$ then we say that ϕ is conditionally strictly positive definite of order m ($\phi \in CSPD(m)$).

For augmented RBF interpolation (2.6) it is usual to insist that the geometry of the locations satisfy the following mild property.

Definition 2.5 (*Unisolvency*) Let m be a positive integer and let $M = \dim \Pi_{m-1}(\mathbb{R}^d)$. A set of distinct points $\{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ is said to be unisolvent with respect to $\Pi_{m-1}(\mathbb{R}^d)$ if the only element of $\Pi_{m-1}(\mathbb{R}^d)$ to vanish at each $\mathbf{x}_i$ is the zero polynomial.

The following theorem establishes a unique solution to the augmented interpolation problem (2.9) in the case where $\phi \in CSPD(m)$.

Theorem 2.2 *Let* $\phi \in CSPD(m)$ *and* $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ *denote a set of N distinct data points in* $\mathbb{R}^d$ *such that*

(i) $N \geq M = \dim \Pi_{m-1}(\mathbb{R}^d)$,
(ii) *X contains a subset that is unisolvent with respect to* $\Pi_{m-1}(\mathbb{R}^d)$.

Then the augmented interpolation problem (2.9) *has a unique solution.*

Proof It is sufficient to show that, if $\boldsymbol{\lambda} \in \mathbb{R}^N$ and $\boldsymbol{\mu} \in \mathbb{R}^M$ satisfy the homogeneous linear system

$$\begin{pmatrix} A_\phi & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{2.24}$$

then they are zero vectors. We note that $P^T\boldsymbol{\lambda} = 0$ implies that $\boldsymbol{\lambda} \in V_{m-1}$, and thus

$$\mathbf{0} = \boldsymbol{\lambda}^T(A_\phi\boldsymbol{\lambda} + P\boldsymbol{\mu}) = \boldsymbol{\lambda}^T A_\phi\boldsymbol{\lambda} + (\boldsymbol{\mu}^T(P^T\boldsymbol{\lambda}))^T = \boldsymbol{\lambda}^T A_\phi\boldsymbol{\lambda}.$$

Since A_ϕ induces a positive definite form on V_{m-1}, this implies that $\boldsymbol{\lambda} = \mathbf{0}$.

Let $\{p_1, \dots, p_M\}$ denote the basis of $\Pi_{m-1}(\mathbb{R}^d)$ used to define the matrix P (2.10) and let $p^*(\mathbf{x}) = \sum_{j=1}^M \mu_j p_j(\mathbf{x})$. Now, since $\boldsymbol{\lambda} = \mathbf{0}$ we have that

$$P\boldsymbol{\mu} = (p^*(\mathbf{x}_1), \dots, p^*(\mathbf{x}_N))^T = \mathbf{0},$$

and so, using the unisolvency of X, we conclude that $\boldsymbol{\mu} = \mathbf{0}$. $\square$

Inspired by Schoenberg's characterisation theorem, Micchelli proved the following important extension.

Theorem 2.3 (Micchelli) *A continuous function* $\phi : [0, \infty) \to \mathbb{R}$, *belongs to* $CPD(m)$ *if the function* $f = \phi(\sqrt{\cdot})$ *is such that* $(-1)^m f^{(m)}$ *is completely monotone on* $(0, \infty)$. *Moreover, if, in addition,* $f \in C^{m-1}[0, \infty)$ *and is not a polynomial of degree at most* m, *then* ϕ *belongs to* $CSPD(m)$.

We note in passing that, as Micchelli suspected, the converse of this theorem is also true and this was settled in 1993 by Guo et al. [GHS93]. As it stands, Micchelli's theorem serves as an important source of applicable RBFs. The popular choices are the generalised Duchon splines

$$\begin{aligned} &\phi(r) = (-1)^{k+1} r^{2k}\log r && \in CSPD(k+1), && k \in \mathbb{N}, \\ &\phi(r) = (-1)^{\lfloor\beta\rfloor} r^{2\beta} && \in CSPD(\lfloor\beta\rfloor + 1), && \beta > 0 \text{ and } \beta \notin \mathbb{N}, \end{aligned} \tag{2.25}$$

and the generalised multiquadrics

$$\begin{aligned} &\phi(r) = (-1)^{\lfloor\beta\rfloor+1}(r^2 + c^2)^\beta && \in CSPD(\lfloor\beta\rfloor + 1), && \beta > 0 \text{ and } \beta \notin \mathbb{N}, \\ &\phi(r) = (r^2 + c^2)^\beta && \in SPD, && \beta < 0. \end{aligned} \tag{2.26}$$

In addition, Micchelli also proved the following important theorem concerning $CSPD(1)$ functions.

Theorem 2.4 *Let* $-\phi$ *be* $CSPD(1)$ *with* $\phi(0) \geq 0$ *then the corresponding interpolation matrix* A_ϕ *given by* (2.3) *is non-singular.*

Proof By definition the matrix A_ϕ induces a positive definite form on the $N-1$ dimensional hyperplane, given by

$$V_0 = \{\boldsymbol{\lambda} = (\lambda_1, \dots \lambda_N)^T \in \mathbb{R}^N : \sum_{i=1}^N \lambda_i = 0\}.$$

Thus A_ϕ has at least $N-1$ positive eigenvalues. However trace$(A_\phi) \leq 0$, and so the remaining eigenvalue must therefore be negative. $\square$

In view of Franke's numerical findings this theorem is particularly important, for it establishes the solvability of Hardy's original multiquadric interpolation scheme. Indeed, Micchelli's overall contribution has encouraged a large communities of mathematicians to study the properties of RBFs.

One of the most important of the post-Micchelli discoveries is the so-called variational approach, i.e., that every RBF interpolant can be viewed as the solution to a minimal norm interpolation problem in some reproducing kernel Hilbert space (commonly called the Native space). This approach, which can be viewed as a generalization of Duchon's work, was developed by Madych and Nelson in the early 1980 s and finally published in 1990 [MN90]. Over the years many researchers have studied the original Madych-Nelson approach and, as a result, a sound theoretical framework for RBF interpolation has emerged where error estimates can easily be delivered. To give a flavour of the Madych-Nelson theory, we consider the following definition.

Definition 2.6 Let $\Phi \in C(\mathbb{R}^d)$ be of polynomial growth, i.e., there exists $k \in \mathbb{N}_0$, such that $|\Phi(\mathbf{x})| = O(\|\mathbf{x}\|^k)$ as $\|\mathbf{x}\| \to \infty$. A continuous function $\widehat{\Phi} : \mathbb{R}^d \backslash \{\mathbf{0}\} \to \mathbb{R}$ is said to be the generalised Fourier transform of Φ if there exists $m \in \mathbb{N}_0$, such that

$$\int_{\mathbb{R}^d} \Phi(\mathbf{x})\widehat{\gamma}(\mathbf{x})d\mathbf{x} = \int_{\mathbb{R}^d} \widehat{\Phi}(\boldsymbol{\omega})\gamma(\boldsymbol{\omega})d\boldsymbol{\omega} \tag{2.27}$$

holds for all functions γ from the subspace

$$\mathcal{S}_{m-1}(\mathbb{R}^d) = \{\gamma \in \mathcal{S}(\mathbb{R}^d) : \int_{\mathbb{R}^d} \gamma(\mathbf{x})p(\mathbf{x})d\mathbf{x} = 0 \text{ for all } p \in \Pi_{m-1}(\mathbb{R}^d)\}.$$

Furthermore, the minimal choice of m is called the order of $\widehat{\Phi}$.

We now quote a specialisation of a result due to Iske, which can be found in [I95].

Theorem 2.5 *Let $\phi \in C[0, \infty)$ and assume that $\Phi(\boldsymbol{x}) = \phi(\|x\|)$ is of polynomial growth, then the following are equivalent*

(i) $\phi \in CPD(m)$;

(ii) Φ possesses a generalised Fourier transform $\widehat{\Phi}$ of order m which is non-negative and not identically zero on $\mathbb{R}^d \backslash \{\boldsymbol{0}\}$.

It turns out that the generalised Fourier transforms of the most commonly used basis functions $\phi \in CSPD(m)$ are positive on $\mathbb{R}^d \backslash \{\mathbf{0}\}$ [SW01]. This fact allows us to define the so-called *native space* by

$$\mathcal{H}_\phi = \{f \in L_2(\mathbb{R}^d) : |f|_\phi^2 = \int_{\mathbb{R}^d} \frac{1}{\widehat{\Phi}(\boldsymbol{\omega})}|\widehat{f}(\boldsymbol{\omega})|^2 d\boldsymbol{\omega} < \infty\}, \tag{2.28}$$

where $|\cdot|_\phi$ is a semi-norm whose kernel is $\Pi_{m-1}(\mathbb{R}^d)$.

For a given $\phi \in CSPD(m)$, the nature of $\mathcal{H}_\phi$ is largely determined by the decay rate of $\widehat{\Phi}(\boldsymbol{\omega})$. Specifically, if $\widehat{\Phi}(\boldsymbol{\omega})$ has a polynomial rate of decay then $\mathcal{H}_\phi$ is closely related to a certain Sobolev space. On the other hand, if $\widehat{\Phi}(\boldsymbol{\omega})$ decays exponentially quickly then $\mathcal{H}_\phi$ is a smaller space of C^∞ functions. Madych and Nelson were the first to illustrate the importance of the native spaces. Specifically, they showed that given any $f \in \mathcal{H}_\phi$, the solution to the following Duchon-like variational problem

$$\text{minimise}\{|s|_\phi : s \in \mathcal{H}_\phi \text{ and } s(x_i) = f(x_i) \ \ 1 \leq i \leq N\}, \tag{2.29}$$

is precisely the unique $\phi-$based RBF interpolant.

Micchelli's non-singularity results and the Madych-Nelson variational framework are the fundamental starting points from which a whole host of theoretical and practical advances have been made; the reader is encouraged to consult the textbooks highlighted earlier to discover more.

All of the candidate RBFs that we have encountered so far are globally supported. When implementing algorithms on large data sets the global support can be a drawback; the associated dense interpolation matrices can be poorly conditioned and also the evaluation of resulting interpolants can be expensive. In the mid 1990 s several researchers set about overcoming these issues by constructing tailor made strictly positive definite RBFs that have compact support, for these examples the interpolation matrices are sparse and better conditioned and the evaluation of their interpolants is simpler as it requires relatively few evaluations of the RBF. Unlike their global counterparts the compactly supported RBFs are dimension dependent so Michelli's theorem does not apply. Instead we appeal to a famous and more general result of Bochner which tells us that a candidate RBF ϕ is strictly positive definite on $\mathbb{R}^d$ (where d is fixed) whenever the $d-$dimensional Fourier transform of its induced kernel $\Phi(\mathbf{x}) = \phi(\|\mathbf{x}\|)$ $(\mathbf{x} \in \mathbb{R}^d)$ is positive, i.e., whenever

$$\widehat{\Phi}(\boldsymbol{\omega}) > 0, \quad \text{for all } \ \boldsymbol{\omega} \in \mathbb{R}^d.$$

One of the most commonly used families of compactly supported RBFs examples are the Wendland functions (named after their discoverer). In order to introduce these functions we begin by investigating the following family of parameterised basis functions defined by:

$$\phi_{\mu,\alpha}(r) := \frac{1}{2^{\alpha-1}\Gamma(\alpha)} \int_r^1 (1-t)^\mu \, t \left(t^2 - r^2\right)^{\alpha-1} \mathrm{d}t \quad \text{for} \quad r \in [0, 1], \tag{2.30}$$

where $\mu > -1$, $\alpha > 0$. It well known (see [Gne02]) that if $\alpha = k \in \{0, 1, 2, \ldots\}$ then the function $\phi_{\mu,k}$ generates a strictly positive definite function on $\mathbb{R}^d$ if and only if $\mu \geq \frac{d+1}{2} + k$. In [Wen95] Wendland considers the case where

$$\mu = \ell := \left\lfloor \frac{d}{2} \right\rfloor + k + 1, \tag{2.31}$$

i.e., the smallest allowable integer that still allows positive definiteness. In practical cases it is usual to introduce a support parameter $\epsilon > 0$ and define

$$\phi_{\ell,k}^{(\epsilon)}(r) = \frac{1}{2^{k-1}(k-1)!}\int_{\epsilon r}^{1} (1-t)^{\ell}\, t\, \left(t^2 - (\epsilon r)^2\right)^{\alpha-1} \mathrm{d}t \quad \text{for} \quad r \in \left[0, \frac{1}{\epsilon}\right].$$

It is known that the functions $\phi_{\ell,k}^{(\epsilon)}$ are polynomial of degree $2k+\ell$ on $0 \le r \le 1/\epsilon$, and furthermore it is shown in [Hub12] that they are given explicitly by

$$\begin{aligned}\phi_{\ell,k}^{(\epsilon)}(r) = (-1)^k 2^k \ell! \Bigg[&\sum_{i=0}^{k+\lfloor \frac{\ell}{2} \rfloor} \frac{\Gamma\left(i+\frac{1}{2}\right)(\epsilon r)^{2i}}{\Gamma\left(i-k+\frac{1}{2}\right)(\ell+2k-2i)!(2i)!} \\ &- \sum_{i=0}^{\lfloor \frac{\ell-1}{2} \rfloor} \frac{(k+i)!(\epsilon r)^{2k+2i+1}}{i!(\ell-2i-1)!(2k+2i+1)!} \Bigg].\end{aligned} \tag{2.32}$$

By construction each Wendland function induces a radial kernel on the appropriate $\mathbb{R}^d$ whose $d-$dimensional Fourier transform is positive. Furthermore, it can also be shown that there exists positive constants C_1 and C_2 such that

$$C_1 \frac{\epsilon^{2k+1}}{\|\boldsymbol{\omega}\|^{d+1+2k}} \le \widehat{\Phi_{\ell,k}^{(\epsilon)}}(\boldsymbol{\omega}) \le C_2 \frac{\epsilon^{2k+1}}{\|\boldsymbol{\omega}\|^{d+1+2k}}. \tag{2.33}$$

The Madych-Nelson variational theory applies equally to the more general dimension dependent RBFs. In particular, in the case of the Wendland functions we note that the polynomial decay rates of their Fourier transforms (2.33) ensures that their corresponding Native spaces (2.28) are norm equivalent to certain Sobolev spaces.

2.3 The Spherical Basis Function Method

The global spherical interpolation problem is as follows:

Problem 2 Given a set $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ of distinct data points on S^{d-1} and a target function $f : S^{d-1} \to \mathbb{R}$, find a function $s : S^{d-1} \to \mathbb{R}$ that satisfies the interpolation conditions

$$s(\boldsymbol{\xi}_i) = f(\boldsymbol{\xi}_i), \quad 1 \le i \le N. \tag{2.34}$$

In this setting we can consider specializing the RBF method to the sphere by considering an interpolant of the form

$$s(\boldsymbol{\xi}) = \sum_{j=1}^{N} \alpha_i \psi(g(\boldsymbol{\xi}, \boldsymbol{\xi}_j)), \quad \boldsymbol{\xi} \in S^{d-1}, \tag{2.35}$$

where g denotes the geodesic metric on S^{d-1}

$$g(\boldsymbol{\xi}, \boldsymbol{\eta}) = \cos^{-1}(\boldsymbol{\xi}^T \boldsymbol{\eta}), \quad \boldsymbol{\xi}, \boldsymbol{\eta} \in S^{d-1}. \tag{2.36}$$

and where $\psi : [0, \pi] \to \mathbb{R}$ is a continuous function which we will call the spherical basis function (SBF). Applying the interpolation conditions (2.34) provides the linear system

$$A_\psi \boldsymbol{\alpha} = f, \text{ where } A_\psi \in \mathbb{R}^{N \times N} : \; A_{\psi,ij} = \psi(g(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j)), \quad 1 \leq i, j \leq N. \tag{2.37}$$

Thus a unique SBF interpolant exists if and only if the interpolation matrix A_ψ is non-singular.

Just as polynomial reproduction is important in Euclidean data fitting problems it is also common, in the spherical setting, that one requires that an interpolant should reproduce the low order spherical harmonics. Following the RBF approach we can conveniently add to s (2.35) a spherical harmonic of order k, which gives the form

$$s(\boldsymbol{\xi}) = \sum_{j=1}^{N} \alpha_j \psi(g(\boldsymbol{\xi}, \boldsymbol{\xi}_j)) + \sum_{j=1}^{M} \beta_j \mathcal{Y}_j(\boldsymbol{\xi}), \quad \boldsymbol{\xi} \in S^{d-1}, \tag{2.38}$$

where $M = \dim \mathcal{H}_k(S^{d-1})$, and $\{\mathcal{Y}_1, \ldots, \mathcal{Y}_M\}$ is a basis for $\mathcal{H}_k(S^{d-1})$.

The interpolation conditions (2.34) now provide N linear equations in $N + M$ unknowns, and so, following RBF theory, it is usual to impose M linear constraints

$$\sum_{j=1}^{N} \alpha_j \mathcal{Y}_i(\boldsymbol{\xi}_j) = 0, \quad 1 \leq i \leq M, \tag{2.39}$$

which leads to the augmented linear system

$$\begin{pmatrix} A_\psi & Y \\ Y^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{0} \end{pmatrix} \tag{2.40}$$

where A_ψ is as in (2.37) and $Y \in \mathbb{R}^{N \times M}$ is given by

$$Y_{ij} = \mathcal{Y}_j(\boldsymbol{\xi}_i), \quad \text{where } 1 \leq i \leq N, \text{ and } 1 \leq j \leq M. \tag{2.41}$$

Thus a unique augmented SBF interpolant exists if and only if the augmented interpolation matrix in (2.40) is non-singular. In order to make a transfer of the RBF machinery to the spherical setting we require the spherical analogue of Michelli's discovery, i.e., a notion and characterization of positive and conditionally positive definite functions on spheres. Recasting from the Euclidean setting we have the following definitions.

Definition 2.7 (*SPD functions on spheres*) A continuous function $\psi : [0, \pi] \to \mathbb{R}$ is said to be strictly positive definite on S^{d-1} ($\psi \in SPD(S^{d-1})$) if, for any set $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ of distinct points on S^{d-1}, the quadratic form

$$\boldsymbol{\alpha}^T A_\psi \boldsymbol{\alpha} = \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \psi(g(\boldsymbol{\xi}_j \boldsymbol{\xi}_k)) \tag{2.42}$$

is positive on $\mathbb{R}^N \backslash \{\mathbf{0}\}$.

Definition 2.8 (*CSPD functions on spheres*) Let m be a positive integer. A continuous function $\psi : [0, \pi] \to \mathbb{R}$ is said to be conditionally strictly positive definite of order m on S^{d-1} ($\psi \in CSPD_m(S^{d-1})$) if, for any set $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ of distinct points on S^{d-1}, the quadratic form (2.42) is positive on the subspace

$$W_{m-1} = \{\boldsymbol{\alpha} \in \mathbb{R}^N \backslash \{\mathbf{0} : \sum_{i=1}^N \alpha_i \mathcal{Y}(\boldsymbol{\xi}_i) = 0 \;\text{ for all }\; \mathcal{Y} \in \mathcal{H}_{m-1}(S^{d-1})\}. \tag{2.43}$$

Following the story of the RBF method it is clear that any $\psi \in SPD(S^{d-1})$ function can be used to provide a unique interpolant of the form (2.35). Furthermore, it is straightforward to show that $\psi \in CSPD_m(S^{d-1})$ can be used to provide a unique augmented interpolant (2.38) provided that the following spherical unisolvency condition holds:

Definition 2.9 (*Unisolvency on the sphere*) Let m be a positive integer and let $M = \dim \mathcal{H}_{m-1}(S^{d-1})$. A set of distinct points $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^M$ is said to be unisolvent with respect to $\mathcal{H}_{m-1}(S^{d-1})$ if the only element of $\mathcal{H}_{m-1}(S^{d-1})$ to vanish at each $\boldsymbol{\xi}_i$ is the zero spherical harmonic.

So far we have a method *in theory*. The following partial characterization theorem is the result which allows us to practically implement this method and to perform in depth analysis of its properties, it can be viewed as a modification/extension of Schoenberg's pioneering work from the early 1940s.

Theorem 2.6 *If $\psi \in CSPD_m(S^{d-1})$, then it has the following form*

$$\psi(\theta) = \sum_{k=0}^\infty a_\psi(k) P_{k,d}(\cos\theta), \tag{2.44}$$

such that

$$a_\psi(k) \geq 0 \;\; \textit{for} \;\; k \geq m \;\; \textit{and} \;\; \sum_{k=0}^\infty a_\psi(k) < \infty, \tag{2.45}$$

where $\{P_{k,d}\}$ denote the $d-$dimensional Legendre polynomials (1.24).

The above theorem gives rise to several remarks and observations. Firstly, given that $\cos(g(\boldsymbol{\xi},\boldsymbol{\eta})) = \boldsymbol{\xi}^T\boldsymbol{\eta}$ we may view our SBF ψ as a function of the inner product. Secondly, we note that SPD functions are contained in Theorem 2.6 by considering the case $m = 0$; in view of this we shall take $\psi \in CSPD_0(S^{d-1})$ to mean $\psi \in SPD(S^{d-1})$. Thirdly, and finally, we remark that a complete characterization of functions of the form (2.44) satisfying (2.45) has been established for $d \geq 3$ by Chen et al. [CMS03] who show that, in this case, a necessary and sufficient condition is that the set $\{k \in \mathbb{N}_0 \backslash \{0, 1, \dots, m-1\} : a_k > 0\}$ must contain infinitely many odd and infinitely many even integers. The case of $d = 2$ remains an open problem. For our purposes we will only consider SBFs whose Legendre coefficients satisfy the sufficient condition that they are all positive for $k \geq m$.

We now turn to developing a variational setting for SBF interpolants. This begins with an application of the addition formula (1.24) which shows that for every $\psi \in SCPD_m(S^{d-1})$ we can associate a zonal kernel $\Psi(\boldsymbol{\xi},\boldsymbol{\eta}) = \psi(\boldsymbol{\xi}^T\boldsymbol{\eta})$ which has a unique spherical Fourier expansion, given by

$$\Psi(\boldsymbol{\xi},\boldsymbol{\eta}) = \sum_{k=0}^{\infty}\sum_{\ell=1}^{N_{k,d}} \widehat{\psi}_k \mathcal{Y}_{k,\ell}(\boldsymbol{\xi})\mathcal{Y}_{k,\ell}(\boldsymbol{\eta}), \tag{2.46}$$

where $(\widehat{\psi}_k)_{k\geq 0}$ denote the spherical Fourier coefficients (1.33) of Ψ. We say that the coefficients decay at a polynomial rate as $k \to \infty$ if there exist positive constants A_1, A_2 and α independent of k such that

$$A_1(1+k)^{-(d-1+\alpha)} \leq \widehat{\psi}_k \leq A_2(1+k)^{-(d-1+\alpha)}, \qquad k \geq m, \tag{2.47}$$

otherwise they decay exponentially quickly. We remark that if (2.47) holds for the spherical Fourier coefficients then, using formula (1.33) together with (1.19), we can deduce that there exists constant $\mathcal{A}_1$ and $\mathcal{A}_2$ (again independent of k) such that

$$\mathcal{A}_1(1+k)^{-(1+\alpha)} \leq a_\psi(k) \leq \mathcal{A}_2(1+k)^{-(1+\alpha)}, \quad \alpha > 0 \quad k \geq m. \tag{2.48}$$

With this insight we define the so-called Native space for the SBF ψ as follows.

Definition 2.10 (*Native space of the SBF*) Let $\psi \in SCPD_m(S^{d-1})$ and let $\{\widehat{\psi}_k\}_{k\geq m}$ denote the spherical Fourier coefficients of its associated zonal kernel (2.46). We define the native space of ψ to be

$$\mathcal{H}_{\psi,m} := \{f \in L_2(S^{d-1}) : |f|^2_{\psi,m} = \sum_{k=m}^{\infty}\sum_{\ell=1}^{N_{k,d}} \frac{|\widehat{f}_{k,\ell}|^2}{\widehat{\psi}_k} < \infty\}, \tag{2.49}$$

where $|\cdot|_{\psi,m}$ is a (semi-)norm induced via the (semi-)inner product

$$(f, g)_{\psi,m} = \sum_{k=m}^{\infty} \sum_{\ell=1}^{N_{k,d}} \frac{\widehat{f}_{k,\ell}\widehat{g}_{k,\ell}}{\widehat{\psi}_k}. \tag{2.50}$$

We note that this definition is analogous to the RBF native spaces (2.28) where $\widehat{\Phi}(\boldsymbol{\omega})$, $\boldsymbol{\omega} \in \mathbb{R}^d$ is a counterpart to $(\widehat{\psi}_k)_{k\geq 0}$ and $\widehat{f}(\boldsymbol{\omega})$ corresponds to $\widehat{f}_{k,\ell}$. The Sobolev spaces $W^s(S^{d-1})$, $s > 0$ (see Definition 1.54) are a special instance of native spaces generated by the kernel whose Fourier coefficients are $\widehat{\psi}_k = (1 + \lambda_k)^{-s}$. One can easily show that there exists constants A_1 and A_2 such that

$$A_1(1 + k)^{-2s} \leq (1 + \lambda_k)^{-s} \leq A_2(1 + k)^{-2s}, \quad k \geq 0,$$

and, in view of this, we can see that the native space of an SBF $\psi \in SPD(S^{d-1})$ for which $(\widehat{\psi}_k)_{k\geq 0}$ (the Fourier coefficients of its induced kernel) decay at a polynomial rate (2.47) is norm equivalent to the Sobolev space $W^s(S^{d-1})$, $s = (d-1+\alpha)/2$. The Sobolev embedding theorem guarantees that this is a space of continuous functions or, in the language of Chap. 1, the pair $\left(H_{\psi,0}, (\cdot, \cdot)_{\psi,0}\right)$ is a RKHS.

When $m > 0$ the native space is a semi-Hilbert space with the spherical harmonics $\mathcal{H}_{m-1}(S^{d-1})$ being the null space of the induced semi-norm. In this case, in order to make use of the Hilbert space theory as presented in Chap. 1, it is common to modify $(f, g)_{\psi,m}$ so that it becomes a genuine inner product; we do this by defining an appropriate inner product for the null space and add this to the semi-inner product. One common approach is to select a set $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_M\}$ which is unisolvent with respect to $\mathcal{H}_{m-1}(S^{d-1})$ and use this to define the following inner product

$$\langle f, g\rangle_{\mathcal{H}_{m-1}(S^{d-1})} = \sum_{i=1}^{M} f(\boldsymbol{\xi}_i)g(\boldsymbol{\xi}_i), \quad f, g \in \mathcal{H}_{m-1}(S^{d-1}). \tag{2.51}$$

We now propose the following modified native space.

Definition 2.11 (*Native Hilbert space of the SBF*) Let $\psi \in SCPD_m(S^{d-1})$ and let $\{\widehat{\psi}_k\}_{k\geq m}$ denote the spherical Fourier coefficients of its associated zonal kernel (2.46). We define the native Hilbert space of ψ to be

$$\mathcal{H}_\psi := \left\{ f \in L_2(S^{d-1}) : \|f\|_\psi^2 = \sum_{i=1}^{M} \big(f(\boldsymbol{\xi}_i)\big)^2 + \sum_{k=m}^{\infty} \sum_{\ell=1}^{N_{k,d}} \frac{|\widehat{f}_{k,\ell}|^2}{\widehat{\psi}_k} < \infty \right\}, \tag{2.52}$$

where $\|\cdot\|_\psi$ is the norm induced via the inner product

$$\langle f, g\rangle_\psi = \langle f, g\rangle_{\mathcal{H}_{m-1}(S^{d-1})} + (f, g)_{\psi,m}. \tag{2.53}$$

We remark that as all norms are equivalent on finite dimensional spaces, we can use the same arguments as above (for the $m = 0$ case) to deduce that if the spherical Fourier coefficients exhibit the polynomial decay rate (2.47) then $\mathcal{H}_\psi$ is norm equivalent to $W^s(S^{d-1})$, with $s = (d-1+\alpha)/2$. This means both spaces coincide as sets and there exist constants $0 < k_{eq} < K_{eq}$, such that

$$k_{eq}\|\cdot\|_{W^s(S^{d-1})} \leq \|\cdot\|_\psi \leq K_{eq}\|\cdot\|_{W^s(S^{d-1})}. \tag{2.54}$$

When the coefficients decay exponentially quickly then $\mathcal{H}_\psi$ is a much smaller subspace of infinitely differentiable functions.

With this preparation we now investigate the following variational problem:

Proposition 2.1 (Optimal interpolation in the native space) *Let $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ denote a set of distinct points on S^{d-1} and $\psi \in CSPD_m(S^{d-1})$. Assume further that $\{\boldsymbol{\xi}_1, \ldots \boldsymbol{\xi}_M\} \subset \Xi$ is unisolvent with respect to $\mathcal{H}_{m-1}(S^{d-1})$, and let $\left(\mathcal{H}_\psi, \langle\cdot,\cdot\rangle_\psi\right)$ denote the native Hilbert space of ψ (2.52). Then, for any $f \in \mathcal{H}_\psi$, the solution to the variational problem:*

$$\text{minimise}\Big\{\ \|s\|_\psi : \text{ subject to } \ s \in \mathcal{H}_\psi \ \text{ and } \ s(\boldsymbol{\xi}_i) = f(\boldsymbol{\xi}_i), \quad \boldsymbol{\xi}_i \in \Xi\Big\}, \tag{2.55}$$

is the unique $\psi-$based SBF interpolant to f at Ξ.

Proof Given that $\left(\mathcal{H}_\psi, \langle\cdot,\cdot\rangle_\psi\right)$ is a Hilbert function space we know, from Sect. 1.3, that the solution has the form

$$s_f(\boldsymbol{\xi}) := \sum_{k=1}^N \lambda_k \mathtt{K}(\boldsymbol{\xi}, \boldsymbol{\xi}_k), \tag{2.56}$$

where $\mathtt{K}$ is the reproducing kernel of $\mathcal{H}_\psi$. Thus, we need to compute $\mathtt{K}$. We begin this process by defining a projection operator from $\mathcal{H}_\psi$ onto $\mathcal{H}_{m-1}(S^{d-1})$. Specifically we use the subset of Ξ that is unisolvent with respect to $\mathcal{H}_{m-1}(S^{d-1})$ to define the unique "Lagrange" basis $\{\widehat{\mathcal{Y}}_1, \ldots, \widehat{\mathcal{Y}}_M\}$ for $\mathcal{H}_{m-1}(S^{d-1})$ satisfying

$$\widehat{\mathcal{Y}}_i(\boldsymbol{\xi}_j) = \begin{cases} 1, & \text{if } \ i = j, \\ 0, & \text{if } \ i \neq j. \end{cases} \tag{2.57}$$

The "Lagrange" projection $\mathcal{P} : \mathcal{H}_\psi \to \mathcal{H}_{m-1}(S^{d-1})$ is given by

$$(\mathcal{P}f)(\boldsymbol{\xi}) = \sum_{j=1}^M \widehat{\mathcal{Y}}_j(\boldsymbol{\xi}) f(\boldsymbol{\xi}_j), \quad \text{where } \ \boldsymbol{\xi} \in S^{d-1}. \tag{2.58}$$

This operator maps each $f \in \mathcal{H}_\psi$ to its unique spherical harmonic interpolant based on $\{\boldsymbol{\xi}_j\}_{j=1}^M$. In particular, $(\mathcal{P}\mathcal{Y})(\boldsymbol{\xi}) = \mathcal{Y}(\boldsymbol{\xi})$ for all $\mathcal{Y} \in \mathcal{H}_{m-1}(S^{d-1})$, and we have the following decomposition

$$\mathcal{H}_\psi := \widetilde{H}_\psi \oplus \mathcal{H}_{m-1}(S^{d-1}), \tag{2.59}$$

where

$$\widetilde{H}_\psi = (I - \mathcal{P})H_\psi = \{f \in \mathcal{H}_\psi : f(\boldsymbol{\xi}_j) = 0 \ \text{ for } \ j = 1, \ldots, M\}. \tag{2.60}$$

We note that $(\widetilde{H}_\psi, (\cdot,\cdot)_{\psi,m})$ and $(\mathcal{H}_{m-1}(S^{d-1}), \langle\cdot,\cdot\rangle_{\mathcal{H}_{m-1}(S^{d-1})})$ are complementary orthogonal subspaces of $(H_\psi, \langle\cdot,\cdot\rangle_\psi)$ in the sense of Proposition 1.1. The reproducing kernel $\mathtt{K}_1$ of $\widetilde{H}_\psi$ was computed in [S99], and we quote

$$\mathtt{K}_1(\boldsymbol{\xi}, \boldsymbol{\eta}) := (I - \mathcal{P})_{\boldsymbol{\xi}}(I - \mathcal{P})_{\boldsymbol{\eta}}\Psi(\boldsymbol{\xi}, \boldsymbol{\eta}), \tag{2.61}$$

where Ψ is the zonal kernel associated with ψ, and where the subscript denotes the variable to which the operator applies. Furthermore, it is easily verified that

$$\mathtt{K}_2(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sum_{k=1}^{M} \widehat{\mathcal{Y}}_k(\boldsymbol{\xi})\widehat{\mathcal{Y}}_k(\boldsymbol{\eta}) \tag{2.62}$$

is the reproducing kernel of $\mathcal{H}_{m-1}(S^{d-1})$. Using Proposition 1.1, the reproducing kernel of $(\mathcal{H}_\psi, \langle\cdot,\cdot\rangle_\psi)$ is given by $\mathtt{K} = \mathtt{K}_1 + \mathtt{K}_2$, that is,

$$\mathtt{K}(\boldsymbol{\xi}, \boldsymbol{\eta}) = \Psi(\boldsymbol{\xi}, \boldsymbol{\eta}) - \sum_{j=1}^{M} \widehat{\mathcal{Y}}_j(\boldsymbol{\eta})\Psi(\boldsymbol{\xi}, \boldsymbol{\xi}_j) + \left(-\mathcal{P}_{\boldsymbol{\xi}}(I - \mathcal{P})_{\boldsymbol{\eta}}\Psi(\boldsymbol{\xi}, \boldsymbol{\eta}) + \sum_{k=1}^{M} \widehat{\mathcal{Y}}_k(\boldsymbol{\xi})\widehat{\mathcal{Y}}_k(\boldsymbol{\eta})\right).$$

We observe that the term in the brackets, as a function of $\boldsymbol{\xi}$, is simply an element of $\mathcal{H}_{m-1}(S^{d-1})$ whose coefficients depend upon $\boldsymbol{\eta}$, we denote this as $\mathcal{Y}_{\boldsymbol{\eta}}(\boldsymbol{\xi})$ and rewrite the reproducing kernel as

$$\mathtt{K}(\boldsymbol{\xi}, \boldsymbol{\eta}) = \Psi(\boldsymbol{\xi}, \boldsymbol{\eta}) - \sum_{j=1}^{M} \widehat{\mathcal{Y}}(\boldsymbol{\eta})\Psi(\boldsymbol{\xi}, \boldsymbol{\xi}_j) + \mathcal{Y}_{\boldsymbol{\eta}}(\boldsymbol{\xi}). \tag{2.63}$$

It is known that $\mathtt{K}$ belongs to $SPD(S^{d-1})$ (see [S99, Sect. 6]) and thus (2.55) has a unique solution of the form

$$s_f(\boldsymbol{\xi}) = \sum_{k=1}^{N} \lambda_k \mathtt{K}(\boldsymbol{\xi}, \boldsymbol{\xi}_k) = \sum_{k=1}^{N} \lambda_k \left(\Psi(\boldsymbol{\xi}, \boldsymbol{\xi}_k) - \sum_{j=1}^{M} \mathcal{Y}_j(\boldsymbol{\xi}_k)\Psi(\boldsymbol{\xi}, \boldsymbol{\xi}_j)\right) + \sum_{i=1}^{N} \lambda_i \mathcal{Y}_{\boldsymbol{\xi}_i}(\boldsymbol{\xi}).$$

The final sum in the above expression is an element of $\mathcal{H}_{m-1}(S^{d-1})$ which we can express in terms of the Lagrange basis. This observation, together with a little further manipulation, allows us to write

$$s_f(\boldsymbol{\xi}) = \sum_{j=1}^{N} \alpha_j \Psi(\boldsymbol{\xi}, \boldsymbol{\xi}_j) + \sum_{j=1}^{M} \beta_j \widehat{\mathcal{Y}}_j(\boldsymbol{\xi}), \tag{2.64}$$

where

$$\alpha_j = \begin{cases} \lambda_j - \sum_{k=1}^{N} \lambda_k \widehat{\mathcal{Y}}_j(\boldsymbol{\xi}_k), & \text{if } 1 \le j \le M, \\ \lambda_j, & \text{if } M+1 \le j \le N. \end{cases} \tag{2.65}$$

It is easy to check that the α_j satisfy the SBF side conditions given by (2.39), specifically let $\mathcal{Y} \in \mathcal{H}_{m-1}(S^{d-1})$ then

$$\begin{aligned} \sum_{j=1}^{N} \alpha_j \mathcal{Y}(\boldsymbol{\xi}_j) &= \sum_{j=1}^{N} \lambda_j \mathcal{Y}(\boldsymbol{\xi}_j) - \sum_{k=1}^{N} \lambda_k \sum_{j=1}^{M} \widehat{\mathcal{Y}}_j(\boldsymbol{\xi}_k) \mathcal{Y}(\boldsymbol{\xi}_j) \\ &= \sum_{j=1}^{N} \lambda_j \mathcal{Y}(\boldsymbol{\xi}_j) - \sum_{k=1}^{N} \lambda_k (\mathcal{P}\mathcal{Y})(\boldsymbol{\xi}_k) = 0. \end{aligned}$$

Thus, s_f* is precisely the unique $\psi-$based SBF interpolant to f at Ξ.

We close this section with a result that provides two important properties of the optimal SBF interpolants.

Lemma 2.1 *For a given $f \in \mathcal{H}_\psi$ let s_f denote its optimal $\psi-$based SBF interpolant, then we have*

(i) $\|f - s_f\|_\psi^2 = \langle f, f - s_f\rangle_\psi$ $\quad (ii)$ $\|f - s_f\|_\psi \le \|f\|_\psi$ $\quad (iii)$ $\|s_f\|_\psi \le \|f\|_\psi$

Proof To prove (i) consider

$$\|f - s_f\|_\psi^2 = \langle f, f - s_f\rangle_\psi - \langle s_f, f - s_f\rangle_\psi.$$

Using (2.56) and the reproducing kernel property, we have

$$\langle s_f, f - s_f\rangle_\psi = \langle \sum_{k=1}^{N} \lambda_k \mathrm{K}(\cdot, \boldsymbol{\xi}_k), f - s_f\rangle_\psi = \sum_{k=1}^{N} \lambda_k (f - s_f)(\boldsymbol{\xi}_k) = 0,$$

and thus (i) follows. To prove (ii) and (iii) we use (i) to provide:

$$\begin{aligned} \|s_f\|_\psi^2 + \|f - s_f\|_\psi^2 &= \|s_f\|_\psi^2 + \langle f, f - s_f\rangle \\ &= \|s_f\|_\psi^2 + \|f\|_\psi^2 - \langle f, s_f\rangle_\psi = \|f\|_\psi^2, \end{aligned}$$

where the final equality comes from $\langle s_f, f - s_f\rangle_\psi = 0$; inequalities (ii) and (iii) follow from this. $\square$

2.4 Framework for Pointwise Error Estimates

The undoubted appeal of the variational approach to SBF interpolation is that it provides a physical interpretation of the construction process. The surfaces generated minimize a certain energy measure (bending energy in the case of cubic and thin plate splines) and this gives a reassuring sense that they will be sensibly shaped and well-behaved. In addition, the variational approach also provides a rather nice framework for delivering error bounds and the following development show how easily it is to access such bounds. We begin by choosing $\psi \in CSPD_m(S^{d-1})$ and let s_f denote the unique $\psi-$based SBF interpolant to a given target function $f \in \mathcal{H}_\psi$. We observe that the error function $f - s_f$ belongs to the Hilbert function space $(\widetilde{H}_\psi, |\cdot|_{\psi,m})$ (2.60) and so, for any $\boldsymbol{\xi} \in S^{d-1}$, we have

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| = |(\mathtt{K}_1(\boldsymbol{\xi}, \cdot), f - s_f)_{\psi,m}|,$$

where $\mathtt{K}_1$ is the reproducing kernel of $\widetilde{H}_\psi$ given by (2.61). Applying the Cauchy-Schwarz inequality we have

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| \le |\mathtt{K}_1(\boldsymbol{\xi}, \cdot)|_{\psi,m} |f - s_f|_{\psi,m} = |\mathtt{K}_1(\boldsymbol{\xi}, \cdot)|_{\psi,m} \|f - s_f\|_\psi. \quad (2.66)$$

The factor $|\mathtt{K}_1(\boldsymbol{\xi}, \cdot)|_{\psi,m}$ is called the *Lagrange power function* for ψ, and we write $L_\psi(\boldsymbol{\xi}) = |\mathtt{K}_1(\boldsymbol{\xi}, \cdot)|_{\psi,m}$. The square of this function can be computed explicitly since

$$|\mathtt{K}_1(\boldsymbol{\xi}, \cdot)|^2_{\psi,m} = (\mathtt{K}_1(\boldsymbol{\xi}, \cdot), \mathtt{K}_1(\boldsymbol{\xi}, \cdot))_{\psi,m} = \mathtt{K}_1(\boldsymbol{\xi}, \boldsymbol{\xi}) = L^2_\psi(\boldsymbol{\xi}).$$

Thus, employing (2.61), we have

$$L^2_\psi(\boldsymbol{\xi}) := \sum_{i=1}^{M} \sum_{j=1}^{M} \widehat{\mathcal{Y}}_i(\boldsymbol{\xi}) \widehat{\mathcal{Y}}_j(\boldsymbol{\xi}) \psi(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) - 2 \sum_{i=1}^{M} \widehat{\mathcal{Y}}_i(\boldsymbol{\xi}) \psi(\boldsymbol{\xi}^T \boldsymbol{\xi}_i) + \psi(1), \quad (2.67)$$

where $\{\widehat{\mathcal{Y}}_1, \dots, \widehat{\mathcal{Y}}_M\}$ is the Lagrange basis for $\mathcal{H}_{m-1}(S^{d-1})$. This ensures spherical harmonic reproduction via

$$\mathcal{Y}(\boldsymbol{\xi}) = \sum_{i=1}^{N} \widehat{\mathcal{Y}}_i(\xi) \mathcal{Y}(\boldsymbol{\xi}_i), \quad \text{for all } \mathcal{Y} \in \mathcal{H}_{m-1}(S^{d-1}). \quad (2.68)$$

In particular, in view of (2.66) we have

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| \le L_\psi(\boldsymbol{\xi}) \|f - s_f\|_\psi, \quad \boldsymbol{\xi} \in S^{d-1}. \quad (2.69)$$

The Lagrange power function clearly provides a bound on the pointwise interpolation error. However, it only makes use of information based on the subset of $\varXi$ that is unisolvent with respect to $\mathcal{H}_{m-1}(S^{d-1})$. Intuitively, we would expect an improvement if the function were allowed to depend upon the whole of Ξ. In view of this, we fix $\boldsymbol{\xi} \in S^{d-1}$ and generalize (2.68) by selecting N real coefficients $\{\gamma_i\}_{i=1}^N$ so that

$$\mathcal{Y}(\boldsymbol{\xi}) = \sum_{i=1}^{N} \gamma_i \mathcal{Y}(\boldsymbol{\xi}_i), \quad \text{for all } \mathcal{Y} \in \mathcal{H}_{m-1}(S^{d-1}). \tag{2.70}$$

In addition, we define a bounded linear functional on $\mathcal{H}_\psi$ by

$$\Lambda_{\boldsymbol{\xi}}(f) = (\delta_{\boldsymbol{\xi}} - \sum_{i=1}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i})(f), \quad \text{for all } f \in \mathcal{H}_\psi. \tag{2.71}$$

Using Lemma 1.1, the Riesz representor of $\Lambda_{\boldsymbol{\xi}}$ in $\mathcal{H}_\psi$, is given by

$$k_{\Lambda_{\boldsymbol{\xi}}}(\cdot) = \mathtt{K}(\boldsymbol{\xi}, \cdot) - \sum_{i=1}^{N} \gamma_i \mathtt{K}(\boldsymbol{\xi}_i, \cdot), \tag{2.72}$$

where $\mathtt{K}$ is the reproducing kernel of $\mathcal{H}_\psi$, see (2.63). Now, applying the same analysis as before we find

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| = |\Lambda_{\boldsymbol{\xi}}(f - s_f)| = |\langle k_{\Lambda_{\boldsymbol{\xi}}}, f - s_f \rangle_\psi| \le \|k_{\Lambda_{\boldsymbol{\xi}}}\|_\psi \|f - s_f\|_\psi.$$

We can evoke, again, the Riesz representation theorem, to deduce that

$$\|k_{\Lambda_{\boldsymbol{\xi}}}\|_\psi = \|\Lambda_{\boldsymbol{\xi}}\|_{\psi^*} = \|\delta_{\boldsymbol{\xi}} - \sum_{i=1}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i}\|_{\psi^*}, \tag{2.73}$$

where $\|\cdot\|_{\psi^*}$ denotes the usual dual space norm given by

$$\|T\|_{\psi^*} = \sup\{|Tf| : \|f\|_\psi \le 1\}. \tag{2.74}$$

The factor $\|k_{\Lambda_{\boldsymbol{\xi}}}\|_\psi$ is said to be a *power function* for ψ at $\boldsymbol{\xi}$, and we write $P_{\psi,\gamma}(\boldsymbol{\xi}) = \|k_{\Lambda_{\boldsymbol{\xi}}}\|_\psi$. The square of this function can be computed explicitly since

$$\|k_{\Lambda_{\boldsymbol{\xi}}}\|_\psi^2 = \langle k_{\Lambda_{\boldsymbol{\xi}}}, k_{\Lambda_{\boldsymbol{\xi}}} \rangle_\psi = \Lambda_{\boldsymbol{\xi}}(k_{\Lambda_{\boldsymbol{\xi}}}) = P_{\psi,\gamma}^2(\boldsymbol{\xi}),$$

indeed this calculation was made in [LLRS99] and, again, we quote

$$P_{\psi,\gamma}(\boldsymbol{\xi}) = \left(\sum_{i=1}^{N} \sum_{j=1}^{N} \gamma_i \gamma_j \psi(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) - 2 \sum_{i=1}^{N} \gamma_i \psi(\boldsymbol{\xi}^T \boldsymbol{\xi}_i) + \psi(1) \right)^{1/2}.$$

To improve the presentation, we set $\boldsymbol{\xi}_0 = \boldsymbol{\xi}$ and $\gamma_0 = -1$, this enables us to write the power function in the more compact form

$$P_{\psi,\gamma}(\boldsymbol{\xi}) = \left(\sum_{i=0}^{N} \sum_{j=0}^{N} \gamma_i \gamma_j \psi(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) \right)^{1/2} = \left\| \sum_{i=0}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i} \right\|_{\psi^*}. \tag{2.75}$$

For a given $\boldsymbol{\xi} \in S^{d-1}$, each selection of coefficients $\{\gamma_i\}_{i=1}^N$ satisfying (2.70), gives rise to its own power function $P_{\psi,\gamma}$, which, in turn, provides the following error bound

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| \leq P_{\psi,\gamma}(\boldsymbol{\xi}) \cdot \|f - s_f\|_\psi. \tag{2.76}$$

Stated in this way, it is clear that a close investigation of $P_{\psi,\gamma}$, and especially the choice of coefficients, ought to provide an insight into the accuracy of the SBF interpolation method. In particular, in [WuS93], Wu and Schaback solve the linearly constrained optimisation problem of choosing the optimal coefficients $\{\gamma_i^*\}_{i=1}^N$ which minimises (2.75) subject to (2.70). In view of this we define

$$P_{\psi,\gamma^*}(\boldsymbol{\xi}) = \min \left\{ \left\| \sum_{i=0}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i} \right\|_{\psi^*} : \ \{\gamma_i\}_{i=1}^N \text{ satisfy condition (2.70)} \right\}, \tag{2.77}$$

to be the *optimal power function* for ψ at $\boldsymbol{\xi}$.

We remark that the error bound (2.76) may be viewed as a specific instance of the following more general result for functions with vanishing conditions.

Proposition 2.2 *Let $\psi \in CSPD_m(S^{d-1})$ and $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ denote a set of distinct points on S^{d-1}. For any $\boldsymbol{\xi} \in S^{d-1}$ we have the following bound*

$$|f(\boldsymbol{\xi})| \leq P_{\psi,\gamma^*}(\boldsymbol{\xi}) \cdot \|f\|_\psi, \ \text{ where } \ f \in \mathcal{H}_\psi \text{ and } f(\boldsymbol{\xi}_i) = 0, \ \ i = 1, \ldots, N. \tag{2.78}$$

Proof Let $\{\gamma_i\}_{i=1}^N$ denote a choice of real coefficients satisfying (2.70). Set $\boldsymbol{\xi}_0 = \boldsymbol{\xi}$ and $\gamma_0 = -1$, then, for any $f \in \mathcal{H}_\psi$ that satisfies $f(\boldsymbol{\xi}_i) = 0$, $i = 1, \ldots, N$, we have

$$|f(\boldsymbol{\xi})| = |\delta_{\boldsymbol{\xi}}(f)| = \left| \sum_{i=0}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i}(f) \right| \leq \left\| \sum_{i=0}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i} \right\|_{\psi^*} \cdot \|f\|_\psi,$$

taking the infimum over all such choices of $\{\gamma_i\}_{i=1}^N$ completes the proof. □

2.5 Pointwise Error Estimate I

Let $\psi \in CSPD_m(S^{d-1})$ denote an SBF whose Legendre coefficients decay like (2.48), i.e., whose native Hilbert space $\mathcal{H}_\psi$ is equivalent to the Sobolev space $W^s(S^{d-1})$, $s = (d-1+\alpha)/2$. In this section we aim to estimate the uniform rate at which the interpolant converges to its target function as the interpolation points fill the surface of the sphere. To measure the relative density of the point set Ξ in S^{d-1} we use the so-called mesh norm

$$h := h(\Xi, S^{d-1}) := \sup_{\boldsymbol{\eta} \in S^{d-1}} \min\{g(\boldsymbol{\eta}, \boldsymbol{\xi}_i) = \cos^{-1}(\boldsymbol{\eta}^T \boldsymbol{\xi}_i) : \boldsymbol{\xi}_i \in \Xi\}, \tag{2.79}$$

and our specific aim is then to estimate the value p such

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| = O(h^p), \quad \text{for all } f \in \mathcal{H}_\psi \text{ and } \boldsymbol{\xi} \in S^{d-1}. \tag{2.80}$$

The strategy we employ is again borrowed from RBF theory where we attempt to bound the optimal power function of ψ by a function of h. We begin our analysis with the following remarkable result from [JSW99].

Lemma 2.2 (Jetter, Stöckler and Ward) *Let $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ denote a set of distinct data points on S^{d-1} with mesh-norm h, and let K be the positive integer satisfying*

$$\frac{1}{K+1} \le 2h \le \frac{1}{K}. \tag{2.81}$$

Let $\boldsymbol{\xi} \in S^{d-1}$, then there exist coefficients $\{\gamma_i\}_{i=1}^N$ such that

$$\mathcal{Y}(\boldsymbol{\xi}) = \sum_{i=1}^N \gamma_i \mathcal{Y}(\boldsymbol{\xi}_i), \quad \text{for all } \mathcal{Y} \in \mathcal{H}_K(S^{d-1}), \tag{2.82}$$

and such that

$$\sum_{i=1}^N |\gamma_i| \le 2. \tag{2.83}$$

This result has played an important role in advancing our understanding of the SBF interpolation method. Specifically, it enables us to deliver our first error estimate.

Theorem 2.7 *Let $\psi \in CSPD_m(S^{d-1})$ be an SBF whose Legendre coefficients decay like (2.48) for some $\alpha > 0$. Let $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ denote a set distinct data points on S^{d-1} whose mesh-norm h satisfies (2.81) for some positive integer $K \ge m-1$. Let $f \in \mathcal{H}_\psi$ and s_f denote its unique SBF interpolant. Then, for any $\boldsymbol{\xi} \in S^{d-1}$, we have*

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| \le \mathcal{C} \cdot h^{\frac{\alpha}{2}} \|f - s_f\|_\psi, \tag{2.84}$$

where $\mathcal{C}$ is a positive constant independent of h.

Proof The assumption $K \geq m - 1$ allows us to deduce that the coefficients $\{\gamma_i\}_{i=1}^N$ from Lemma 2.2 also satisfy condition (2.70), and so, in view of (2.77), they can be used to bound the optimal power function of ψ :

$$
\begin{aligned}
P_{\psi,\gamma^*}(\boldsymbol{\xi}) \leq\ & \Big\| \sum_{i=0}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i} \Big\|_{\psi^*} = \sup_{f \in \mathcal{H}_\psi:\, \|f\|_\psi \leq 1} \Big| (\sum_{i=0}^{N} \gamma_i \delta_{\boldsymbol{\xi}_i})(f) \Big| \\
&= \sup_{f \in \mathcal{H}_\psi:\, \|f\|_\psi \leq 1} \Big| \sum_{i=0}^{N} \gamma_i \sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \widehat{f}_{k,\ell} \mathcal{Y}_{k,\ell}(\boldsymbol{\xi}_i) \Big| \\
&\leq \sup_{f \in \mathcal{H}_\psi:\, \|f\|_\psi \leq 1} \sum_{i=0}^{N} |\gamma_i| \cdot \Big| \sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \widehat{f}_{k,\ell} \mathcal{Y}_{k,\ell}(\boldsymbol{\xi}_i) \Big| \\
&\leq \sup_{f \in \mathcal{H}_\psi:\, \|f\|_\psi \leq 1} \left(\sum_{i=0}^{N} |\gamma_i| \right) \cdot \max_{i \in \{0,\ldots,N\}} \Big| \sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \widehat{f}_{k,\ell} \mathcal{Y}_{k,\ell}(\boldsymbol{\xi}_i) \Big| \\
&\leq 3 \cdot \sup_{f \in \mathcal{H}_\psi:\, \|f\|_\psi \leq 1} \max_{i \in \{0,\ldots,N\}} | \sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \widehat{f}_{k,\ell} \mathcal{Y}_{k,\ell}(\boldsymbol{\xi}_i) |. \qquad (2.85)
\end{aligned}
$$

To bound the maximum value we can employ the Cauchy-Schwarz inequality together with the addition formula (1.24). Specifically, for any $\boldsymbol{\xi} \in S^{d-1}$, we have

$$
\begin{aligned}
\left(\sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \widehat{f}_{k,\ell} \mathcal{Y}_{k,\ell}(\boldsymbol{\xi}) \right)^2 &\leq \left(\sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \frac{\widehat{f}_{k,\ell}^2}{\widehat{\psi}_k} \right) \cdot \left(\sum_{k>K} \sum_{\ell=1}^{N_{k,d}} \widehat{\psi}_k \mathcal{Y}_{k,\ell}^2(\boldsymbol{\xi}) \right) \\
&\leq \|f\|_\psi^2 \cdot \sum_{k>K} \frac{N_{k,d} \widehat{\psi}_k}{\omega_{d-1}} \leq \sum_{k>K} \frac{N_{k,d} \widehat{\psi}_k}{\omega_{d-1}} = \sum_{k>K} a_\psi(k),
\end{aligned}
$$

where, by (1.33), the $a_\psi(k)$ denotes the k^{th} Legendre expansion coefficient of ψ. We can continue bounding from above by using (2.48) and (2.81) to give

$$
\begin{aligned}
\sum_{k>K} a_\psi(k) &\leq C_a \sum_{k>K} \frac{1}{(1+k)^{\alpha+1}} \leq C_a \int_K^\infty \frac{dx}{(1+x)^{\alpha+1}} \\
&= \frac{C_a}{\alpha} \cdot \frac{1}{(1+K)^\alpha} \leq C \cdot h^\alpha, \quad \text{where } C = \frac{2^\alpha C_a}{\alpha}. \qquad (2.86)
\end{aligned}
$$

Now, linking (2.85) and (2.86) together gives

$$P_{\psi,\gamma^*}(\boldsymbol{\xi}) \le \mathcal{C} \cdot h^{\alpha/2}, \quad \boldsymbol{\xi} \in S^{d-1}.$$

Thus, for any $\boldsymbol{\xi} \in S^{d-1}$, we can use (2.76) to deduce that

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| \le \mathcal{C} \cdot h^{\alpha/2} \cdot \|f - s_f\|_{\psi}, \quad \text{for all } f \in \mathcal{H}_{\psi},$$

where $\mathcal{C}$ is a positive constant independent of h. □

2.6 Pointwise Error Estimate II

In this section we will examine a similar but local approach to bounding the pointwise interpolation error. This development was initiated by Golitschek and Light [GL01] and later refined in [M01]. As usual the basic idea is to bound the optimal power function of ψ at $\boldsymbol{\xi} \in S^{d-1}$; however, the new approach shows that this can be done by using only those data points $\boldsymbol{\xi}_i \in \Xi$ which lie within a certain neighbourhood of $\boldsymbol{\xi}$. This is in contrast to Theorem 2.7 where every location in Ξ is used to bound the power function.

In our previous result we have used the mesh-norm h to measure the relative density of a set of data points $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^{N}$ in S^{d-1}. Geometrically speaking, h represents the radius of the largest spherical cap (open geodesic ball) which can be placed on S^{d-1} without covering any $\boldsymbol{\xi}_i$. In [GL01], von Golitschek and Light use the height h_d of the maximal spherical cap as an alternative mesh-norm; that is, they define h_d to be the smallest number such that

$$\inf_{\boldsymbol{\eta} \in S^{d-1}} \max\{\boldsymbol{\eta}^T \boldsymbol{\xi}_i : \boldsymbol{\xi}_i \in \Xi\} > 1 - h_d, \tag{2.87}$$

is satisfied. We shall call h_d the "dot product" mesh norm of Ξ. Using some elementary trigonometry we can show that $h_d = 2\sin^2(h/2)$. Furthermore, if $h \in (0, 2\pi/3)$ then we can apply the small angle result for $\sin(h/2)$ to give

$$\frac{h^2}{8} \le h_d \le \frac{h^2}{2} \tag{2.88}$$

that is, h_d is equivalent to h^2. The idea of using the dot product as an alternative measure of distance will prove to be a useful one.

Definition 2.12 (*Dot product neighbourhood*) For every $\boldsymbol{\xi} \in S^{d-1}$ we define an associated a dot product distance function

$$d_{\boldsymbol{\xi}} : S^{d-1} \to [-1, 1], \quad \text{given by} \quad d_{\boldsymbol{\xi}}(\boldsymbol{\eta}) = \boldsymbol{\xi}^T \boldsymbol{\eta}.$$

Furthermore, we can define a dot product neighbourhood of $\boldsymbol{\xi}$ by

$$N(\boldsymbol{\xi}, r_d) = \{\boldsymbol{\eta} \in S^{d-1} : d_{\boldsymbol{\xi}}(\boldsymbol{\eta}) > 1 - r_d\}, \quad \text{where} \;\; r_d \in (0, 1). \tag{2.89}$$

Proposition 2.3 *Let $\boldsymbol{\xi} \in S^{d-1}$ be a fixed point and let $r_d \in (0, 1)$, then*

$$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in N(\boldsymbol{\xi}, r_d) \Longrightarrow 1 - \boldsymbol{\eta}_1^T \boldsymbol{\eta}_2 < 4r_d.$$

Proof For any $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in S^{d-1}$ we have the following useful relation

$$2 - 2\boldsymbol{\eta}_1^T \boldsymbol{\eta}_2 = \|\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2\|^2.$$

Furthermore, if $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in N(\boldsymbol{\xi}, r_d)$ then we also know that

$$2 - 2\boldsymbol{\eta}_i^T \boldsymbol{\xi} < 2r_d, \quad i \in \{1, 2\},$$

which allows us to deduce that

$$\begin{aligned} \sqrt{2 - 2\boldsymbol{\eta}_1^T \boldsymbol{\eta}_2} &= \|(\boldsymbol{\eta}_1 - \boldsymbol{\xi}) + (\boldsymbol{\xi} - \boldsymbol{\eta}_2)\| \le \|\boldsymbol{\eta}_1 - \boldsymbol{\xi}\| + \|\boldsymbol{\xi} - \boldsymbol{\eta}_2\| \\ &= \sqrt{2 - 2\boldsymbol{\eta}_1^T \boldsymbol{\xi}} + \sqrt{2 - 2\boldsymbol{\eta}_2^T \boldsymbol{\xi}} < 2\sqrt{2r_d}, \end{aligned}$$

and the proof is complete. □

The following crucial result is quoted from [GL01].

Lemma 2.3 (von Golitschek and Light) *Let $\boldsymbol{\xi} \in S^{d-1}$ and let J be a fixed positive integer. Let $\Xi = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$ denote a set of N distinct data points on S^{d-1} with dot product mesh-norm h_d. There is a number $h_0 \in (0, 1)$ such that if $h_d < h_0$, then there exist coefficients $\{\gamma_i\}_{i=1}^N$ such that*

1. $\mathcal{Y}(\boldsymbol{\xi}) = \sum_{i=1}^{N} \gamma_i \mathcal{Y}(\boldsymbol{\xi}_i)$, *for all* $\mathcal{Y} \in \mathcal{H}_{J-1}(S^{d-1})$,
2. *there exists a constant K_1 (independent of $\boldsymbol{\xi}$ and h_d) such that if $\boldsymbol{\xi}_i \notin N(\boldsymbol{\xi}, K_1 h_d)$, then $\gamma_i = 0$, and*
3. *there exists a constant K_2 (independent of $\boldsymbol{\xi}$ and h_d) such that* $\sum_{i=1}^{N} |\gamma_i| \le K_2$.

It is pertinent to mention that Lemma 2.3 is similar in spirit to Lemma 2.2. To illustrate this, we provide the following useful comparison list.

C1. For a sufficiently dense set of data-points, both Lemmata supply coefficients $\{\gamma_i\}_{i=1}^N$ which satisfy condition (2.70) and, in both cases, the quantity $\sum_{i=1}^{N} |\gamma_i|$ is suitably bounded.

C2. For a given $\boldsymbol{\xi} \in S^{d-1}$, the coefficients $\{\gamma_i\}_{i=1}^N$ arising from Lemma 2.3 are said to be "local" since $\gamma_i \neq 0$ if and only if $\boldsymbol{\xi}_i \in N(\boldsymbol{\xi}, K_1 h_d)$.

C3. Lemma 2.3 is stated for a "fixed" positive integer J, whereas Lemma 2.2 is stated for an integer K which depends upon the mesh-norm h of Ξ. In both cases the result for m, and hence condition (2.70), follows if we assume that $m - 1 \leq \max\{K, J\}$.

The main aim, again, is to provide a suitable bound on the optimal power function of ψ and hence, using (2.76), deduce error estimate results for SBF interpolation. However, in contrast to the previous attempt, we will pursue a different approach which relies heavily on Taylor series analysis; for the convenience of the reader we briefly compose the key arguments.

$\diamond$ **Analysis via Taylor series.** Let $\psi \in CSPD_{m-1}(S^{d-1})$ be an SBF whose Legendre coefficients decay like (2.48) for some $\alpha > 0$. Let Ψ denote the zonal kernel induced by ψ. For a fixed $\boldsymbol{\xi} \in S^{d-1}$, we consider the function $F_{\boldsymbol{\xi}} : S^{d-1} \to \mathbb{R}$ given by $F_{\boldsymbol{\xi}}(\boldsymbol{\eta}) = \Psi(\boldsymbol{\xi}, \boldsymbol{\eta})$. In particular, we can write

$$F_{\boldsymbol{\xi}}(\boldsymbol{\eta}) = \sum_{k=m}^{\infty} a_k P_{k,d}(\boldsymbol{\xi}^T \boldsymbol{\eta}). \tag{2.90}$$

Our aim is to investigate the behaviour of $F_{\boldsymbol{\xi}}$ in a local neighbourhood $N(\boldsymbol{\xi}, r_d)$ and, in view of (2.90), we can do this by studying the local behaviour of the $d-$dimensional Legendre polynomials. Specifically, we choose a suitable positive integer J and consider the Taylor expansion

$$P_{k,d}(t) = \sum_{r=0}^{J-1} \frac{P_{k,d}^{(r)}(1)}{r!}(1-t)^r + R_J(k,t), \quad t \in (1 - r_d, 1], \tag{2.91}$$

where the remainder term $R_J(k,t)$ satisfies

$$|R_J(k,t)| \leq \frac{(1-t)^J}{J!} \sup_{t \in (1-r,1]} |P_{k,d}^{(J)}(t)|, \quad \text{for } k \geq J.$$

For all $d \geq 2$, we can use (1.26) together with Markov's inequality for algebraic polynomials ([DL93, Chap. 4], to deduce that

$$|R_J(k,t)| \leq \frac{(1-t)^J}{J!} k^{2J}, \quad \text{for } k \geq J. \tag{2.92}$$

As a final remark we note that, for each $d \geq 2$, the polynomials $\{P_{k,d}\}_{k=0}^r$ form a basis for the space of univariate polynomials on $[-1, 1]$ of degree at most r. In particular, for a given $r \in \mathbb{N}_0$, there exists real coefficients $\{\alpha_{rs}\}_{s=0}^r$ such that

$$(1-t)^r = \sum_{s=0}^{r} \alpha_{rs} P_{s,d}(t). \tag{2.93}$$

With this preparation we are now in position to prove our next error estimate.

Theorem 2.8 *Let $\psi \in CSPD_m(S^{d-1})$ be an SBF whose Legendre coefficients decay like (2.48) for some $\alpha > 0$. Let $\Xi = \{\boldsymbol{\xi}_i\}_{i=1}^N$ denote a set distinct points on S^{d-1} with mesh-norm h. Set*

$$J = \max\left\{m, \frac{\lceil \alpha + 1 \rceil}{2}\right\}, \tag{2.94}$$

where $\lceil x \rceil$ denotes the smallest integer $\geq x$, and assume that the dot product mesh-norm h_d (2.87) of Ξ satisfies

$$\frac{1}{(K+1)^2} \leq h_d < \frac{1}{K^2}, \tag{2.95}$$

where $K > J$ is a positive integer. Let $f \in \mathcal{H}_\psi$ and s_f denote its unique SBF interpolant. Then, for any $\boldsymbol{\xi} \in S^{d-1}$, we have

$$|f(\boldsymbol{\xi}) - s_f(\boldsymbol{\xi})| \leq \mathcal{C} \cdot h^{\alpha/2} \cdot \|f - s_f\|_\psi, \tag{2.96}$$

where $\mathcal{C}$ is a positive constant independent of h.

Proof The choice of integer J (2.94), allows us to evoke Lemma 2.3 to provide, for any $\boldsymbol{\xi} \in S^{d-1}$, a neighbourhood $N(\boldsymbol{\xi}, K_1 h_d)$ and a set of local coefficients $\{\gamma_i\}_{i \in I_{\ell oc}}$, where $I_{\ell oc} := \{i : \boldsymbol{\xi}_i \in \Xi \cap N(\boldsymbol{\xi}, K_1 h_d)\}$, which satisfy condition (2.70). Furthermore, these coefficients can be used to define a local power function which, in turn, provides a bound on the optimal power function of ψ :

$$P^2_{\psi,\gamma^*}(\boldsymbol{\xi}) \leq P^2_{\psi,\ell oc}(\boldsymbol{\xi}) = \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j \psi(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) = \sum_{k=m}^{\infty} a_\psi(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j P_{k,d}(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j),$$

where we have employed the Legendre expansion of ψ. For our investigation it is useful to split the above sum into two parts; that is, we shall consider

$$\underbrace{\sum_{k=m}^{K} a_\psi(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j P_{k,d}(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j)}_{sum\ 1} + \underbrace{\sum_{k=K+1}^{\infty} a_\psi(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j P_{k,d}(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j)}_{sum\ 2}. \tag{2.97}$$

We begin by considering "sum 1" of (2.97). In particular, substituting in the Taylor expansion (2.91) of the Legendre polynomials yields

$$\sum_{k=m}^{K} a_\psi(k) \left(\sum_{r=0}^{J-1} \frac{P_{k,d}^{(r)}(1)}{r!} \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j (1 - \boldsymbol{\xi}_i^T \boldsymbol{\xi}_j)^r + \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j R_J(k, \boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) \right).$$

We continue our development by analysing the first sum appearing in the brackets. Using identity (2.93) followed by an application of addition formula (1.24), gives

$$\begin{aligned}\sum_{i,j\in I_{\ell oc}\cup\{0\}} \gamma_i\gamma_j(1-\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j)^r &= \sum_{s=0}^{r}\alpha_{rs}\sum_{i,j\in I_{\ell oc}\cup\{0\}}\gamma_i\gamma_j P_{s,d}(t)\\ &= \sum_{s=0}^{r}\alpha_{rs}\frac{\omega_{d-1}}{N_{s,d}}\sum_{i,j\in I_{\ell oc}\cup\{0\}}\gamma_i\gamma_j\sum_{\ell=0}^{N_{s,d}}\mathcal{Y}_{s,\ell}(\boldsymbol{\xi}_i)\mathcal{Y}_{k,\ell}(\boldsymbol{\xi}_j)\\ &= \sum_{s=0}^{r}\alpha_{rs}\frac{\omega_{d-1}}{N_{s,d}}\sum_{\ell=0}^{N_{s,d}}\left(\sum_{i\in I_{\ell oc}\cup\{0\}}\gamma_i\mathcal{Y}_{s,\ell}(\boldsymbol{\xi}_i)\right)^2.\end{aligned}$$

This expression vanishes for $0 \le r \le J-1$, by part (i) of Lemma 2.3, and thus, the bound reduces to

$$\sum_{k=m}^{K} a_\psi(k)\sum_{i,j\in I_{\ell oc}\cup\{0\}}\gamma_i\gamma_j R_J(k,\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j) = \sum_{k=J}^{K} a_\psi(k)\sum_{i,j\in I_{\ell oc}\cup\{0\}}\gamma_i\gamma_j R_J(k,\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j),$$

since the remainder $R_J(k,t)$ is zero for $k \le J-1$.

Now, for any $i, j \in I_{\ell oc}$ we have $\boldsymbol{\xi}_i, \boldsymbol{\xi}_j \in N(\boldsymbol{\xi}, K_1 h_d)$, and so, by Proposition 2.3, it follows that $(1-\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j)^J < (4K_1h_d)^J$. We can use this fact, together with (2.92), to deduce

$$\begin{aligned}\sum_{k=J}^{K} a_\psi(k)\sum_{i,j\in I_{\ell oc}\cup\{0\}}\gamma_i\gamma_j R_J(k,\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j) &\le \sum_{k=J}^{K} a_\psi(k)\sum_{i,j\in I_{\ell oc}\cup\{0\}}|\gamma_i\gamma_j|\frac{(1-\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j)^J}{J!}k^{2J}\\ &\le \sum_{k=J}^{K} a_\psi(k)\sum_{i,j\in I_{\ell oc}\cup\{0\}}|\gamma_i\gamma_j|\frac{(4K_1h_d)^J}{J!}k^{2J}\\ &= \sum_{k=J}^{K} a_\psi(k)\left(\sum_{i\in I_{\ell oc}\cup\{0\}}|\gamma_i|\right)^2\frac{(4K_1h_d)^J}{J!}k^{2J}.\end{aligned}$$

We can now apply part (iii) of Lemma 2.3 to give

$$\sum_{k=m}^{K} a_\psi(k)\sum_{i,j\in I_{\ell oc}\cup\{0\}}\gamma_i\gamma_j P_{k,d}(\boldsymbol{\xi}_i^T\boldsymbol{\xi}_j) \le (1+K_2)^2\frac{(4K_1h_d)^J}{J!}\sum_{k=J}^{K} a_\psi(k)k^{2J}. \quad (2.98)$$

Using the assumed decay rate of the Legendre coefficients we can write

$$\sum_{k=J}^{K} a_{\psi}(k)k^{2J} \leq C_a \sum_{k=J}^{K} \frac{k^{2J}}{(1+k)^{\alpha+1}} \leq C_a \sum_{k=J}^{K} k^{2J-(\alpha+1)}.$$

The definition of J guarantees that the function $x \mapsto x^{2J-(\alpha+1)}$ is non-decreasing on $[0, \infty)$ and hence we have the bound

$$\sum_{k=J}^{K} a_{\psi}(k)k^{2J} \leq C_a \int_J^K x^{2J-(\alpha+1)}dx \leq C_a \cdot K^{2J-\alpha} \leq C_a \cdot h_d^{-J+\frac{\alpha}{2}}, \qquad (2.99)$$

where the final inequality follows from (2.95). Linking (2.98) and (2.99) together gives us our final bound for "sum 1" that is,

$$\sum_{k=m}^{K} a_{\psi}(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j P_{k,d}(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) \leq C_{sum1} \cdot h_d^{\frac{\alpha}{2}}, \qquad (2.100)$$

where

$$C_{sum1} = C_a \cdot (1+K_2)^2 \cdot \frac{(4K_1)^J}{J!} \qquad (2.101)$$

is independent of h_d.

We now turn to "sum 2" of (2.97) which is easier to bound. Specifically, we use (1.26) followed by part (iii) of Lemma 2.3 to yield

$$\sum_{k=K+1}^{\infty} a_{\psi}(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j P_{k,d}(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) \leq \sum_{k=K+1}^{\infty} a_{\psi}(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} |\gamma_i \gamma_j|$$

$$= \sum_{k=K+1}^{\infty} a_{\psi}(k) \left(\sum_{i \in I_{\ell oc} \cup \{0\}} |\gamma_i| \right)^2 \leq (K_2+1)^2 \sum_{k=K+1}^{\infty} a_{\psi}(k). \qquad (2.102)$$

Again we use the decay of the Legendre coefficients together (2.95) to deduce that

$$\sum_{k=K+1}^{\infty} a_{\psi}(k) \leq C_a \sum_{k=K+1}^{\infty} \frac{1}{(1+k)^{\alpha+1}} \leq C_a \int_K^{\infty} \frac{dx}{(1+x)^{\alpha+1}}$$

$$= \frac{C_a}{\alpha} \frac{1}{(K+1)^{\alpha}} \leq \frac{C_a}{\alpha} \cdot h_d^{\frac{\alpha}{2}}. \qquad (2.103)$$

Linking (2.102) and (2.103) together provides the following bound for "sum 2"

$$\sum_{k=K+1}^{\infty} a_{\psi}(k) \sum_{i,j \in I_{\ell oc} \cup \{0\}} \gamma_i \gamma_j P_{k,d}(\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j) \leq C_{sum2} \cdot h_d^{\frac{\alpha}{2}}, \tag{2.104}$$

where

$$C_{sum2} = (1 + K_2)^2 \cdot \frac{C_a}{\alpha}, \tag{2.105}$$

is independent of h_d.

We are now in a position to provide a more meaningful bound on the optimal power function. In particular, in view of (2.100), (2.104) and the mesh-norm equivalence relation (2.88), we choose to set $\mathcal{C} = 2^{-\alpha/2} \max\{C_{sum1}, C_{sum2}\}$ and deduce

$$P^2_{\psi,\gamma^*}(\boldsymbol{\xi}) \leq P^2_{\psi,\ell oc}(\boldsymbol{\xi}) \;\leq\; \max\{C_{sum1}, C_{sum2}\} \cdot h_d^{\frac{\alpha}{2}} \;\leq\; \mathcal{C} \cdot h^{\alpha}, \tag{2.106}$$

the proof is then completed by employing this bound in (2.76). □

Spherical Radial Basis Functions, Theory and Applications
Hubbert, S.; Le Gia, Q.T.; Morton, T.M.
2015, X, 143 p. 7 illus., 3 illus. in color., Softcover
ISBN: 978-3-319-17938-4