Università degli Studi di Padova
Dipartimento di Matematica "Tullio Levi-Civita"
Corso di Laurea Magistrale in Matematica

Rescaled localized radial basis functions and fast decaying polynomial reproduction

Supervisor:
Prof. Stefano De Marchi
Co-supervisor:
Prof. Holger Wendland

Candidate:
Giacomo Cappellazzo
ID number 2056525

Academic year 2022/2023
September 22, 2023

Abstract

Approximating a set of data can be a difficult task but it is very useful in applications. Through a linear combination of basis functions we want to reconstruct an unknown quantity from partial information. We study radial basis functions (RBFs) to obtain an approximation method that is meshless, provides a data dependent approximation space and generalization to larger dimensions is not an obstacle. We analyze a rational approximation method with compactly supported radial basis functions (Rescaled localized radial basis function method). The method reproduces exactly the constants and the density of the interpolation nodes influences the support of the RBFs. There is a proof of the convergence in a quasi-uniform setting up to a conjecture: we can determine a lower bound for the approximant of the constant function 1 uniformly with respect to the size of the support of the kernel. We investigate the statement of the conjecture and bring some practical and theoretical results to support it. We study the Runge phenomenon on the approximant and obtain uniform estimates on the cardinal functions. We extend the distinguishing features of the method reproducing exactly larger polynomial spaces. We replace local polynomial reproduction with basis functions that decrease rapidly and approximate exactly a polynomial space. This change releases the basis functions from the compactness of the support and guarantees the same convergence rate (the oversampling problem does not appear). The rescaled localized radial basis function method can be interpreted in this new framework because the cardinal functions have global support even if the kernel has compact support. The decay of the basis functions undertake convergence and stability. In this analysis the smoothness of the approximant is not important, what matters is the "locality" provided by the fast decay. With a moving least squares approach we provide an example of a smooth quasi-interpolant. We continue trying to improve the performance of the method even when the weight functions do not have compact support. All the new theoretical results introduced in this work are also supported by numerical evidence.

Contents

Introduction

Approximating a set of data can be a difficult task but it is also very useful and necessary in applications. Through a linear combination of basis functions we want to reconstruct an unknown quantity from partial information to study its behavior or to make predictions about the future. The basis functions in principle must be "simple": from an academic point of view we must be able to analyze them to obtain results of uniqueness, convergence and stability, furthermore numerically they must be implementable effectively and efficiently.

To be more precise we briefly introduce the mathematical setting. Let $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ a data set, $\{f_1, \ldots, f_N\}$ the discrete information we know on X of a continuous quantity $f : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}$ and $\{u_1, \ldots, u_N\} \subseteq \mathcal{C}(\Omega)$ are the basis functions. Our goal is to determine $a \in \mathbb{R}^N$ such that

$$s_{f,X}(x_i) = \sum_{j=1}^{N} a_j u_j(x_i) \approx f_i \quad \text{for } i = 1, \ldots, N.$$

If we replace $\approx$ with $=$ then we are studying an interpolation method, otherwise we call $s_{f,X}$ quasi-interpolant. Some themes of fundamental importance immediately arise: Is the method well defined? Is the solution unique? Does the approximant accurately reconstruct the function f?

We can easily answer these questions using a polynomial basis, but we have to impose some conditions on the initial data. This request can not always be satisfied in applications (it is too expensive to resample the function f, the event we want to study can not be repeated or the dimensionality of the problem is too high). The approximation method that we decide to use, in addition to satisfy conditions of technical nature, must be able to deal with arbitrary data distributions and generalization to larger dimensions must not be an obstacle. Moreover we must be aware of which functions we can reconstruct exactly, which functions we can approximate and above all how many nodes we need to obtain coherent results.

This is precisely the intention of Chapter 1 and Chapter 2. We study radial basis functions (RBFs) to obtain an approximation method that is meshless and provides a data dependent approximation space. We aim to investigate the interconnections between reproducing-kernel Hilbert spaces and positive definite kernels (which are a generalization of radial basis functions). In native spaces, which are the natural environment in which to study approximation methods with RBFs, we can recognize, generalize and abstract classical functional spaces (e.g. Sobolev spaces). We reserved in-depth analysis to conditionally positive definite functions, that allow us to reproduce exactly a finite-dimensional vector subspace of continuous functions (in this work we considered polynomial spaces).

In Chapter 3 we use compactly supported radial basis functions as weight functions to determine convergent approximation methods. We review concepts like local polynomial reproduction and moving least squares to obtain a quadratic convex optimization problem that gives us basis functions that guarantee a convergence rate of $\mathcal{O}(h_{X,\Omega}^{m+1})$ if polynomials up to degree m can be reconstructed exactly (we work in a quasi-uniform setting). With these techniques we can gain convergent and stable methods (we are able to determine the computational costs exactly) but we are restricting ourselves to basis functions with compact support and oversampling problems can arise. Although compact support is effective in practice it is difficult to reduce numerical schemes in this framework.

We continue with Chapter 4 and Chapter 5 to study the convergence error for RBF methods and the so-called trade-off principle, linking approximation error, numerical stability, eigenvalues of the interpolation matrix, fill distance and separation radius. Long story short, when the approximation improves, i.e. the power function decreases, then also the minimal eigenvalue of the interpolation matrix decreases (hence the conditioning of the matrix increases). A substantial part of these sections is devoted to prove estimates on the conditioning of the interpolation matrix when the native space coincides with a Sobolev space and when the basis functions are scaled with a parameter δ, which controls their support.

Chapter 6 summarizes several optimality properties of the RBF interpolant and of the cardinal functions when observed through the native space and its norm.

The discussion concludes with Chapter 7, which also contains most of the original contributions. In [1] is described a rational approximation method with compactly supported radial basis functions (Rescaled localized radial basis function method). The proposed method, which can be considered as an instance of Shepard's method, reproduces exactly the constants and the main peculiarity is the variability of the support of the RBFs (the size of the support depends on the density of the interpolation nodes and their mutual position). This method aims to make the most of the locality guaranteed by the compactness of the support. If the size of the support remains fixed and the number of nodes increases then the problem becomes global again. Only numerical evidence were provided to verify the well-posedness of the method and the linear convergence with respect to the fill distance. The proof of the convergence in a quasi-uniform setting is due to [2] up to a conjecture: we can determine a lower bound for the approximant of the constant function 1 uniformly with respect to the size of the support of the kernel under analysis. It is important to underline that convergence can not be proved with classical inequalities for RBFs because in a quasi-uniform context the scaled version of the power function does not decrease with the fill distance. We investigated the statement of the conjecture and brought some practical and theoretical results to support it. The research is based on the comparison between the cardinal functions observed from different perspectives (on the one hand the classic cardinal functions and on the other the cardinal functions of a method that reproduces exactly the constants). We analyzed the Runge phenomenon on the approximant and obtained uniform estimates on the cardinal functions (the uniform norm is bounded and the constants, i.e. the polynomials of degree zero, which allow the constant functions to be reproduced exactly tend to zero as the number of nodes increases). These results can be obtained with two different approaches, one considers the abstract definition of native space (the space of functions that make some functionals continuous), while the other uses the matrix algebra. We also studied the behavior of the constants that appear in [2].

A careful review of [2], with the goal of reproducing exactly larger polynomial spaces, led to a generalization of the results of Chapter 3. We replaced local polynomial reproduction with basis functions that decrease rapidly and approximate exactly the same polynomial space. This change releases the basis functions from the compactness of the support and guarantees the same convergence rate (the oversampling problem does not appear). The rescaled localized radial basis function method (RL-RBF) can be interpreted in this new framework because the cardinal functions have global support even if the kernel has compact support. The decay of the basis functions, governed by a scalar function of distance and of the separation radius, undertake convergence and stability. In this analysis the smoothness of the approximant is not important, what matters is the "locality" provided by the fast decay.

With a moving least squares approach we provided an example of a smooth ($\mathcal{C}^\infty(\mathbb{R}^d)$) quasi-interpolant (the Lebesgue constant can be obtained explicitly and the computational cost is linear in the number of nodes). We want to remark that with the standard construction with compactly supported RBF the smoothness of the approximant is inherited by the smoothness of the RBF (if we use Wendland's functions then the smoothness is limited). We continued the work trying to improve the performance of the method even when the weight functions do not have compact support (the matrices that are involved, even if with small dimensions when the polynomial space to be reproduced is not too large, can be dense and led to numerical instability). To address the obstacle we replaced a quadratic optimization problem with a linear program on a polyhedron (warm start techniques, column generation techniques and the choice of appropriate solvers provide speed in the execution of the approximation algorithm and a bound on the number of basis functions different from zero in each point of the domain).

All the new theoretical results introduced in this work are also supported by numerical evidence (most of the code was produced with MATLAB while for the part concerning the optimization AMPL was used).

Chapter 1

Native Spaces

1.1 Riesz representation theorem

For a better understanding of the reproducing kernels we need to introduce some definitions and results about metric and duality properties of Hilbert spaces. We will follow the constructions in [3, 4, 5] to introduce these concepts.

Definition 1.1 *(Convex sets and distance to a set). A convex set is a subset U of a vector space V such that for all $u, v \in U$, $tu + (1-t)v \in U$ for all $t \in [0, 1]$. When V is a normed vector space, we say that the distance from a vector p to a subset U is defined by $dist(p, U) = \inf_{q \in U} \|p - q\|$.*

The following elementary identity of the inner product will be useful to prove a projection theorem.

Theorem 1.1 *(The parallelogram equality) If V is an inner product space with the norm induced by the inner product then for all vectors $u, v \in V$*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Proof
Distributing over the sums $u + v$ and $u - v$, we have

$$\begin{aligned} \|u + v\|^2 + \|u - v\|^2 &= \langle u + v, u + v\rangle + \langle u - v, u - v\rangle = \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle + \langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle = \\ &= 2\langle u, u\rangle + 2\langle v, v\rangle = 2\|u\|^2 + 2\|v\|^2. \end{aligned}$$

□

The following fact does not hold in general for Banach spaces, and indeed the following proof relies on the parallelogram equality.

Theorem 1.2 *(Hilbert Projection theorem) For a Hilbert spaces V and a closed convex subset U, the distance to p from U, as in the definition 1.1, is attained by a unique element of U.*

Proof
Let $\{q_n\}_{n\in\mathbb{N}}$ be a sequence of vectors in U whose distance to p approach $\text{dist}(p, U) = \inf_{q\in U} \|p - q\|$. We want to prove that this sequence is Cauchy. Applying the parallelogram equality (Theorem 1.1) to the pair $p - q_n$ and $p - q_m$ we have

$$\|(p - q_n) + (p - q_m)\|^2 + \|(p - q_n) - (p - q_m)\|^2 = 2\|p - q_n\|^2 + 2\|p - q_m\|^2.$$

From this we obtain

$$\|q_m - q_n\|^2 = 2\|p - q_n\|^2 + 2\|p - q_m\|^2 - 4\left\|p - \frac{q_n + q_m}{2}\right\|^2.$$

Noting that $\frac{q_n+q_m}{2} \in U$ by convexity, we get

$$\|q_m - q_n\|^2 \leq 2\|p - q_n\|^2 + 2\|p - q_m\|^2 - 4\text{dist}(p, U)^2,$$

from $\text{dist}(p, U) \leq \|p - \frac{q_n+q_m}{2}\|$. Since $\|p - q_n\|$ and $\|p - q_m\|$ approach $\text{dist}(p, U)$, the right side of the inequality can be made arbitrarily small by choosing large enough n and m. This proves that the sequence $\{q_n\}_{n\in\mathbb{N}}$ is Cauchy. Since V is an Hilbert space and U is a closed subset of V, the limit q of the sequence $\{q_n\}_{n\in\mathbb{N}}$ is an element of U. The continuity of the norm ensures that

$$\|p - q\| = \lim_{n\to\infty} \|p - q_n\| = \text{dist}(p, U).$$

To prove that q in unique, consider two such vectors q and q^*. With the same computation as above we obtain

$$\|q - q^*\|^2 \leq 2\|p - q\|^2 + 2\|p - q^*\|^2 - 4\text{dist}(p, U)^2 = 0,$$

which implies $q = q^*$.

□

Definition 1.2 *(Orthogonal projections) For a vector v and a closed convex subset U of a Hilbert space V, we use v_U to denote the distance-minimizing element of U, called the orthogonal projection of v into U.*

Proposition 1.3 *Let V be an Hilbert spaces and U a closed convex subset of V, then $\langle v - v_U, u\rangle = 0$ for all $v \in V$ and $u \in U$.*

Proof
For all scalars λ we have $\|v - v_U\|^2 \leq \|v - (v_U - \lambda u)\|^2$, as $v_U - \lambda u \in U$ and v_U is distance-minimizing. So for any $t > 0$, choosing $\lambda = -t\langle v - v_U, u\rangle$, we obtain

$$\|v - v_U\|^2 \leq \langle v - v_U + \lambda u, v - v_U + \lambda u\rangle = \|v - v_U\|^2 + |\lambda|^2\|u\|^2 + 2\lambda\langle v - v_U, u\rangle.$$

Substituting for λ, we get

$$\|v - v_U\|^2 \le \|v - v_U\|^2 + t^2 |\langle v - v_U, u\rangle|^2 \|u\|^2 - 2t|\langle v - v_U, u\rangle|^2$$

which implies, dividing by $t > 0$,

$$2|\langle v - v_U, u\rangle|^2 \le t|\langle v - v_U, u\rangle|^2 \|u\|^2.$$

If $t \to 0^+$ we can conclude $\langle v - v_U, u\rangle = 0$.

□

Definition 1.3 *(Orthogonal complemets) For a subset U of an inner product space V, we denote by $U^\perp$ the space of vectors orthogonal to U, called the orthogonal complement of U. More precisely*

$$U^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ for all } u \in U\}.$$

The following proposition will be fundamental when U is a dense subset of V.

Proposition 1.4 *Let U be a subset of a Hilbert space V, then*

1. *$U^\perp$ is a closed subspace of V,*
2. *$U^\perp = \overline{U}^\perp$ (where the bar denotes the metric closure),*
3. *$(U^\perp)^\perp = \overline{U}$, if U is a subspace of V.*

Proof
(1) Let $\{v_n\}_{n\in\mathbb{N}}$ be a converging sequence of elements of $U^\perp$. Since $v_n \to v \in V$ we have to prove that $v \in U^\perp$. By the continuity of the inner product we obtain

$$0 = \lim_{n\to\infty} \langle v_n, u\rangle = \langle v, u\rangle \quad \text{for all } u \in U.$$

(2) Since $U \subseteq \overline{U}$ we obtain the following inclusion:

$$\overline{U}^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ for all } u \in \overline{U}\} \subseteq \{v \in V : \langle v, u\rangle = 0 \text{ for all } u \in U\} = U^\perp.$$

To prove the reverse inclusion we note that for each $u \in \overline{U}$ there exists a sequence $\{u_n\}_{n\in\mathbb{N}}$ of elements of U converging to u. By the continuity of the inner product we obtain

$$0 = \lim_{n\to\infty} \langle v, u_n\rangle = \langle v, u\rangle \quad \text{for all } u \in \overline{U} \text{ and } v \in U^\perp.$$

(3) We begin to prove that $U \subseteq (U^\perp)^\perp$: by Definition 1.3 if $u \in U$ then $\langle u, v\rangle = 0$ for all $v \in U^\perp$. With (2) we get $\overline{U} \subseteq (\overline{U}^\perp)^\perp = (U^\perp)^\perp$.
We observe that $\overline{U}$ is a closed subspace when U is a subspace:

- If $\{u_n\}_{n\in\mathbb{N}}$ is a converging sequence of elements of U and its limit is u, then $\|\lambda u - \lambda u_n\| \leq |\lambda|\|u - u_n\| \to 0$ as $n \to +\infty$ for each $\lambda \in \mathbb{R}$. This proves that $\overline{U}$ is closed under multiplication by scalar since $\lambda u_n \in U$ for each $n \in \mathbb{N}$.

- If $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ are converging sequences of elements of U and their limits are respectively a and b, then $\|a+b-(a_n+b_n)\| \leq \|a-a_n\| + \|b-b_n\| \to 0$ as $n \to +\infty$. This proves that $\overline{U}$ is closed under addition since $a_n + b_n \in U$ for each $n \in \mathbb{N}$.

Suppose $v \in (U^\perp)^\perp$. Since $\overline{U}$ is a closed (convex) subspace we have that $v_{\overline{U}} \in \overline{U} \subseteq (U^\perp)^\perp$, so $v - v_{\overline{U}} \in (U^\perp)^\perp$. By Proposition 1.3 and (2) we have $v - v_{\overline{U}} \in \overline{U}^\perp = U^\perp$. Since $v - v_{\overline{U}} \in U^\perp \cap (U^\perp)^\perp$ we have $\|v - v_{\overline{U}}\|^2 = \langle v - v_{\overline{U}}, v - v_{\overline{U}}\rangle = 0$, which implies $v = v_{\overline{U}} \in \overline{U}$.

□

Corollary 1.5 *A subspace U of a Hilbert Space V is dense if and only if $U^\perp = \langle 0\rangle$.*

Proof
(⇒) By Proposition 1.4 we can write $(U^\perp)^\perp = \overline{U} = V$, which implies $U^\perp = \langle 0\rangle$.

(⇐) By Proposition 1.4 we obtain $\overline{U} = (U^\perp)^\perp = \langle 0\rangle^\perp = V$.

□

Eventually we can prove a fundamental representation result. In general, infinite-dimensional vector spaces are not isometrically isomorphic to their own dual space, but this is true for Hilbert spaces.

Theorem 1.6 *(Riesz representation theorem) For a continuous linear functional ϕ on a Hilbert space V, there exists a unique $u \in V$ such that $\phi(v) = \langle u, v\rangle$ for all $v \in V$. Furthermore, $\|u\|_V = \|\phi\|_{V^*}$.*

Proof
If ϕ is the zero functional, take $u = 0$. So assume ϕ is non-zero. By continuity, $\ker(\phi) = \{v \in V : \phi(v) = 0\} = \overleftarrow{\phi}(0)$ is a closed subspace of V. Since $\ker(\phi)$ is closed and assumed not to be all of V from Corollary 1.5 we have that $\ker(\phi)^\perp \neq \langle 0\rangle$. There exists $w \in V \setminus \{0\}$ such that $\|w\| = 1$ and $\langle w, v\rangle = 0$ for each $v \in V$ such that $\phi(v) = 0$. Choose $u = \phi(w)w$ and observe that $\phi(w) \neq 0$ since $w \notin \ker(\phi)$.

Since w is unit length, we have $\|u\| = |\phi(w)|$ and $\phi(u) = \phi(w)^2 = \|u\|^2$. For all $v \in V$, we have $\langle u, v\rangle = \langle u, v - \frac{\phi(v)}{\|u\|^2}u\rangle + \langle u, \frac{\phi(v)}{\|u\|^2}u\rangle$. Since ϕ applied to $v - \frac{\phi(v)}{\|u\|^2}u$ equals 0, and since u is orthogonal to any vector in $\ker(\phi)$, we have $\langle u, v\rangle = \langle u, \frac{\phi(v)}{\|u\|^2}u\rangle = \phi(v)$.

To prove uniqueness consider two such vectors u and u^*:

$$\|u - u^*\|^2 = \langle u - u^*, u - u^*\rangle = \langle u, u - u^*\rangle - \langle u^*, u - u^*\rangle = \phi(u - u^*) - \phi(u - u^*) = 0,$$

that implies $u = u^*$.

If $v \in V \setminus \{0\}$ the Cauchy-Schwarz inequality ensures that

$$\frac{|\phi(v)|}{\|v\|} = \frac{|\langle u, v\rangle|}{\|v\|} \leq \frac{\|u\|\|v\|}{\|v\|} = \|u\| \Rightarrow \sup_{v \in V\setminus\{0\}} \frac{|\phi(v)|}{\|v\|} \leq \|u\|.$$

The last implication is trivial if $u = 0$, otherwise the upper-bound is obtained when $v = u$. This proves that $\|u\|_V = \|\phi\|_{V^*}$.

□

We note that if $\phi(v) = \langle u, v\rangle$ then $\ker(\phi) = \{v \in V : \langle u, v\rangle = 0\} = \langle u\rangle^\perp$. This implies that $\ker(\phi)^\perp = \overline{\langle u\rangle} = \langle u\rangle$. Also for infinite-dimensional vector space the orthogonal complement of $\ker(\phi)$ is exactly a one-dimensional vector space. We proved this without explicitly construct an orthonormal bases like there was in the finite-dimensional case. We have no suggestion how to construct u.

1.2 Reproducing-kernel Hilbert spaces

We study vector spaces $\mathcal{F}$ consisting of real-valued functions $f : \Omega \to \mathbb{R}$ defined on a quite arbitrary region $\Omega \subseteq \mathbb{R}$. For the error analysis of the interpolation process we want to study the correspondence between real-valued positive-definite kernels and Hilbert spaces of real-valued functions. The concept of reproducing-kernel Hilbert space plays a fundamental role in numerical analysis. Reproducing-kernel Hilbert spaces have been introduced in [6, 7]. A modern approach with links on radial basis functions can be found in [8, 9, 10, 11]. The method we decided to follow can be found in [12].

Definition 1.4 *Let $\mathcal{F}$ be a Hilbert space of functions $f : \Omega \to \mathbb{R}$. A function $\Phi : \Omega \times \Omega \to \mathbb{R}$ is called reproducing kernel for $\mathcal{F}$ if*

- $\Phi(\cdot, y) \in \mathcal{F} \quad \forall\, y \in \Omega,$
- $f(y) = \langle f, \Phi(\cdot, y)\rangle_{\mathcal{F}} \quad \forall\, f \in \mathcal{F},\, y \in \Omega.$

Proposition 1.7 *Suppose that $\mathcal{F}$ is a Hilbert space of functions $f : \Omega \to \mathbb{R}$. If Φ_1 and Φ_2 are reproducing kernel for $\mathcal{F}$ then $\Phi_1 = \Phi_2$.*

Proof
From the definition 1.4 we obtain that $\langle f, \Phi_1(\cdot, y) - \Phi_2(\cdot, y)\rangle_{\mathcal{F}} = \langle f, \Phi_1(\cdot, y)\rangle_{\mathcal{F}} - \langle f, \Phi_2(\cdot, y)\rangle_{\mathcal{F}} = f(y) - f(y) = 0$ for all $y \in \Omega$. Setting $f = \Phi_1(\cdot, y) - \Phi_2(\cdot, y) \in \mathcal{F}$ we have that $\langle \Phi_1(\cdot, y) - \Phi_2(\cdot, y), \Phi_1(\cdot, y) - \Phi_2(\cdot, y)\rangle_{\mathcal{F}} = 0$. This implies $\Phi_1(\cdot, y) - \Phi_2(\cdot, y)\, \forall y \in \Omega$.

□

Theorem 1.8 *Suppose that $\mathcal{F}$ is a Hilbert space of functions $f : \Omega \to \mathbb{R}$. Then the following statements are equivalent:*

- *the point evaluation functionals are continuous, i.e.* $\delta_y \in \mathcal{F}^*$ *for all* $y \in \Omega$*;*
- $\mathcal{F}$ *has a reproducing kernel.*

Proof
($\Rightarrow$) By Riesz representation theorem (Theorem 1.6) and the continuity of the point evaluation functionals, for every $y \in \Omega$ we can find a $\Phi_y \in \mathcal{F}$ such that $f(y) = \delta_y(f) = \langle \Phi_y, f \rangle_{\mathcal{F}}$ for all $f \in \mathcal{F}$. Thus if we define $\Phi(x, y) := \Phi_y(x)$ we obtain $f(y) = \langle f, \Phi_y \rangle_{\mathcal{F}} = \langle f, \Phi(\cdot, y) \rangle_{\mathcal{F}}$. This proves that Φ is a reproducing kernel for $\mathcal{F}$.

($\Leftarrow$) If $\mathcal{F}$ has a reproducing kernel we have that $\delta_y(f) = f(y) = \langle f, \Phi(\cdot, y) \rangle_{\mathcal{F}}$ for all $y \in \Omega$. Since the point evaluation functional δ_y is linear, we check that it is bounded.

$$|\delta_y(f)| = |\langle f, \Phi(\cdot, y) \rangle_{\mathcal{F}}| \leq \|f\|_{\mathcal{F}} \|\Phi(\cdot, y)\|_{\mathcal{F}}.$$

□

If a Hilbert space has a reproducing kernel then the point-wise convergence is necessary for the convergence in norm.

Theorem 1.9 *Suppose that* $\mathcal{F}$ *is a Hilbert space of functions* $f : \Omega \to \mathbb{R}$ *with reproducing kernel* Φ*. Then we have*

- $\Phi(x, y) = \langle \Phi(\cdot, x), \Phi(\cdot, y) \rangle_{\mathcal{F}} = \langle \delta_x, \delta_y \rangle_{\mathcal{F}^*}$ *for* $x, y \in \Omega$*,*
- $\Phi(x, y) = \Phi(y, x)$ *for* $x, y \in \Omega$*,*
- *if* $f, \{f_n\}_{n \in \mathbb{N}}$ *are functions in* $\mathcal{F}$ *such that* $f_n \xrightarrow[n \to +\infty]{\|\cdot\|_{\mathcal{F}}} f$ *then* $f_n(y) \xrightarrow[n \to +\infty]{} f(y)$ *for* $y \in \Omega$*.*

Proof
For the point evaluation functional the Riesz representation (Theorem 1.6) $F : \mathcal{F}^* \to \mathcal{F}$ is explicit: $\delta_y(f) = f(y) = \langle f, \Phi(\cdot, y) \rangle_{\mathcal{F}}$, which implies $F(\delta_y) = \Phi(\cdot, y)$. Since F is an isometry we obtain

$$\langle \delta_x, \delta_y \rangle_{\mathcal{F}^*} = \langle F(\delta_x), F(\delta_y) \rangle_{\mathcal{F}} = \langle \Phi(\cdot, x), \Phi(\cdot, y) \rangle_{\mathcal{F}} = \Phi(y, x).$$

The last inequality follows from Definition 1.4. Since $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ is symmetric we can obtain (1) and (2). Property (3) is a consequence of Definition 1.4 and the Cauchy–Schwarz inequality:

$$|f_n(x) - f(x)| = |\langle f_n - f, \Phi(\cdot, x) \rangle| \leq \|f_n - f\|_{\mathcal{F}} \|\Phi(\cdot, x)\|_{\mathcal{F}}.$$

□

The following result connects reproducing-kernel Hilbert spaces and positive definite kernels.

Definition 1.5 *A continuous function $\Phi : \mathbb{R}^d \to \mathbb{C}$ is called positive semi-definite if, for all $N \in \mathbb{N}$, all sets of pairwise distinct centers $X = \{x_1, \dots, x_N\} \subseteq \mathbb{R}^d$, and all $\alpha \in \mathbb{C}^N$, the quadratic form*

$$\sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k)$$

is non-negative. The function Φ is called positive definite if the quadratic form is positive for all $\alpha \in \mathbb{C}^N \setminus \{0\}$.
If we restrict to even real-valued function we can check that the real matrix $A_{\Phi,X} = (\Phi(x_j - x_k))_{1\leq j,k\leq N} \in M_N(\mathbb{R})$ is positive (semi-)definite [13].

In this work we will use even real-valued positive definite functions and this more general definition allow us to use techniques such as Fourier transforms more naturally.

Theorem 1.10 *Suppose $\mathcal{F}$ is a reproducing-kernel Hilbert function space with reproducing kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$. Then Φ is positive semi-definite. Moreover, Φ is positive definite if and only if the point evaluation functionals are linearly independent in $\mathcal{F}^*$.*

Proof
If $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is a set of pairwise distinct centers and $\alpha \in \mathbb{R}^N \setminus \{0\}$, from Theorem 1.9 we get

$$\sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \alpha_k \Phi(x_j, x_k) = \sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \alpha_k \langle \delta_{x_j}, \delta_{x_k} \rangle_{\mathcal{F}^*} =$$
$$= \left\langle \sum_{j=1}^{N} \alpha_j \delta_{x_j}, \sum_{k=1}^{N} \alpha_k \delta_{x_k} \right\rangle_{\mathcal{F}^*} = \left\| \sum_{j=1}^{N} \alpha_j \delta_{x_j} \right\|_{\mathcal{F}^*}^2 \geq 0.$$

We observe that

$$\sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \alpha_k \Phi(x_j, x_k) = \left\| \sum_{j=1}^{N} \alpha_j \delta_{x_j} \right\|_{\mathcal{F}^*}^2 = 0 \Leftrightarrow \sum_{j=1}^{N} \alpha_j \delta_{x_j} = 0,$$

which means that $\{\delta_{x_i}\}_{i=1,\dots,N}$ are linearly dependent.

□

This result shows us that the reproducing kernel of a function space $\mathcal{F}$ leads to a real-valued positive semi-definite kernel. From Definition 1.4, we know that $\mathcal{F}$ contains all functions of the form $f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j)$ with $\{x_j\}_{j=1,\dots,N} \subseteq \Omega$. With Theorem 1.9, we have

$$\begin{aligned} \|f\|_{\mathcal{F}}^2 &= \left\langle \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^{N} \alpha_k \Phi(\cdot, x_k) \right\rangle_{\mathcal{F}} = \\ &= \sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \alpha_k \langle \Phi(\cdot, x_j), \Phi(\cdot, x_k) \rangle_{\mathcal{F}} = \sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j \alpha_k \Phi(x_j, x_k). \end{aligned} \tag{1.1}$$

This expression will be meaningful when we will construct a reproducing-kernel Hilbert space for a given positive definite kernel.

The following result on the invariance properties of the Hilbert space $\mathcal{F}$ show that radial basis functions arise quite naturally within the framework of reproducing kernels.

Definition 1.6 *Let $\mathcal{T}$ be a group of transformations $T : \Omega \to \Omega$. We say that $\mathcal{F}$ is invariant under the group $\mathcal{T}$ if*

- *$f \circ T \in \mathcal{F}$ for all $f \in \mathcal{F}$ and $T \in \mathcal{T}$,*
- *$\langle f \circ T, g \circ T \rangle_{\mathcal{F}} = \langle f, g \rangle_{\mathcal{F}}$ for all $f, g \in \mathcal{F}$ and $T \in \mathcal{T}$.*

The reproducing kernel of $\mathcal{F}$ inherits the invariance of the function space.

Theorem 1.11 *Suppose that the reproducing-kernel Hilbert function space $\mathcal{F}$ is invariant under the transformations of $\mathcal{T}$, then the reproducing kernel Φ satisfies*

$$\Phi(T(x), T(y)) = \Phi(x, y) \text{ for all } x, y \in \Omega \text{ and all } T \in \mathcal{T}.$$

Proof
From Definition 1.4 and 1.6 we get

$$f(y) = f \circ T^{-1}(T(y)) = \langle f \circ T^{-1}, \Phi(\cdot, T(y)) \rangle_{\mathcal{F}} = \langle f, \Phi(T(\cdot), T(y)) \rangle_{\mathcal{F}}, \quad y \in \Omega.$$

Since $\Phi(T(\cdot), T(y)) = \Phi(\cdot, T(y)) \circ T \in \mathcal{F}$ and Proposition 1.7 we can write $\Phi(x, y) = \Phi(T(x), T(y))$ for all $x, y \in \Omega$.

□

Example 1.12 *Suppose $\Omega = \mathbb{R}^d$. Let $\mathcal{T}$ be the group of translations on $\mathbb{R}^d$. If we choose the translation $T(\xi) = \xi - y$ for a fixed $y \in \mathbb{R}^d$ then Theorem 1.11 gives us*

$$\Phi_0(x - y) := \Phi(x - y, 0) = \Phi(x, y),$$

i.e. the kernel is translation invariant.

If $\Omega = \mathbb{R}^d$ and $\mathcal{T}$ consists of affine orthogonal transformations. If we fix $x, y \in \Omega$ we can find an orthogonal transformation $A \in O_d(\mathbb{R})$ such that $A(x - y) = \|x - y\|_2 \nu$, where ν is an arbitrary unit vector in $\mathbb{R}^d$. With the previous construction and Theorem 1.11 we have

$$\phi(\|x-y\|_2) := \Phi_0(\|x-y\|_2 \nu) = \Phi_0(A(x-y)) = \Phi_0(A(x)-A(y)) = \Phi(A(x), A(y)) = \Phi(x, y),$$

i.e. Φ is radial.

1.3 Inner product space completion

Following the results in [14] we introduce the completion of a inner product space. This argument will be useful to give examples of reproducing-kernel Hilbert spaces starting from positive-definite kernels. The goal of this section is to give a proof of the following theorem.

Theorem 1.13 *(Completion) If $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ is any inner product space, then there exists a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ and a map $U : \mathcal{V} \to \mathcal{H}$ such that*

- *U is injective*
- *U is linear*
- $\langle U(x), U(y) \rangle_{\mathcal{H}} = \langle x, y \rangle_{\mathcal{V}}$
- *$U(\mathcal{V}) = \{U(x) : x \in \mathcal{V}\}$ is dense in $\mathcal{H}$. If $\mathcal{V}$ is complete, then $U(\mathcal{V}) = \mathcal{H}$.*

$\mathcal{H}$ is called the completion of $\mathcal{V}$. U gives a isometric correspondence between elements of $\mathcal{V}$ and elements of $U(\mathcal{V})$ so we can think of $\mathcal{V}$ as being $U(\mathcal{V}) \subseteq \mathcal{H}$. Using this point of view Theorem 1.13 says that any inner product space can be extended to a complete inner product space, i.e. it can have its holes filled in.

To guess the shape of $\mathcal{H}$ we use a backwards strategy. Suppose that, somehow, we have found a suitable $\mathcal{H}$ with $\mathcal{V} \subseteq \mathcal{H}$. Because $\mathcal{V}$ is dense in $\mathcal{H}$, each element of $\mathcal{H}$ can be written as the limit of a sequence in $\mathcal{V}$, and each such sequence in Cauchy. Thus specifying an element of $\mathcal{H}$ is equivalent to specify a Cauchy sequence in $\mathcal{V}$. But there is not a one-to-one correspondence between elements of $\mathcal{H}$ and Cauchy sequences in $\mathcal{V}$, because many different Cauchy sequences in $\mathcal{V}$ can converge to the same element of $\mathcal{H}$. To get a one-to-one correspondence, we can identify each $x \in \mathcal{H}$ with the set of all Cauchy sequences in $\mathcal{V}$ that converges to x.

First we define the set of all Cauchy sequences in $\mathcal{V}$ as

$$\mathcal{V}' = \left\{ \{x_n\}_{n\in\mathbb{N}} : \{x_n\}_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } \mathcal{V} \right\}.$$

Next we define two Cauchy sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ in $\mathcal{V}$ to be equivalent, written $\{x_n\}_{n\in\mathbb{N}} \sim \{y_n\}_{n\in\mathbb{N}}$, if and only if

$$\lim_{n\to\infty} \|x_n - y_n\|_{\mathcal{V}} = 0.$$

From this expression we can exploit that any two convergent sequences have the same limit if and only if they are equivalent: suppose $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ have as limit x and y respectively.

($\Rightarrow$) Suppose $x = y$: $\|x_n - y_n\|_{\mathcal{V}} \leq \|x_n - x\|_{\mathcal{V}} + \|y - y_n\|_{\mathcal{V}} \xrightarrow{n\to\infty} 0$.

($\Leftarrow$) Suppose $\{x_n\}_{n\in\mathbb{N}} \sim \{y_n\}_{n\in\mathbb{N}}$: $\|x-y\|_{\mathcal{V}} \leq \|x-x_n\|_{\mathcal{V}} + \|x_n - y_n\|_{\mathcal{V}} + \|y_n - y\|_{\mathcal{V}} \xrightarrow{n\to\infty} 0$, which implies $x = y$.

Next, if $\{x_n\}_{n\in\mathbb{N}} \in \mathcal{V}'$, we define the equivalence class of $\{x_n\}_{n\in\mathbb{N}}$ to be the set of all Cauchy sequence that are equivalent to $\{x_n\}_{n\in\mathbb{N}}$:

$$[\{x_n\}_{n\in\mathbb{N}}] = \{\{y_n\}_{n\in\mathbb{N}} \in \mathcal{V}' : \{x_n\}_{n\in\mathbb{N}} \sim \{y_n\}_{n\in\mathbb{N}}\}.$$

This definition makes sense because $\sim$ is an equivalence relation.

Proposition 1.14 *$\sim$ is an equivalence relation. In particular, if $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \in \mathcal{V}'$ then either $[\{x_n\}_{n\in\mathbb{N}}] = [\{y_n\}_{n\in\mathbb{N}}]$ or $[\{x_n\}_{n\in\mathbb{N}}] \cap [\{y_n\}_{n\in\mathbb{N}}] = \emptyset$.*

Proof
(1) Reflexivity:
$\{x_n\}_{n\in\mathbb{N}} \sim \{x_n\}_{n\in\mathbb{N}}$ because $\|x_n - x_n\|_{\mathcal{V}} = 0$ for each $n \in \mathbb{N}$.

(2) Symmetry:
If $\{x_n\}_{n\in\mathbb{N}} \sim \{y_n\}_{n\in\mathbb{N}} \Rightarrow \lim_{n\to\infty} \|x_n - y_n\|_{\mathcal{V}} = 0 \Rightarrow \lim_{n\to\infty} \|y_n - x_n\|_{\mathcal{V}} = 0 \Rightarrow \{y_n\}_{n\in\mathbb{N}} \sim \{x_n\}_{n\in\mathbb{N}}$.

(3) Transitivity: If $\{x_n\}_{n\in\mathbb{N}} \sim \{y_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}} \sim \{z_n\}_{n\in\mathbb{N}} \Rightarrow \|x_n - z_n\|_{\mathcal{V}} \leq \|x_n - y_n\|_{\mathcal{V}} + \|y_n - z_n\|_{\mathcal{V}} \xrightarrow{n\to\infty} 0$.

The last property in the statement is true because equivalence classes of an equivalence relation provide a partition of $\mathcal{V}'$.

□

The Proposition 1.14 suggests us to define

$$\mathcal{H} = \left\{[\{x_n\}_{n\in\mathbb{N}}] : \{x_n\}_{n\in\mathbb{N}} \in \mathcal{V}'\right\}. \tag{1.2}$$

Now the goal is to define on $\mathcal{H}$ a Hilbert space structure.

Proposition 1.15 *If $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \in \mathcal{V}'$ then $\lim_{n\to\infty}\langle x_n, y_n\rangle_{\mathcal{V}}$ exists.*

Proof
We first observe that

$$\begin{aligned}
|\langle x_n, y_n\rangle_{\mathcal{V}} - \langle x_m, y_m\rangle_{\mathcal{V}}| &\leq |\langle x_n, y_n\rangle_{\mathcal{V}} - \langle x_m, y_n\rangle_{\mathcal{V}}| + |\langle x_m, y_n\rangle_{\mathcal{V}} - \langle x_m, y_m\rangle_{\mathcal{V}}| \leq \\
&\leq |\langle x_n - x_m, y_n\rangle_{\mathcal{V}}| + |\langle x_m, y_n - y_m\rangle_{\mathcal{V}}| \\
&\leq \|x_n - x_m\|_{\mathcal{V}}\|y_n\|_{\mathcal{V}} + \|x_m\|_{\mathcal{V}}\|y_n - y_m\|_{\mathcal{V}}
\end{aligned}$$

Since $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ are Cauchy sequences then $\{\|x_n\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ and $\{\|y_n\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ are bounded (from triangle inequality we obtain $|\|x_n\|_{\mathcal{V}} - \|x_m\|_{\mathcal{V}}| \leq \|x_n - x_m\|_{\mathcal{V}}$, so $\{\|x_n\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ is a bounded Cauchy sequence in $\mathbb{R}$) and $\{\|x_n - x_m\|_{\mathcal{V}}\}_{n,m\in\mathbb{N}}$ and $\{\|y_n - y_m\|_{\mathcal{V}}\}_{n,m\in\mathbb{N}}$ converge to 0 as $n, m \to +\infty$. We proved that $\{\langle x_n, y_n\rangle_{\mathcal{V}}\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, the considered sequence converges.

□

Proposition 1.16 *Define, for each* $[\{x_n\}_{n\in\mathbb{N}}], [\{y_n\}_{n\in\mathbb{N}}] \in \mathcal{H}$ *and* $\alpha \in \mathbb{R}$,

$$\begin{aligned}
[\{x_n\}_{n\in\mathbb{N}}] + [\{y_n\}_{n\in\mathbb{N}}] &= [\{x_n + y_n\}_{n\in\mathbb{N}}],\\
\alpha[\{x_n\}_{n\in\mathbb{N}}] &= [\{\alpha x_n\}_{n\in\mathbb{N}}],\\
\langle [\{x_n\}_{n\in\mathbb{N}}], [\{y_n\}_{n\in\mathbb{N}}]\rangle_{\mathcal{H}} &= \lim_{n\to\infty} \langle x_n, y_n\rangle_{\mathcal{V}}.
\end{aligned}$$

Each of these operations is well-defined.

Proof
Let $\{x_n\}_{n\in\mathbb{N}} \sim \{x_n^*\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}} \sim \{y_n^*\}_{n\in\mathbb{N}}$. We obtain

$$\begin{aligned}
|\langle x_n, y_n\rangle_{\mathcal{V}} - \langle x_n^*, y_n^*\rangle_{\mathcal{V}}| &\le |\langle x_n, y_n\rangle_{\mathcal{V}} - \langle x_n^*, y_n\rangle_{\mathcal{V}}| + |\langle x_n^*, y_n\rangle_{\mathcal{V}} - \langle x_n^*, y_n^*\rangle_{\mathcal{V}}| \le\\
&\le |\langle x_n - x_n^*, y_n\rangle_{\mathcal{V}}| + |\langle x_n^*, y_n - y_n^*\rangle_{\mathcal{V}}|\\
&\le \|x_n - x_n^*\|_{\mathcal{V}}\|y_n\|_{\mathcal{V}} + \|x_n^*\|_{\mathcal{V}}\|y_n - y_n^*\|_{\mathcal{V}} \xrightarrow{n\to+\infty} 0.
\end{aligned}$$

The limit relation holds because $\{\|y_n\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ and $\{\|x_n^*\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ are bounded (Proposition 1.15) and $\{\|x_n - x_n^*\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ and $\{\|y_n - y_n^*\|_{\mathcal{V}}\}_{n\in\mathbb{N}}$ converge to 0. To conclude the proof we need the following:

$$\begin{aligned}
\|(x_n + y_n) - (x_n^* + y_n^*)\|_{\mathcal{V}} &\le \|x_n - x_n^*\|_{\mathcal{V}} + \|y_n - y_n^*\|_{\mathcal{V}} \xrightarrow{n\to+\infty} 0\\
\|\alpha x_n - \alpha x_n^*\|_{\mathcal{V}} &\le |\alpha|\|x_n - x_n^*\|_{\mathcal{V}} \xrightarrow{n\to+\infty} 0.
\end{aligned}$$

□

Proposition 1.17 $\mathcal{H}$ *with the operations of Proposition 1.16 is a inner product space.*

Proof
We need to check that $\mathcal{H}$ is a $\mathbb{R}$-vector space and that $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ is a positive definite bilinear form. Let us check the positive definiteness of $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. If $[\{x_n\}_{n\in\mathbb{N}}] \in \mathcal{H}$ and $\langle [\{x_n\}_{n\in\mathbb{N}}], [\{x_n\}_{n\in\mathbb{N}}]\rangle_{\mathcal{H}} = 0$ then

$$0 = \lim_{n\to\infty} \langle x_n, x_n\rangle_{\mathcal{V}} = \lim_{n\to\infty} \|x_n\|_{\mathcal{V}}^2 \Rightarrow 0 = \lim_{n\to\infty} \|x_n\|_{\mathcal{V}} = \lim_{n\to\infty} \|x_n - 0\|_{\mathcal{V}}.$$

We proved that $\{x_n\}_{n\in\mathbb{N}} \sim \{0\}_{n\in\mathbb{N}} \Rightarrow [\{x_n\}_{n\in\mathbb{N}}] = [\{0\}_{n\in\mathbb{N}}]$.

□

With the results of Proposition 1.16 and Proposition 1.17 we can state the useful result

Proposition 1.18 $\mathcal{H}$ *is complete.*

Proof
Let $\{X^{(n)}\}_{n\in\mathbb{N}} \subseteq \mathcal{H}$ be a Cauchy sequence. We need to prove that it has a limit $X \in \mathcal{H}$. Each $X^{(n)}$ is an equivalence class of Cauchy sequence in $\mathcal{V}$. We will use the following notation : $X^{(n)} = [\{x_\ell^{(n)}\}_{\ell\in\mathbb{N}}]$. We will use a classical diagonal argument.

$$\{x_\ell^{(1)}\}_{\ell\in\mathbb{N}} \text{ Cauchy } \Rightarrow \exists\, \ell_1 \in \mathbb{N} \text{ such that } \forall\, \ell \geq \ell_1 \; \|x_\ell^{(1)} - x_{\ell_1}^{(1)}\|_\mathcal{V} < 1. \text{ Choose } x_1 = x_{\ell_1}^{(1)}.$$
$$\{x_\ell^{(2)}\}_{\ell\in\mathbb{N}} \text{ Cauchy } \Rightarrow \exists\, \ell_2 > \ell_1 \text{ such that } \forall\, \ell \geq \ell_2 \; \|x_\ell^{(2)} - x_{\ell_2}^{(2)}\|_\mathcal{V} < \frac{1}{2}. \text{ Choose } x_2 = x_{\ell_2}^{(2)}.$$
$$\{x_\ell^{(3)}\}_{\ell\in\mathbb{N}} \text{ Cauchy } \Rightarrow \exists\, \ell_3 > \ell_2 \text{ such that } \forall\, \ell \geq \ell_3 \; \|x_\ell^{(3)} - x_{\ell_3}^{(3)}\|_\mathcal{V} < \frac{1}{3}. \text{ Choose } x_3 = x_{\ell_3}^{(3)}.$$
$$\vdots$$

The example below shows that at each step we choose an element of a Cauchy sequence.

$$\begin{array}{ccccc}
X^{(1)} & X^{(2)} & X^{(3)} & X^{(4)} & \cdots \\
x_1^{(1)}\bullet & x_1^{(2)}\circ & x_1^{(3)}\circ & x_1^{(4)}\circ & \\
x_2^{(1)}\circ & x_2^{(2)}\bullet & x_2^{(3)}\circ & x_2^{(4)}\circ & \\
x_3^{(1)}\circ & x_3^{(2)}\circ & x_3^{(3)}\bullet & x_3^{(4)}\circ & \\
x_4^{(1)}\circ & x_4^{(2)}\circ & x_4^{(3)}\circ & x_4^{(4)}\bullet & \\
\vdots & & & &
\end{array}$$

Now we prove that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. Let $\varepsilon > 0$ by the triangle inequality we have

$$\|x_n - x_m\|_\mathcal{V} \leq \|x_{\ell_n}^{(n)} - x_\ell^{(n)}\|_\mathcal{V} + \|x_\ell^{(n)} - x_\ell^{(m)}\|_\mathcal{V} + \|x_\ell^{(m)} - x_{\ell_m}^{(m)}\|_\mathcal{V} = \tag{1.3}$$
$$= \|x_{\ell_n}^{(n)} - x_\ell^{(n)}\|_\mathcal{V} + \Big(\|x_\ell^{(n)} - x_\ell^{(m)}\|_\mathcal{V} - \|X^{(n)} - X^{(m)}\|_\mathcal{H}\Big) + \|X^{(n)} - X^{(m)}\|_\mathcal{H} + \|x_\ell^{(m)} - x_{\ell_m}^{(m)}\|_\mathcal{V}$$

for any $\ell \in \mathbb{N}$.

- If $\ell \geq \ell_n$ the first term is smaller than $\frac{1}{n}$.
- By Proposition 1.16, $\|X^{(n)} - X^{(m)}\|_\mathcal{H} = \lim_{\ell\to\infty} \|x_\ell^{(n)} - x_\ell^{(m)}\|_\mathcal{V}$. So there exists a natural number $N_{n,m}$ such that the second term is smaller than $\frac{\varepsilon}{4}$ whenever $\ell \geq N_{n,m}$.
- By hypothesis, the sequence $\{X^{(n)}\}_{n\in\mathbb{N}}$ is a Cauchy sequence. So there exists a natural number $\tilde{N}$ such that the third term is smaller than $\frac{\varepsilon}{4}$ whenever $n, m \geq \tilde{N}$.
- If $\ell \geq \ell_m$ the last term is smaller than $\frac{1}{m}$.

Choose any natural number $N \geq \max\{\tilde{N}, \frac{4}{\varepsilon}\}$. We want to prove that $\|x_n - x_m\|_{\mathcal{V}} < \varepsilon$ for $n, m \geq N$. Fix such n, m and choose ℓ greater than $\max\{N_{n,m}, \ell_n, \ell_m\}$. Since $\frac{1}{n} \leq \frac{1}{N} \leq \frac{\varepsilon}{4}$ with the previous choice the claim holds because the four term in equation (1.3) are each smaller than $\frac{\varepsilon}{4}$.

Now we claim that $X \coloneqq [\{x_n\}_{n\in\mathbb{N}}] = \lim_{n\to\infty} X^{(n)}$. Let $\varepsilon > 0$. By Proposition 1.16 we get

$$\|X - X^{(n)}\|_{\mathcal{H}} = \lim_{m\to\infty} \|x_m - x_m^{(n)}\|_{\mathcal{V}} = \lim_{m\to\infty} \|x_{\ell_m}^{(m)} - x_m^{(n)}\|_{\mathcal{V}}.$$

By the triangle inequality we have

$$\|x_{\ell_m}^{(m)} - x_m^{(n)}\|_{\mathcal{V}} \leq \|x_{\ell_m}^{(m)} - x_{\ell_n}^{(n)}\|_{\mathcal{V}} + \|x_{\ell_n}^{(n)} - x_m^{(n)}\|_{\mathcal{V}}. \tag{1.4}$$

- Since the sequence $\{x_n = x_{\ell_n}^{(n)}\}_{n\in\mathbb{N}}$ is a Cauchy sequence, there exists $N' \in \mathbb{N}$ such that the first term is smaller than $\frac{\varepsilon}{2}$ for $n, m \geq N'$.
- The second term is smaller than $\frac{1}{n}$ whenever $m \geq \ell_n$.

To complete the proof we choose $N \geq \max\{N', \frac{2}{\varepsilon}\}$. We claim $\|X - X^{(n)}\|_{\mathcal{H}} < \varepsilon$ if $n \geq N$. Let $n \geq N$. Since $\frac{1}{n} \leq \frac{1}{N} \leq \frac{\varepsilon}{2}$, for all m bigger than $\max\{N', \ell_n\}$, the two terms in equation (1.4) are each smaller than $\frac{\varepsilon}{2}$.

□

Finally with the following construction we can prove Theorem 1.13.

We define $U : \mathcal{V} \to \mathcal{H}$ by

$$U(x) = [\{x, x, x, \dots, x, \dots\}]. \tag{1.5}$$

Proof of Theorem 1.13

1. U is linear and $\langle U(x), U(y)\rangle_{\mathcal{H}} = \langle x, y\rangle_{\mathcal{V}}$ for all $x, y \in \mathcal{V}$.

   The linearity of U follow from equation (1.5) and Proposition 1.16.

   $$\langle U(x), U(y)\rangle_{\mathcal{H}} = \langle [\{x, x, x, \dots\}], [\{y, y, y, \dots\}]\rangle_{\mathcal{H}} = \lim_{n\to\infty} \langle x, y\rangle_{\mathcal{V}} = \langle x, y\rangle_{\mathcal{V}}.$$

2. U is injective.

   $$U(x) = U(y) \Leftrightarrow [\{x, x, x, \dots\}] = [\{y, y, y, \dots\}] \Leftrightarrow \lim_{n\to\infty} \|x - y\|_{\mathcal{V}} = 0 \Leftrightarrow x = y.$$

3. $U(\mathcal{V})$ is dense in $\mathcal{H}$.

   Let $X = [\{x_n\}_{n\in\mathbb{N}}] \in \mathcal{H}$. We claim that the sequence $\{U(x_n)\}_{n\in\mathbb{N}}$ converges in $\mathcal{H}$ to X. By Proposition 1.16

   $$\|X - U(x_n)\|_{\mathcal{H}} = \lim_{m\to\infty} \|x_m - (U(x_n))_m\|_{\mathcal{V}} = \lim_{m\to\infty} \|x_m - x_n\|_{\mathcal{V}}.$$

   Since $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{V}$, the expression above converges to 0 as $n \to +\infty$.

4. If $\mathcal{V}$ is complete, then $U(\mathcal{V}) = \mathcal{H}$.

   Fix $X \in \mathcal{H}$. We want to find $x \in \mathcal{V}$ with $U(x) = X$. By (3) we obtained that $U(\mathcal{V})$ is dense in $\mathcal{H}$, so there exists $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathcal{V}$ such that

$$
\begin{aligned}
& X = \lim_{n\to\infty} U(x_n) \\
\Rightarrow \quad & \{U(x_n)\}_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } \mathcal{H} \\
\Rightarrow \quad & \text{By (1) } \{x_n\}_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } \mathcal{V} \\
\Rightarrow \quad & \text{Since } \mathcal{V} \text{ is complete, there exists } \lim_{n\to\infty} x_n = x \in \mathcal{V} \\
\Rightarrow \quad & \text{By (1) } \lim_{n\to\infty} U(x_n) = U(x) \\
\Rightarrow \quad & U(x) = X \text{ because a metric space is Hausdorff (T}_2\text{ space).}
\end{aligned}
$$

□

1.4 Native spaces for positive definite kernels

From Theorem 1.10 we proved that the reproducing kernel of a Hilbert function space is a positive definite function. In our context it is useful to start with a positive definite function and it becomes necessary to find the associated function space (i.e. find the Hilbert function space that has this fixed kernel as its reproducing kernel).

We start with a symmetric positive definite kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$. Studying equation (1.1) it is natural to define the following vector space

$$F_\Phi(\Omega) := \langle \Phi(\cdot, y) : y \in \Omega \rangle, \tag{1.6}$$

where $\{\Phi(\cdot, y)\}_{y\in\Omega}$ are linearly independent. Let $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ a set of pairwise distinct centers, if we suppose that there esists $\alpha \in \mathbb{R}^N$ such that

$$\sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) = 0 \Rightarrow \sum_{k=1}^{N} \alpha_k \underbrace{\sum_{j=1}^{N} \alpha_j \Phi(x_k, x_j)}_{=0} = \sum_{k=1}^{N} \sum_{j=1}^{N} \alpha_k \alpha_j \Phi(x_k, x_j) = 0,$$

that forces α to be 0.

We can equip the space defined in equation (1.6) with the bilinear form

$$\left\langle \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^{M} \beta_k \Phi(\cdot, y_k) \right\rangle_\Phi = \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(x_j, y_k). \tag{1.7}$$

Theorem 1.19 *If $\Phi : \Omega \times \Omega \to \mathbb{R}$ is a symmetric positive definite kernel then the bilinear form $\langle \cdot, \cdot \rangle_\Phi$ defined in equation (1.7) is a inner product on $F_\Phi(\Omega)$. Furthermore, $F_\Phi(\Omega)$ is a inner product space with reproducing kernel Φ.*

Proof
The symmetry of $\langle \cdot, \cdot \rangle_\Phi$ immediately follows from the symmetry of Φ. Let us prove that $\langle \cdot, \cdot \rangle_\Phi$ is positive definite: fix $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \neq 0$, which means that $\alpha \in \mathbb{R}^N \setminus \{0\}$ if $x_j \neq x_i$ for each $j \neq i$.

$$\langle f, f \rangle_\Phi = \left\langle \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^N \alpha_k \Phi(\cdot, x_k) \right\rangle_\Phi = \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \Phi(x_j, x_k) > 0,$$

because Φ is a positive definite kernel. The reproducing kernel property follows from

$$\langle f, \Phi(\cdot, y) \rangle_\Phi = \left\langle \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j), \Phi(\cdot, y) \right\rangle_\Phi = \sum_{j=1}^N \alpha_j \Phi(x_j, y) = \sum_{j=1}^N \alpha_j \Phi(y, x_j) = f(y).$$

□

With the construction provided in section 1.3 we can isometrically embed the inner product space $F_\Phi(\Omega)$ in the Hilbert space $\mathcal{F}_\Phi(\Omega)$. We note that with the construction provide in section 1.3 $\mathcal{F}_\Phi(\Omega)$ does not contain functions from Ω to $\mathbb{R}$. To interpret this as a candidate Hilbert function space with reproducing kernel Φ we will use a continuous extension of the point-evaluation functionals (compare Theorem 1.8).

By the reproducing property of Φ (Theorem 1.19) the point-evaluation functional δ_x is continuous on $F_\Phi(\Omega)$:

$$|\delta_x(f)| = |f(x)| = |\langle f, \Phi(\cdot, x) \rangle_\Phi| \leq \|f\|_\Phi \|\Phi(\cdot, x)\|_\Phi.$$

Since $\mathcal{F}_\Phi(\Omega) = \overline{F_\Phi(\Omega)}$ (Theorem 1.13) is a Hilbert space, the point-evaluation functional δ_x with $x \in \Omega$ can be uniquely extend as a continuous functional on $\mathcal{F}_\Phi(\Omega)$.

To be more precise and clear we define the liner map $R : \mathcal{F}_\Phi(\Omega) \to \mathcal{C}(\Omega)$

$$\begin{aligned} R(f) : \Omega &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle f, \Phi(\cdot, x) \rangle_\Phi. \end{aligned} \tag{1.8}$$

We stated that $R(f)(\cdot)$ is a continuous function in Ω because

$$|R(f)(x) - R(f)(y)| = |\langle f, \Phi(\cdot, x) - \Phi(\cdot, y) \rangle_\Phi| \leq \|f\|_\Phi \|\Phi(\cdot, x) - \Phi(\cdot, y)\|_\Phi.$$

From continuity and symmetry of Φ with equation (1.7) we can conclude

$$\|\Phi(\cdot, x) - \Phi(\cdot, y)\|_\Phi^2 = \Phi(x, x) + \Phi(y, y) - 2\Phi(x, y).$$

The function $R : \mathcal{F}_\Phi(\Omega) \to \mathcal{C}(\Omega)$ let us to think the elements of $\mathcal{F}_\Phi(\Omega)$ as continuous functions on Ω, indeed

$$R(f)(x) = \langle f, \Phi(\cdot, x) \rangle_\Phi = f(x) \tag{1.9}$$

for each $f \in F_\Phi(\Omega)$. To make this last observation precise we need the following

Theorem 1.20 *The linear map $R : \mathcal{F}_\Phi(\Omega) \to \mathcal{C}(\Omega)$ defined by equation (1.8) is injective.*

Proof

The map R is linear, so we check that $\ker(R) = 0$. Suppose $R(f) = 0$ for $f \in \mathcal{F}_\Phi(\Omega)$, which means that $\langle f, \Phi(\cdot, x)\rangle_\Phi = 0$ for all $x \in \Omega$. Such $f \in F_\Phi(\Omega)^\perp = \overline{F_\Phi(\Omega)}^\perp = \mathcal{F}_\Phi(\Omega)^\perp = \langle 0\rangle$. This proves that $f = 0$.

□

Definition 1.7 *The native Hilbert function space corresponding to the symmetric positive definite kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ is defined by*

$$\mathcal{N}_\Phi(\Omega) \coloneqq R(\mathcal{F}_\Phi(\Omega)).$$

Where it is defined a inner product inherited from $\mathcal{F}_\Phi(\Omega)$ by

$$\langle f, g\rangle_{\mathcal{N}_\Phi(\Omega)} \coloneqq \langle R^{-1}(f), R^{-1}(g)\rangle_\Phi.$$

To justify the previous definition we can note that $\Phi(\cdot, x) = R(\Phi(\cdot, x))$ because of equation (1.9). Moreover, if $f \in \mathcal{N}_\Phi(\Omega)$ and $x \in \Omega$

$$\begin{aligned}\langle f, \Phi(\cdot, x)\rangle_{\mathcal{N}_\Phi(\Omega)} &= \langle R^{-1}(f), R^{-1}(\Phi(\cdot, x))\rangle_\Phi = \\ &= \langle R^{-1}(f), \Phi(\cdot, x)\rangle_\Phi = R(R^{-1}(f))(x) = f(x).\end{aligned}$$

With Theorem 1.10 and Definition 1.7 we established a connection between positive definite kernels and reproducing kernels of Hilbert function spaces.
If we recall from equation (1.9) that $R(F_\Phi(\Omega)) = F_\Phi(\Omega)$ and that

$$R : (\mathcal{F}_\Phi(\Omega), \langle\cdot,\cdot\rangle_\Phi) \longrightarrow (\mathcal{N}_\Phi(\Omega), \langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega)}) \tag{1.10}$$

is an isometry, then $F_\Phi(\Omega)$ is dense in $\mathcal{N}_\Phi(\Omega)$ and

$$\|f\|_\Phi = \|f\|_{\mathcal{N}_\Phi(\Omega)} \tag{1.11}$$

for all $f \in F_\Phi(\Omega)$.

Because of Proposition 1.7 it is interesting to investigate the uniqueness of the native space.

Theorem 1.21 *Suppose that Φ is a symmetric positive definite kernel and that $\mathcal{G}$ is a Hilbert space of function $f : \Omega \to \mathbb{R}$ with reproducing kernel Φ. Then $\mathcal{G}$ is the native space $\mathcal{N}_\Phi(\Omega)$ and the inner products are the same.*

Proof

From definition of reproducing kernel (Definition 1.4) and equation (1.1) we know that $F_\Phi(\Omega) \subseteq \mathcal{G}$ and $\|f\|_\mathcal{G} = \|f\|_\Phi = \|f\|_{\mathcal{N}_\Phi(\Omega)}$ for all $f \in F_\Phi(\Omega)$ (see also equation (1.11)).

Now fix $f \in \mathcal{N}_\Phi(\Omega)$, then there exists a Cauchy sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq F_\Phi(\Omega)$ converging to f in $\mathcal{N}_\Phi(\Omega)$. By Theorem 1.9 we have that

$$f(y) = \lim_{n\to\infty} f_n(y).$$

Since $\{f_n\}_{n\in\mathbb{N}} \subseteq F_\Phi(\Omega)$ is converging in $\mathcal{N}_\Phi(\Omega)$ then it is also Cauchy in $\mathcal{G}$. This sequence converges to $g \in \mathcal{G}$ because it is complete. The reproducing kernel property of Theorem 1.9 gives

$$g(y) = \lim_{n\to\infty} f_n(y) = f(y)$$

for each $y \in \Omega$. This proves $\mathcal{N}_\Phi(\Omega) \subseteq \mathcal{G}$ and in particular

$$\|f\|_\mathcal{G} = \|g\|_\mathcal{G} = \lim_{n\to\infty} \|f_n\|_\mathcal{G} = \lim_{n\to\infty} \|f_n\|_{\mathcal{N}_\Phi(\Omega)} = \|f\|_{\mathcal{N}_\Phi(\Omega)}$$

for each $f \in \mathcal{N}_\Phi(\Omega)$. From this properties we can state that $\mathcal{N}_\Phi(\Omega)$ is closed in $\mathcal{G}$ ($\mathcal{N}_\Phi(\Omega)$ is an Hilbert space and $\|\cdot\|_\mathcal{G}$ and $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$ coincide on $\mathcal{N}_\Phi(\Omega)$).

Suppose by contradiction that $\mathcal{N}_\Phi(\Omega) \subsetneq \mathcal{G}$. If $\mathcal{N}_\Phi(\Omega)^{\perp_\mathcal{G}} = \langle 0 \rangle$, then by Corollary 1.5 we will have $\mathcal{N}_\Phi(\Omega) = \overline{\mathcal{N}_\Phi(\Omega)}^\mathcal{G} = \mathcal{G}$ (it is a contradiction). In this setting we can find an element $g \in \mathcal{G} \setminus \{0\}$ orthogonal to $\mathcal{N}_\Phi(\Omega)$. With the reproduction property this means $g(y) = \langle g, \Phi(\cdot, y)\rangle_\mathcal{G} = 0$ for all $y \in \Omega$. We proved $g = 0$ and it is a contradiction. We can conclude that $\mathcal{N}_\Phi(\Omega) = \mathcal{G}$ and that $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)} = \|\cdot\|_\mathcal{G}$. We can conclude the proof by recalling an equality for normed vector spaces with norm induced by a inner product:

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

□

The following results will have a particular interest in our work because it let us to study interpolation processes in classical functional spaces.

We need to recall some classical definition and result on Fourier Transform.

Definition 1.8 *For $f \in L^1(\mathbb{R}^d)$ we define its Fourier transform [15] by*

$$\widehat{f}(x) = f^\wedge(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(\omega) e^{-i\langle x,\omega\rangle} d\omega$$

and its inverse Fourier transform by

$$f^\vee(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(\omega) e^{i\langle x,\omega\rangle} d\omega.$$

We underline the fact that with dominated convergence theorem we can prove that $\widehat{f}$ is uniformly continuous in $\mathbb{R}^d$.

Theorem 1.22 *(Plancherel) There exists an isomorphism mapping $T : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ such that:*

- $\|T(f)\|_{L^2(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)}$ *for all* $f \in L^2(\mathbb{R}^d)$,
- $T(f) = \widehat{f}$ *for all* $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$,
- $T^{-1}(f) = f^\vee$ *for all* $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

The isomorphism is uniquely determined by these properties.

Now we can state an important result.

Theorem 1.23 *Suppose that* $\Phi \in \mathcal{C}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ *is a real-valued positive definite function. Define*

$$\mathcal{G} := \left\{ f \in L^2(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d) : \frac{\widehat{f}}{\sqrt{\widehat{\Phi}}} \in L^2(\mathbb{R}^d) \right\}$$

and equip this space with the bilinear form

$$\langle f, g \rangle_{\mathcal{G}} := (2\pi)^{-d/2} \left\langle \frac{\widehat{f}}{\sqrt{\widehat{\Phi}}}, \frac{\widehat{g}}{\sqrt{\widehat{\Phi}}} \right\rangle_{L^2(\mathbb{R}^d)} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)}}{\widehat{\Phi}(\omega)} d\omega.$$

Then $\mathcal{G}$ *is a real Hilbert space with inner product* $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ *and reproducing kernel* $\Phi(\cdot - \cdot)$. *Hence* $\mathcal{G}$ *is the native space of* Φ, *i.e.* $\mathcal{G} = \mathcal{N}_\Phi(\mathbb{R}^d)$, *and both inner product coincide. In particular, every* $f \in \mathcal{G}$ *can be recovered from its Fourier transform* $\widehat{f} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

To make this result more clear we must mention:

Theorem 1.24 *Suppose that* $\Phi \in L^1(\mathbb{R}^d)$ *is a continuous function. Then* Φ *is positive definite if and only if* Φ *is bounded and its Fourier transform is non-negative and non-vanishing. Moreover,* $\widehat{\Phi} \in L^1(\mathbb{R}^d)$.

Proof of Theorem 1.23
From Theorem 1.24 we get that $\widehat{\Phi} \in L^1(\mathbb{R}^d)$. For $f \in \mathcal{G}$ this means in particular that $\widehat{f} \in L^1(\mathbb{R}^d)$ because

$$\int_{\mathbb{R}^d} |\widehat{f}(\omega)| \mathrm{d}\,\omega = \int_{\mathbb{R}^d} \frac{|\widehat{f}(\omega)|}{\sqrt{\widehat{\Phi}(\omega)}} \sqrt{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega \le \left(\int_{\mathbb{R}^d} \frac{|\widehat{f}(\omega)|^2}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} \widehat{\Phi}(\omega) \mathrm{d}\,\omega \right)^{\frac{1}{2}} < +\infty.$$

Plancherel's theorem (Theorem 1.22) and the continuity of f and $(\widehat{f})^\vee$ allow us to recover f point-wise from its Fourier transform with

$$f(x) = T^{-1}(T(f))(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} T(f)(\omega) e^{i\langle x, \omega \rangle} \mathrm{d}\,\omega = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{f}(\omega) e^{i\langle x, \omega \rangle} \mathrm{d}\,\omega.$$

where T is the unique isometry of Theorem 1.22. Naturally we can identify $T(f)$ with $\widehat{f}$.

Since $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ is $\mathbb{R}$-linear, we need to prove that it is real-valued and positive definite: for a real-valued $f \in L^2(\mathbb{R}^d)$ its fourier transfrom satisfies $\overline{\widehat{f}(x)} = \widehat{f}(-x)$ almost everywhere in $\mathbb{R}^d$. For the next computation it will be helpful to decompose $\mathbb{R}^d$ (up to negligible set) as

$$\mathbb{R}^d = \Gamma \cup -\Gamma$$

where $\Gamma = \{(\omega_1, \omega_2, \dots, \omega_d) : \omega_i \in \mathbb{R} \text{ for } i = 1, \dots, d \text{ and } \omega_1 > 0\}$. We can compute the bilinear form as

$$\int_{\mathbb{R}^d} \frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)}}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega = \int_{\Gamma} \frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)}}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega + \int_{-\Gamma} \frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)}}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega =$$

$$= \int_{\Gamma} \frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)}}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega + \underbrace{\int_{\Gamma} \frac{\widehat{f}(-\omega)\overline{\widehat{g}(-\omega)}}{\widehat{\Phi}(-\omega)} \mathrm{d}\,\omega}_{\substack{\text{change of variable:} \\ \omega \mapsto -\omega}} = \int_{\Gamma} \left(\frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)}}{\widehat{\Phi}(\omega)} + \frac{\widehat{f}(-\omega)\overline{\widehat{g}(-\omega)}}{\widehat{\Phi}(-\omega)} \right) \mathrm{d}\,\omega =$$

$$= \int_{\Gamma} \frac{\widehat{f}(\omega)\overline{\widehat{g}(\omega)} + \overline{\widehat{f}(\omega)}\widehat{g}(\omega)}{\widehat{\Phi}(\omega)} = 2 \int_{\Gamma} \frac{\Re[\widehat{f}(\omega)\overline{\widehat{g}(\omega)}]}{\widehat{\Phi}(\omega)} \in \mathbb{R}.$$

The $\mathbb{R}$-bilinear form is positive definite because:

If $\langle f, f \rangle_{\mathcal{G}} = 0 \Rightarrow \left\| \frac{\widehat{f}}{\sqrt{\widehat{\Phi}}} \right\|_{L^2(\mathbb{R}^d)} = 0 \Rightarrow \frac{\widehat{f}}{\sqrt{\widehat{\Phi}}} = 0$ a.e. in $\mathbb{R}^d \Rightarrow \widehat{f} = 0$ a.e. in $\mathbb{R}^d$, we can conclude that $f = 0$ from Theorem 1.22.

Our next goal is to prove that $\mathcal{G}$ is complete. Let us fix a Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{G}$. This lead us to say that $\left\{ \frac{\widehat{f_n}}{\sqrt{\widehat{\Phi}}} \right\}_{n\in\mathbb{N}}$ is Cauchy in $L^2(\mathbb{R}^d)$. Since $L^2(\mathbb{R}^d)$ is complete, there exists $g \in L^2(\mathbb{R}^d)$ such that

$$\frac{\widehat{f_n}}{\sqrt{\widehat{\Phi}}} \xrightarrow[n\to+\infty]{\|\cdot\|_{L^2(\mathbb{R}^d)}} g.$$

Such g has the following property: $g\sqrt{\widehat{\Phi}} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

$$\int_{\mathbb{R}^d} \left| g(\omega)\sqrt{\widehat{\Phi}(\omega)} \right| \mathrm{d}\,\omega \le \left(\int_{\mathbb{R}^d} |g(\omega)|^2 \mathrm{d}\,\omega \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} \widehat{\Phi}(\omega) \mathrm{d}\,\omega \right)^{\frac{1}{2}} < +\infty.$$

and

$$\int_{\mathbb{R}^d} \left| g(\omega)\sqrt{\widehat{\Phi}(\omega)} \right|^2 \mathrm{d}\,\omega \le \|\widehat{\Phi}\|_{L^\infty(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} |g(\omega)|^2 \mathrm{d}\,\omega \right) < +\infty$$

because $\Phi \in L^1(\mathbb{R}^d)$. With Theorem 1.22 we can define

$$f(x) := T^{-1}\left(g\sqrt{\widehat{\Phi}}\right)(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} g(\omega)\sqrt{\widehat{\Phi}(\omega)} e^{i\langle x,\omega\rangle} \mathrm{d}\,\omega.$$

for each $x \in \mathbb{R}^d$. With this definition $f \in L^2(\mathbb{R}^d)$, it is well-defined, continuous and satisfies

$$\frac{\widehat{f}}{\sqrt{\widehat{\Phi}}} = g \in L^2(\mathbb{R}^d).$$

Since $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{G}$ we can bound

$$\begin{aligned}|f(x) - f_n(x)| &\leq (2\pi)^{-d/2} \int_{\mathbb{R}^d} \left| g(\omega)\sqrt{\widehat{\Phi}(\omega)} - \widehat{f_n}(\omega) \right| \mathrm{d}\,\omega = \\ &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{|g(\omega)\sqrt{\widehat{\Phi}(\omega)} - \widehat{f_n}(\omega)|}{\sqrt{\widehat{\Phi}(\omega)}} \sqrt{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega \leq \\ &\leq (2\pi)^{-d/2} \left\| g - \frac{\widehat{f_n}}{\sqrt{\widehat{\Phi}}} \right\|_{L^2(\mathbb{R}^d)} \left\| \widehat{\Phi} \right\|^{\frac{1}{2}}_{L^1(\mathbb{R}^d)},\end{aligned}$$

which proves that $f \in \mathcal{G}$ because it is real-valued.

$\mathcal{G}$ is complete because

$$\|f - f_n\|_{\mathcal{G}} = (2\pi)^{-d/4} \left\| \frac{\widehat{f}}{\sqrt{\widehat{\Phi}}} - \frac{\widehat{f_n}}{\sqrt{\widehat{\Phi}}} \right\|_{L^2(\mathbb{R}^d)} = (2\pi)^{-d/4} \left\| g - \frac{\widehat{f_n}}{\sqrt{\widehat{\Phi}}} \right\|_{L^2(\mathbb{R}^d)} \xrightarrow{n\to+\infty} 0.$$

To conclude the proof with Theorem 1.21 we need to show that $\Phi(\cdot - \cdot)$ is the reproducing kernel of $\mathcal{G}$. We comment that Φ is bounded by $\Phi(0)$ and in $L^1(\mathbb{R}^d)$ so it is also in $L^2(\mathbb{R}^d)$:

$$\int_{\mathbb{R}^d} |\Phi(x-y)|^2 \mathrm{d}\,x \leq \|\Phi\|_{L^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} |\Phi(x-y)| \mathrm{d}\,x = \|\Phi\|_{L^\infty(\mathbb{R}^d)} \underbrace{\int_{\mathbb{R}^d} |\Phi(x)| \mathrm{d}\,x}_{\substack{\text{change of variable:} \\ x\mapsto x+y}} < +\infty.$$

Now we study

$$\begin{aligned}\frac{\Phi(\cdot - y)^\wedge(x)}{\sqrt{\widehat{\Phi}(x)}} &= \frac{(2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(\omega - y) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega}{\sqrt{(2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(\omega) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega}} = \\ &\underbrace{=}_{\substack{\text{change of variable:} \\ \omega\mapsto\omega+y}} \frac{(2\pi)^{-d/2} (\int_{\mathbb{R}^d} \Phi(\omega) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega) e^{-i\langle x,y\rangle}}{\sqrt{(2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(\omega) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega}}.\end{aligned} \tag{1.12}$$

If we apply $|\cdot|^2$ to the previous expression we obtain

$$\left| \frac{\Phi(\cdot - y)^\wedge(x)}{\sqrt{\widehat{\Phi}(x)}} \right|^2 = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(\omega) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega = \widehat{\Phi}(x),$$

and

$$\left\| \frac{\Phi(\cdot - y)^\wedge}{\sqrt{\widehat{\Phi}}} \right\|^2_{L^2(\mathbb{R}^d)} = \|\widehat{\Phi}\|_{L^1(\mathbb{R}^d)} < +\infty.$$

We proved that $\Phi(\cdot - y) \in \mathcal{G}$ for every $y \in \mathbb{R}^d$.

Following the same steps of equation (1.12) we obtain the reproduction property:

$$\begin{aligned}
\langle f, \Phi(\cdot - y) \rangle_{\mathcal{G}} &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\widehat{f}(\omega)\overline{\Phi(\cdot - y)^\wedge(\omega)}}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega = \\
&= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\widehat{f}(\omega)\overline{\widehat{\Phi}(\omega)e^{-i\langle \omega, y\rangle}}}{\widehat{\Phi}(\omega)} \mathrm{d}\,\omega = \\
&= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{f}(\omega) e^{i\langle \omega, y\rangle} \mathrm{d}\,\omega = f(y),
\end{aligned}$$

because $f \in \mathcal{G}$ and $\widehat{\Phi}$ is real-valued.

□

We can finally state that native spaces are generalization of classical function spaces, e.g. Sobolev spaces. We recall that for $s > \frac{d}{2}$ the Sobolev space of order s is defined as

$$H^s(\mathbb{R}^d) = W^{2,s}(\mathbb{R}^d) = \{f \in L^2(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d) : \widehat{f}(\cdot)(1 + \|\cdot\|_2^2)^{s/2} \in L^2(\mathbb{R}^d)\}.$$

with the following norm

$$\|g\|_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\widehat{g}(\omega)|^2 (1 + \|\omega\|_2^2)^s \mathrm{d}\,\omega. \tag{1.13}$$

If Φ has a Fourier transform that decays only algebraically then its native space is a Sobolev space.

Theorem 1.25 *Suppose that $\Phi \in L^1(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d)$ satisfies*

$$c_1(1 + \|\omega\|_2^2)^{-s} \leq \widehat{\Phi}(\omega) \leq c_2(1 + \|\omega\|_2^2)^{-s} \qquad \omega \in \mathbb{R}^d$$

with $s > \frac{d}{2}$ and two constant $0 < c_1 \leq c_2$. Then the native space $\mathcal{N}_\Phi(\mathbb{R}^d)$ corresponding to Φ coincides with the Sobolev space $H^s(\mathbb{R}^d)$, and the native space norm and the Sobolev norm are equivalent.

Proof

From Theorem 1.23 we recall that

$$\mathcal{N}_\Phi(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d) : \frac{\widehat{f}}{\sqrt{\widehat{\Phi}}} \in L^2(\mathbb{R}^d) \right\}.$$

If we apply $\sqrt{\cdot}$ to the hypothesis on the algebraic decay of the Fourier transform we obtain

$$\frac{c_1^{\frac{s}{2}}}{\sqrt{\widehat{\Phi}(\omega)}} \leq (1+\|\omega\|_2^2)^{\frac{s}{2}} \leq \frac{c_2^{\frac{s}{2}}}{\sqrt{\widehat{\Phi}(\omega)}} \Longrightarrow \frac{c_1^{\frac{s}{2}}\widehat{f}(\omega)}{\sqrt{\widehat{\Phi}(\omega)}} \leq \widehat{f}(\omega)(1+\|\omega\|_2^2)^{\frac{s}{2}} \leq \frac{c_2^{\frac{s}{2}}\widehat{f}(\omega)}{\sqrt{\widehat{\Phi}(\omega)}},$$

which implies $\mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ and

$$c_1^{\frac{1}{2}}\|f\|_{\mathcal{N}_\Phi(\mathbb{R}^d)} \leq (2\pi)^{-\frac{d}{4}}\|f\|_{H^s(\mathbb{R}^d)} \leq c_2^{\frac{1}{2}}\|f\|_{\mathcal{N}_\Phi(\mathbb{R}^d)}.$$

□

For our purposes it is useful to scale the basis function in the following way [16, 17]. Let Φ_δ defined by

$$\Phi_\delta(x) := \Phi\left(\frac{x}{\delta}\right), \qquad \text{with } \delta > 0. \tag{1.14}$$

With this definition the following property on the Fourier transform holds

$$\begin{aligned}
\widehat{\Phi_\delta}(\omega) &= (2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\Phi_\delta(x)e^{-i\langle x,\omega\rangle}\mathrm{d}\,x \\
&= (2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\Phi\left(\frac{x}{\delta}\right)e^{-i\langle x,\omega\rangle}\mathrm{d}\,x = \underbrace{\delta^d(2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\Phi(x)e^{-i\langle \delta x,\omega\rangle}\mathrm{d}\,x}_{\substack{\text{change of variable:}\\ x\mapsto\delta x}} = \\
&= \delta^d\widehat{\Phi}(\delta\omega).
\end{aligned} \tag{1.15}$$

Theorem 1.26 *Suppose that* $\Phi \in L^1(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d)$ *satisfies*

$$c_1(1+\|\omega\|_2^2)^{-s} \leq \widehat{\Phi}(\omega) \leq c_2(1+\|\omega\|_2^2)^{-s} \qquad \textit{for each } \omega \in \mathbb{R},$$

with $s > \frac{d}{2}$ *and two constant* $0 < c_1 \leq c_2$*. For every* $\delta \in (0,1]$ *we have* $\mathcal{N}_{\Phi_\delta}(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ *and the norm equivalence*

$$c_1^{\frac{1}{2}}\delta^{\frac{d}{2}}\|f\|_{\mathcal{N}_{\Phi_\delta}(\mathbb{R}^d)} \leq (2\pi)^{-\frac{d}{4}}\|f\|_{H^s(\mathbb{R}^d)} \leq c_2^{\frac{1}{2}}\delta^{\frac{d}{2}-s}\|f\|_{\mathcal{N}_{\Phi_\delta}(\mathbb{R}^d)}.$$

Proof

From equation (1.15) we get

$$c_1\delta^d(1+\delta^2\|\omega\|_2^2)^{-s} \leq \widehat{\Phi_\delta}(\omega) \leq c_2\delta^d(1+\delta^2\|\omega\|_2^2)^{-s}, \qquad \omega \in \mathbb{R}^d.$$

For $\delta \leq 1$ we can write

$$(1+\|\omega\|_2^2)^s = \delta^{-2s}(\delta^2+\delta^2\|\omega\|_2^2)^s \leq \delta^{-2s}(1+\delta^2\|\omega\|_2^2)^s.$$

If we recall the definition of Sobolev and native space norms (equation 1.13, Theorem 1.23) with the following chains of inequalities we can conclude.

$$\|f\|^2_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2(1+\|\omega\|_2^2)^s \mathrm{d}\,\omega \leq \delta^{-2s}\int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2(1+\delta^2\|\omega\|_2^2)^s \mathrm{d}\,\omega \leq$$
$$\leq c_2\delta^{d-2s}\int_{\mathbb{R}^d} \frac{|\widehat{f}(\omega)|^2}{\widehat{\Phi_\delta}(\omega)}\mathrm{d}\,\omega = (2\pi)^{\frac{d}{2}}c_2\delta^{d-2s}\|f\|^2_{\mathcal{N}_\Phi(\mathbb{R}^d)},$$

$$\|f\|^2_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2(1+\|\omega\|_2^2)^s \mathrm{d}\,\omega \geq \int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2(1+\delta^2\|\omega\|_2^2)^s \mathrm{d}\,\omega \geq$$
$$\geq c_1\delta^{d}\int_{\mathbb{R}^d} \frac{|\widehat{f}(\omega)|^2}{\widehat{\Phi_\delta}(\omega)}\mathrm{d}\,\omega = (2\pi)^{\frac{d}{2}}c_1\delta^{d}\|f\|^2_{\mathcal{N}_\Phi(\mathbb{R}^d)}.$$

□

Chapter 2

Conditionally positive definite kernels

2.1 Native spaces for conditionally positive definite kernels

The first part of this section can be found at [23, 24, 25, 26, 27].

Definition 2.1 *Suppose that $\mathcal{P}$ is a finite-dimensional subspace of $\mathcal{C}(\Omega)$, $\Omega \subseteq \mathbb{R}^d$. A continuous symmetric kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ is said to be conditionally positive definite on Ω with respect to $\mathcal{P}$ if, for any N pairwise distinct centers $x_1, \ldots, x_N \in \Omega$ and all $\alpha \in \mathbb{R}^N \setminus \{0\}$ with*

$$\sum_{j=1}^{N} \alpha_j p(x_j) = 0$$

for all $p \in \mathcal{P}$, then the quadratic form

$$\sum_{j=1}^{N} \sum_{k=1}^{N} \alpha_j \alpha_k \Phi(x_j, x_k) > 0.$$

The domain $\Omega \subseteq \mathbb{R}^d$ can still be quite arbitrary but it should contain at least one $\mathcal{P}-$unisolvent set (Definition 3.4). We will say that a conditionally positive definite kernel has order m if it is conditionally positive definite with respect to $\pi_{m-1}(\mathbb{R}^d)$. We note that a function which is conditionally positive definite of order m then it is also conditionally positive definite of order $\ell \geq m$. This means that every positive definite function is also conditionally positive definite of any order.

If we fix $\langle p_1, \ldots, p_Q \rangle$ a basis of $\mathcal{P}$ the conditional positive definiteness of order Q of a kernel Φ can also be interpreted as the positive definiteness of the matrix $A_{\Phi,X} = (\Phi(x_j - x_k))_{j,k=1,\ldots,N}$ on the vector subspace of $\mathbb{R}^d$ that satisfies the linear constraints

$$\sum_{j=1}^{N} \alpha_j p_\ell(x_j) = 0 \qquad 1 \leq \ell \leq Q = \dim(\mathcal{P}).$$

Now we will explain why it is important that Ω contains a $\mathcal{P}-$unisolvent set. We want to interpolate the function $f : \Omega \to \mathbb{R}$ on the data set $X = \{x_1, \ldots, x_N\}$ with an interpolant of the form

$$s_{f,X}(x) = \sum_{j=1}^{N} \alpha_j \Phi(x - x_j) + \sum_{k=1}^{Q} \beta_k p_k(x). \tag{2.1}$$

Since we have $N + Q$ parameters we need $N + Q$ linear equations:

$$\begin{aligned} s_{f,X}(x_j) &= f(x_j) \qquad 1 \le j \le N \\ \sum_{j=1}^{N} \alpha_j p_k(x_j) &= 0 \qquad 1 \le k \le Q, \end{aligned}$$

in other words we need to solve

$$\widetilde{A}_{\Phi,X} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} A_{\Phi,X} & P \\ P^\top & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} f_{|X} \\ 0 \end{pmatrix} \tag{2.2}$$

where $A_{\Phi,X} = (\Phi(x_j - x_k))_{j,k=1,\ldots,N} \in M_N(\mathbb{R})$ and $P = (p_k(x_j))_{j=1,\ldots,N,k=1,\ldots,Q} \in M_{N,Q}(\mathbb{R})$. We will start to study the uniqueness of the solution.

Theorem 2.1 *Suppose that Φ is conditionally positive definite with respect to $\mathcal{P}$ and $X = \{x_1, \ldots, x_N\}$ is a $\mathcal{P}-$unisolvent set. Then the linear system described in equation* (2.2) *is uniquely solvable.*

Proof
Suppose that $(\alpha, \beta)^\top \in \ker(\widetilde{A}_{\Phi,X})$, so

$$\begin{aligned} A_{\Phi,X}\alpha + P\beta &= 0 \\ P^\top \alpha &= 0. \end{aligned}$$

If we multiply the first equation by $\alpha^\top$ we obtain

$$0 = \alpha^\top A_{\Phi,X}\alpha + \alpha^\top P\beta = \alpha^\top A_{\Phi,X}\alpha + \beta^\top \underbrace{P^\top \alpha}_{=0} = \alpha^\top A_{\Phi,X}\alpha.$$

The second equation and Definition 2.1 imply $\alpha = 0$ and consequently $P\beta = 0$. Since X is $\mathcal{P}-$unisolvent the matrix P is injective and this guarantees $\beta = 0$.

□

To prove uniqueness we used the fact that X is $\mathcal{P}-$unisolvent. This is not necessary if we need only existence. Let us fix some notation: we will denote with $V = \text{range}(P) \subseteq \mathbb{R}^N$ and for brevity $A = A_{\Phi,X}$. The orthogonal complements of V, $V^\top$, is the kernel of $P^\top$. The system (2.2) admits a solution for every $f_{|X}$ if $\mathbb{R}^N = AV^\top + V$. We claim that $\mathbb{R}^N = AV^\top \oplus V$. If $x \in AV^\top \cap V$ we have that $x = A\alpha = P\beta$ with $\alpha \in V^\top$ and $\beta \in \mathbb{R}^Q$. A multiplication by $\alpha^\top$ of the chain of equalities gives us $\alpha^\top A\alpha = \alpha^\top P\beta = \beta^\top(P^\top\alpha) = 0$, which implies $\alpha = 0$ and $x = A\alpha = 0$. We note that $A : V^\top \to V^\top$ is injective: if $Av_1 = Av_2$

with $v_1, v_2 \in V^\top$ then $A(v_1 - v_2) = 0$ and $(v_1 - v_2)^\top A(v_1 - v_2) = 0$. Since $v_1 - v_2 \in V^\top$ we can write $v_1 - v_2 = 0$. We can conclude the claim with a dimension comparison.

$$N \geq \dim(AV^\top + V) = \dim(AV^\top) + \dim(V) = \dim(V^\top) + \dim(V) = N.$$

We note that if $\mathcal{P} = \pi_m(\mathbb{R}^d)$ then the proposed interpolation method reproduces $\pi_m(\mathbb{R}^d)$ exactly.

After justifying our hypotheses we can build the native space. Similarly as in the case of positive definite kernel we start with a linear subspace of $\mathcal{C}(\Omega)$:

$$F_\Phi(\Omega) = \left\{ \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) : N \in \mathbb{N}, \alpha \in \mathbb{R}^N, x_1, \dots, x_N \in \Omega \text{ and } \sum_{j=1}^N \alpha_j p(x_j) = 0 \text{ for all } p \in \mathcal{P} \right\}.$$

We note that every $\alpha \in \mathbb{R}^N$ such that $\sum_{j=1}^N \alpha_j p(x_j) = 0$ for all $p \in \mathcal{P}$ uniquely determines an element of $F_\Phi(\Omega)$:

$$\sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) = 0 \Rightarrow \sum_{k=1}^N \alpha_k \underbrace{\sum_{j=1}^N \alpha_j \Phi(x_k, x_j)}_{=0} = \sum_{k=1}^N \sum_{j=1}^N \alpha_k \alpha_j \Phi(x_k, x_j) = 0, \tag{2.3}$$

which implies $\alpha = 0$. If we define a bilinear form on $F_\Phi(\Omega)$ as

$$\left\langle \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^M \beta_k \Phi(\cdot, y_k) \right\rangle_\Phi = \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \Phi(x_j, y_k),$$

then $F_\Phi(\Omega)$ becomes an inner product space (the constraints on α and β make the bilinear form positive definite).

As in the case of positive definite kernels it is useful for this construction to consider the completion $\mathcal{F}_\Phi(\Omega)$ of $F_\Phi(\Omega)$ (Theorem 1.13), which in general is not a space of functions. This construction is more tricky because it is not true that $\Phi(\cdot, x) \in F_\Phi(\Omega)$ so it is not clear how to extend the point-evaluation functional δ_x in a continuous way.

From now on will be essential that there exists a $\mathcal{P}-$unisolvent set $\Xi = \{\xi_1, \dots, \xi_Q\} \subseteq \Omega$ with $\#\Xi = \dim(\mathcal{P}) = Q$. With respect to Ξ we can also obtain a Lagrange basis $\{p_1, \dots, p_Q\} \subseteq \mathcal{P}$ such that $p_i(\xi_j) = \delta_{i,j}$ for each $i, j = 1, \dots, Q$. We define a modified point-evaluation functional as

$$\delta_{(x)} = \delta_x - \sum_{k=1}^Q p_k(x) \delta_{\xi_k} \quad \text{for } x \in \Omega. \tag{2.4}$$

We apply it to $\Phi(\cdot, y)$ to obtain

$$G(\cdot, x) = \delta_{(x)}^y \Phi(\cdot, y) = \Phi(\cdot, x) - \sum_{k=1}^Q p_k(x) \Phi(\cdot, \xi_k) \quad \text{for } x \in \Omega. \tag{2.5}$$

We note that $\sum_{k=1}^{Q} p_k(x)\Phi(\cdot,\xi_k)$ as a function of $x \in \Omega$ is the unique function in $\mathcal{P}$ that interpolates on Ξ the map $x \mapsto \Phi(\cdot, x)$.

We claim that $G(\cdot, x) \in F_\Phi(\Omega)$ because the vector in $\mathbb{R}^{Q+1}$ that defines the function is $(1, -p_1(x), \ldots, -p_Q(x))^\top$ and it satisfies

$$1 \cdot p(x) - \sum_{k=1}^{Q} p_k(x) p(\xi_k) = 0 \qquad \text{for } p \in \mathcal{P},$$

because $\{p_1, \ldots, p_Q\} \subseteq \mathcal{P}$ is a Lagrange basis for Ξ.

If $f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega)$ then

$$\begin{aligned} \delta_{(x)}(f) = f(x) - \sum_{k=1}^{Q} p_k(x) f(\xi_k) &= \sum_{j=1}^{N} \alpha_j \Phi(x, x_j) - \sum_{k=1}^{Q} p_k(x) \sum_{j=1}^{N} \alpha_j \Phi(\xi_k, x_j) = \\ &= \sum_{j=1}^{N} \alpha_j \Phi(x, x_j) - \sum_{k=1}^{Q}\sum_{j=1}^{N} p_k(x) \alpha_j \Phi(\xi_k, x_j) = \\ &= \left\langle \Phi(\cdot, x) - \sum_{k=1}^{Q} p_k(x)\Phi(\cdot, \xi_k), \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) \right\rangle_\Phi = \\ &= \langle G(\cdot, x), f\rangle_\Phi = \langle f, G(\cdot, x)\rangle_\Phi, \end{aligned} \tag{2.6}$$

so $|\delta_{(x)}(f)| \leq \|f\|_\Phi \|G(\cdot, x)\|_\Phi$, that proves the boundness of $\delta_{(x)}$ in $F_\Phi(\Omega)$.

We can uniquely extends $\delta_{(x)}$ to $\mathcal{F}_\Phi(\Omega)$ and define the linear map $R : \mathcal{F}_\Phi(\Omega) \to \mathcal{C}(\Omega)$ as

$$\begin{aligned} R(f) : \Omega &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle f, G(\cdot, x)\rangle_\Phi \end{aligned} \tag{2.7}$$

with $f \in \mathcal{F}_\Phi(\Omega)$. We claim that $R(f)$ is a continuous function because

$$|R(f)(x) - R(f)(y)| = |\langle f, G(\cdot, x) - G(\cdot, y)\rangle_\Phi| \leq \|f\|_\Phi \|G(\cdot, x) - G(\cdot, y)\|_\Phi$$

and

$$\begin{aligned} \|G(\cdot, x) - G(\cdot, y)\|_\Phi^2 &= \|\Phi(\cdot, x) - \Phi(\cdot, y) - \sum_{k=1}^{Q} (p_k(x) - p_k(y))\Phi(\cdot, \xi_k)\|_\Phi^2 = \\ &= \Phi(x, x) + \Phi(y, y) - 2\Phi(x, y) - \sum_{k=1}^{Q} (p_k(x) - p_k(y))(\Phi(x, \xi_k) - \Phi(y, \xi_k)) + \\ &+ \sum_{j,k=1}^{Q} (p_k(x) - p_k(y))(p_j(x) - p_j(y))\Phi(\xi_j, \xi_k). \end{aligned}$$

Now we claim that $R : \mathcal{F}_\Phi(\Omega) \to \mathcal{C}(\Omega)$ injects $\mathcal{F}_\Phi(\Omega)$ in $\mathcal{C}(\Omega)$.

Theorem 2.2 *The linear map $R : \mathcal{F}_\Phi(\Omega) \to \mathcal{C}(\Omega)$ defined in equation 2.7 is injective.*

Proof
Since R is linear we verify that $\ker(R) = 0$. Suppose that $R(f) = 0$ with $f \in \mathcal{F}_\Phi(\Omega)$. We have that $0 = R(f)(x) = \langle f, G(\cdot, x)\rangle_\Phi$. Fix $h = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega)$ and compute

$$\sum_{j=1}^N \alpha_j G(\cdot, x_j) = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) - \sum_{j=1}^N \alpha_j \sum_{k=1}^Q p_k(x_j)\Phi(\cdot, \xi_k) =$$
$$= h - \sum_{k=1}^Q \Phi(\cdot, \xi_k) \underbrace{\sum_{j=1}^N \alpha_j p_k(x_j)}_{=0} = h.$$

Thus we can say that $\langle f, h\rangle_\Phi = 0$ for all $h \in F_\Phi(\Omega)$, which proves that $f \in F_\Phi(\Omega)^\perp = \overline{F_\Phi(\Omega)}^\perp = \mathcal{F}_\Phi(\Omega)^\perp = \langle 0 \rangle$ (Theorem 1.13).

□

If we follow the same step we have done for positive definite kernels then $R(\mathcal{F}_\Phi(\Omega))$ will be the candidate native space, but in this situation this does not make sense.

Indeed, if $f \in \mathcal{F}_\Phi(\Omega)$ then $R(f)(\xi_k) = \langle f, G(\cdot, \xi_k)\rangle_\Phi = 0$ for $k = 1, \ldots, Q$, because

$$G(\cdot, \xi_k) = \Phi(\cdot, \xi_k) - \sum_{i=1}^Q \underbrace{p_i(\xi_k)}_{\delta_{i,k}} \Phi(\cdot, \xi_i) = \Phi(\cdot, \xi_k) - \Phi(\cdot, \xi_k) = 0. \tag{2.8}$$

In general $f(\xi_k)$ is not 0 for $f \in F_\Phi(\Omega)$ so $R(\mathcal{F}_\Phi(\Omega))$ can not be the candidate native space.

Since we want that the candidate native space contains $F_\Phi(\Omega)$ we study the range of R when it is applied to this vector subspace. If $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega)$ then in equation (2.6) we showed

$$R(f)(x) = \langle f, G(\cdot, x)\rangle_\Phi = f(x) - \sum_{k=1}^Q f(\xi_k)p_k(x) = f(x) - \Pi_\mathcal{P}(f)(x),$$

where $\Pi_\mathcal{P}(f) \in \mathcal{P}$ is the only function in $\mathcal{P}$ that interpolates f on Ξ, because $\{p_1, \ldots, p_Q\}$ is a Lagrange basis for $\mathcal{P}$ with respect to Ξ. For $f \in F_\Phi(\Omega)$ it is important to understand that

$$f = R^{-1}(f - \Pi_\mathcal{P}(f)). \tag{2.9}$$

Finally we can define the candidate native space for conditionally positive definite kernels.

Definition 2.2 *The native space corresponding to a symmetric kernel Φ that is conditionally positive definite on Ω with respect to $\mathcal{P}$ is*

$$\mathcal{N}_\Phi(\Omega) = R(\mathcal{F}_\Phi(\Omega)) \oplus \mathcal{P}.$$

We define on the space $\mathcal{N}_\Phi(\Omega)$ *the degenerate bilinear form*

$$\langle f, g\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), R^{-1}(g - \Pi_{\mathcal{P}}(g))\rangle_\Phi. \tag{2.10}$$

First of all, $R(\mathcal{F}_\Phi(\Omega)) \oplus \mathcal{P}$ is a direct sum because every element of $R(\mathcal{F}_\Phi(\Omega))$ vanishes on Ξ (equation (2.8)), but the only element of $\mathcal{P}$ that is 0 on Ξ is the zero function.

The degenerate bilinear form in equation (2.10) is well-defined because $f - \Pi_{\mathcal{P}} \in R(\mathcal{F}_\Phi(\Omega))$ for $f \in \mathcal{N}_\Phi(\Omega)$. Every $f \in \mathcal{N}_\Phi(\Omega)$ can be written as $f = R(g) + p$ with $g \in \mathcal{F}_\Phi(\Omega), p \in \mathcal{P}$ and $f(\xi_k) = p(\xi_k)$ for $k = 1, \ldots, Q$. This description of f permit us to state that $p = \Pi_{\mathcal{P}}(f)$, thus $f - \Pi_{\mathcal{P}}(f) = R(g) \in R(\mathcal{F}_\Phi(\Omega))$. The last considerations imply that

$$\{f - \Pi_{\mathcal{P}}(f) : f \in \mathcal{N}_\Phi(\Omega)\} = R(\mathcal{F}_\Phi(\Omega)).$$

From this we can prove that the null space of $\langle \cdot, \cdot\rangle_{\mathcal{N}_\Phi(\Omega)}$ is $\mathcal{P}$.

If $\langle f, g\rangle_{\mathcal{N}_\Phi(\Omega)} = 0$ for all $g \in \mathcal{N}_\Phi(\Omega)$ then $\langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), R^{-1}(g - \Pi_{\mathcal{P}}(g))\rangle_\Phi = 0$ for all $g \in \mathcal{N}_\Phi(\Omega)$. Since $\langle \cdot, \cdot\rangle_\Phi$ is non-degenerate and R is linear we obtain than $R^{-1}(f - \Pi_{\mathcal{P}}(f)) = 0$ and consequently $f - \Pi_{\mathcal{P}}(f) = 0$, so $f \in \mathcal{P}$. The inverse inclusion holds because if $p \in \mathcal{P}$ then $p - \Pi_{\mathcal{P}}(p) = 0$.

From equation (2.9) if $f \in F_\Phi(\Omega)$ we can achieve

$$|f|^2_{\mathcal{N}_\Phi(\Omega)} = \langle f, f\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), R^{-1}(f - \Pi_{\mathcal{P}}(f))\rangle_\Phi = \langle f, f\rangle_\Phi = \|f\|^2_\Phi. \tag{2.11}$$

We prove that $\langle \cdot, \cdot\rangle_{\mathcal{N}_\Phi(\Omega)}$ is not an inner product. For this reason we will state a generalization of the reproducing kernel property for positive definite kernels.

Theorem 2.3 *Suppose that* $\Phi : \Omega \times \Omega \to \mathbb{R}$ *is a symmetric kernel that is conditionally positive definite on* Ω *with respect to* $\mathcal{P} \subseteq \mathcal{C}(\Omega)$*. Every* $f \in \mathcal{N}_\Phi(\Omega)$ *is*

$$f(x) = \Pi_{\mathcal{P}}(f)(x) + \langle f, G(\cdot, x)\rangle_{\mathcal{N}_\Phi(\Omega)},$$

with G *defined in equation* (2.5)*.*

Proof
From equation (2.9) we know that $G(\cdot, x) = R^{-1}(G(\cdot, x) - \Pi_{\mathcal{P}}(G(\cdot, x)))$ because $G(\cdot, x) \in F_\Phi(\Omega)$. We proved that every element $f \in \mathcal{N}_\Phi(\Omega)$ can be written as $f = R(g) + \Pi_{\mathcal{P}}(f)$ with $g \in \mathcal{F}_\Phi(\Omega)$ so $R(g)(x) = \langle g, G(\cdot, x)\rangle_\Phi = \langle R^{-1}(R(g)), G(\cdot, x)\rangle_\Phi = \langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), G(\cdot, x)\rangle_\Phi$. We can compute

$$\begin{aligned} f(x) = \Pi_{\mathcal{P}}(f)(x) + R(g)(x) = \Pi_{\mathcal{P}}(f)(x) + \langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), G(\cdot, x)\rangle_\Phi = \\ = \Pi_{\mathcal{P}}(f)(x) + \langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), R^{-1}(G(\cdot, x) - \Pi_{\mathcal{P}}(G(\cdot, x)))\rangle_\Phi = \\ = \Pi_{\mathcal{P}}(f)(x) + \langle f, G(\cdot, x)\rangle_{\mathcal{N}_\Phi(\Omega)}. \end{aligned}$$

□

We can think this representation of $f \in \mathcal{N}_\Phi(\Omega)$ as a Taylor expansion. Theorem 2.3 is a generalization of Definition 1.7 because if we consider a positive definite kernel with $\mathcal{P} = \langle 0 \rangle$ then $G(\cdot, x) = \Phi(\cdot, x)$ and $\mathcal{N}_\Phi(\Omega) = R(\mathcal{F}_\Phi(\Omega))$.

Before continuing we note that if $f \in \mathcal{N}_\Phi(\Omega)$ and $f(\xi_k) = 0$ for $k = 1, \dots, Q$ then $f \in R(\mathcal{F}_\Phi(\Omega))$:

$$f = R(g) + \Pi_\mathcal{P}(f) \Rightarrow f - R(g) = \Pi_\mathcal{P}(f),$$

which implies $\Pi_\mathcal{P}(f) = 0$ because $R(g)_{|\Xi} = 0$ and Ξ is $\mathcal{P}-$unisolvent.

Theorem 2.4 *The bilinear form $\langle \cdot, \cdot \rangle_{\mathcal{N}_\Phi(\Omega)}$ is an inner product on the vector space*

$$X_\Phi(\Omega) = \{f \in \mathcal{N}_\Phi(\Omega) : f(\xi_k) = 0 \text{ for } k = 1, \dots, Q\} = R(\mathcal{F}_\Phi(\Omega)).$$

Moreover $(X_\Phi(\Omega), \langle \cdot, \cdot \rangle_{\mathcal{N}_\Phi(\Omega)})$ is an Hilbert space with reproducing kernel

$$\rho(x, y) = \Phi(x, y) - \sum_{k=1}^{Q} p_k(x)\Phi(\xi_k, y) - \sum_{\ell=1}^{Q} p_\ell(y)\Phi(x, \xi_\ell) + \sum_{k=1}^{Q}\sum_{\ell=1}^{Q} p_k(x)p_\ell(y)\Phi(\xi_k, \xi_\ell).$$

Proof
If $f \in X_\Phi(\Omega)$ such that $\langle f, g \rangle_{\mathcal{N}_\Phi(\Omega)} = 0$ for all $g \in X_\Phi(\Omega) = R(\mathcal{F}_\Phi(\Omega))$ then $\langle f, g \rangle_{\mathcal{N}_\Phi(\Omega)} = 0$ for all $g \in \mathcal{N}_\Phi(\Omega) = R(\mathcal{F}_\Phi(\Omega)) \oplus \mathcal{P}$ because $\mathcal{P}$ is contained in the null space of the bilinear form. This proves that $f \in R(\mathcal{F}_\Phi(\Omega)) \cap \mathcal{P} = \langle 0 \rangle$.

We study

$$\begin{aligned}\langle R(f), R(g) \rangle_{\mathcal{N}_\Phi(\Omega)} &= \langle R^{-1}(R(f) - \Pi_\mathcal{P}(R(f))), R^{-1}(R(g) - \Pi_\mathcal{P}(R(g))) \rangle_\Phi = \\ &= \langle R^{-1}(R(f)), R^{-1}(R(g)) \rangle_\Phi = \langle f, g \rangle_\Phi,\end{aligned}$$

that shows that $R : \mathcal{F}_\Phi(\Omega) \to R(\mathcal{F}_\Phi(\Omega))$ is an isometry (Theorem 2.2). From the fact that $\mathcal{F}_\Phi(\Omega)$ is the Hilbert completion of $F_\Phi(\Omega)$ follows that $R(\mathcal{F}_\Phi(\Omega))$ is a Hilbert space.

We claim that $\rho(x, y) = R(G(\cdot, y))(x)$. Indeed,

$$\begin{aligned}R(G(\cdot, y))(x) &= \langle G(\cdot, y), G(\cdot, x) \rangle_\Phi = \\ &= \left\langle \Phi(\cdot, y) - \sum_{k=1}^{Q} p_k(y)\Phi(\cdot, \xi_k), \Phi(\cdot, x) - \sum_{k=1}^{Q} p_k(x)\Phi(\cdot, \xi_k) \right\rangle_\Phi = \rho(x, y)\end{aligned}$$

At this point we know that $\rho(\cdot, y) \in R(\mathcal{F}_\Phi(\Omega)) = X_\Phi(\Omega)$.

Before proving the reproduction property with Theorem 2.3, we recall from equation (2.9) that

$$G(\cdot, x) = R^{-1}(G(\cdot, x) - \Pi_\mathcal{P}(G(\cdot, x))),$$

because $G(\cdot,x) \in F_\Phi(\Omega)$.

$$\begin{aligned}
\langle f, \rho(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} &= \langle f, R(G(\cdot,x))\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle R(R^{-1}(f)), R(G(\cdot,x))\rangle_{\mathcal{N}_\Phi(\Omega)} \stackrel{R \text{ isometry}}{=} \\
&= \langle R^{-1}(f), G(\cdot,x)\rangle_\Phi = \langle R^{-1}(f), R^{-1}(G(\cdot,x) - \Pi_{\mathcal{P}}(G(\cdot,x)))\rangle_\Phi = \\
&= \underbrace{\Pi_{\mathcal{P}}(f)(x)}_{=0} + \langle R^{-1}(f - \underbrace{\Pi_{\mathcal{P}}(f)}_{=0}), R^{-1}(G(\cdot,x) - \Pi_{\mathcal{P}}(G(\cdot,x)))\rangle_\Phi = \\
&= \Pi_{\mathcal{P}}(f)(x) + \langle f, G(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} = f(x).
\end{aligned}$$

□

The next step is to prove the existence of a reproducing kernel for $\mathcal{N}_\Phi(\Omega)$.

Theorem 2.5 *The native space $\mathcal{N}_\Phi(\Omega)$ of a conditionally positive definite kernel Φ equipped with the inner product*

$$\langle f, g\rangle = \langle f, g\rangle_{\mathcal{N}_\Phi(\Omega)} + \sum_{k=1}^{Q} f(\xi_k) g(\xi_k)$$

is a Hilbert space and it admits a reproducing kernel

$$K(x,y) = \rho(x,y) + \sum_{k=1}^{Q} p_k(x) p_k(y),$$

where ρ is the kernel of Theorem 2.4.

Proof
The bilinear form $\langle\cdot,\cdot\rangle$ is symmetric and non-negative. Let us prove that it is a inner product: if $f \in \mathcal{N}_\Phi(\Omega)$ is such that

$$0 = \langle f, f\rangle = \langle f, f\rangle_{\mathcal{N}_\Phi(\Omega)} + \sum_{k=1}^{Q} |f(\xi_k)|^2,$$

that implies $\langle f, f\rangle_{\mathcal{N}_\Phi(\Omega)} = 0$ and $f(\xi_k) = 0$ for $k = 1, \ldots, Q$.

$$\begin{aligned}
&0 = \langle f, f\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle R^{-1}(f - \Pi_{\mathcal{P}}(f)), R^{-1}(f - \Pi_{\mathcal{P}}(f))\rangle_\Phi \Rightarrow \\
&\Rightarrow R^{-1}(f - \Pi_{\mathcal{P}}(f)) = 0 \Rightarrow f - \Pi_{\mathcal{P}}(f) = 0 \Rightarrow f = \Pi_{\mathcal{P}}(f) \in \mathcal{P}
\end{aligned}$$

This condition with $f(\xi_k) = 0$ for $k = 1, \ldots, Q$ gets $f = 0$.

Recalling that $G(\cdot, \xi_k) = 0$ for $k = 1, \ldots, N$ and the characterization of ρ in the proof of Theorem 2.4 we obtain

$$\rho(\xi_k, \cdot) = \rho(\cdot, \xi_k) = R(G(\cdot, \xi_k)) = 0,$$

thus

$$\sum_{k=1}^{Q} f(\xi_k)K(\xi_k,x) = \sum_{k=1}^{Q} f(\xi_k)\underbrace{\rho(\xi_k,x)}_{=0} + \sum_{k=1}^{Q} f(\xi_k)\sum_{\ell=1}^{Q}\underbrace{p_\ell(\xi_k)}_{=\delta_{\ell k}}p_\ell(x) =$$

$$= \sum_{k=1}^{Q} f(\xi_k)p_k(x) = \Pi_{\mathcal{P}}(f)(x) \quad \text{for } f \in \mathcal{N}_\Phi(\Omega).$$

From equation (2.9) we obtain that

$$\rho(\cdot,x) = R(G(\cdot,x)) = G(\cdot,x) - \Pi_{\mathcal{P}}(G(\cdot,x)),$$

so

$$K(\cdot,x) - G(\cdot,x) = \underbrace{K(\cdot,x) - \rho(\cdot,x)}_{\in\mathcal{P}} + \underbrace{\rho(\cdot,x) - G(\cdot,x)}_{\in\mathcal{P}} \in \mathcal{P}.$$

The representation for $f \in \mathcal{N}_\Phi(\Omega)$ given in Theorem 2.3 gives us

$$\begin{aligned} f(x) &= \Pi_{\mathcal{P}}(f)(x) + \langle f, G(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \Pi_{\mathcal{P}}(f)(x) + \langle f, \underbrace{G(\cdot,x) - K(\cdot,x)}_{\in\mathcal{P}}\rangle_{\mathcal{N}_\Phi(\Omega)} + \langle f, K(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \Pi_{\mathcal{P}}(f)(x) + \langle f, K(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \sum_{k=1}^{Q} f(\xi_k)K(\xi_k,x) + \langle f, K(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \langle f, K(\cdot,x)\rangle. \end{aligned}$$

The chain of equality holds because the null space of $\langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega)}$ is $\mathcal{P}$.

□

As in Theorem 1.21 we study a uniqueness property of the native space for conditionally positive definite kernel.

Theorem 2.6 *Suppose that $\mathcal{G} \subseteq \mathcal{C}(\Omega)$ carries a semi-inner product $\langle\cdot,\cdot\rangle_{\mathcal{G}}$ with null space $\mathcal{P} \subseteq \mathcal{G}$ such that $\mathcal{G}_0 = \{g \in \mathcal{G} : g(\xi_k) = 0, 1 \le k \le Q\}$ is a Hilbert space with reproducing kernel ρ (Theorem 2.4). Then $\mathcal{G}$ is the native space corresponding to Φ on Ω.*

Proof
Since $\Xi = \{\xi_1,\dots,\xi_Q\}$ is $\mathcal{P}$-unisolvent then $\mathcal{G}_0 \cap \mathcal{P} = \langle 0\rangle$: if an element of $\mathcal{P}$ vanishes on Ξ then it is 0. Moreover, for $g \in \mathcal{G}$ we have

$$g = \underbrace{g - \Pi_{\mathcal{P}}(g)}_{\in\mathcal{G}_0} + \underbrace{\Pi_{\mathcal{P}}(g)}_{\in\mathcal{P}},$$

because $(g - \Pi_{\mathcal{P}}(g))(\xi_k) = g(\xi_k) - g(\xi_k) = 0$ for $k = 1,\dots,Q$.

We obtain that $\mathcal{G} = \mathcal{G}_0 \oplus \mathcal{P}$, so we have to prove that $\mathcal{G}_0 = R(\mathcal{F}_\Phi(\Omega))$, since $\mathcal{N}_\Phi(\Omega) = R(\mathcal{F}_\Phi(\Omega)) \oplus \mathcal{P}$. First of all, let us fix $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega)$. We study

$$\sum_{j=1}^N \alpha_j G(x, x_j) = \sum_{j=1}^N \alpha_j \left(\Phi(\cdot, x_j) - \sum_{k=1}^Q p_k(x_j)\Phi(\cdot, \xi_k) \right)(x) =$$
$$= \sum_{j=1}^N \alpha_j \Phi(x, x_j) - \sum_{k=1}^Q \Phi(x, \xi_k) \underbrace{\sum_{j=1}^N \alpha_j p_k(x_j)}_{=0} = f(x),$$

from which we obtain, with equation (2.9) and the representation of ρ in the proof of Theorem 2.4,

$$f(x) - \Pi_{\mathcal{P}}(f)(x) = R(f)(x) = R\left(\sum_{j=1}^N \alpha_j G(\cdot, x_j) \right)(x) =$$
$$= \sum_{j=1}^N \alpha_j R(G(\cdot, x_j))(x) = \sum_{j=1}^N \alpha_j \rho(x, x_j) \in \mathcal{G}_0,$$

because ρ is the reproducing kernel of $\mathcal{G}_0$. We have the inclusion $R(F_\Phi(\Omega)) \subseteq \mathcal{G}_0$.

We compute the bilinear form of the two spaces on the subspace $R(F_\Phi(\Omega))$.

Fix $f_1 = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j)$, $f_2 = \sum_{k=1}^M \beta_k \Phi(\cdot, y_k) \in F_\Phi(\Omega)$ and compute

$$\langle R(f_1), R(f_2) \rangle_{\mathcal{G}} = \left\langle \sum_{j=1}^N \alpha_j \rho(\cdot, x_j), \sum_{k=1}^M \beta_k \rho(\cdot, y_k) \right\rangle_{\mathcal{G}} =$$
$$= \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \langle \rho(\cdot, x_j), \rho(\cdot, y_k) \rangle_{\mathcal{G}} = \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \rho(y_k, x_j),$$

because of the reproducing property of ρ.

Recalling from theorem 2.4 that $R : \mathcal{F}_\Phi(\Omega) \to R(\mathcal{F}_\Phi(\Omega))$ is an isometry we have

$$\langle R(f_1), R(f_2) \rangle_{\mathcal{N}_\Phi(\Omega)} = \left\langle R\left(\sum_{j=1}^N \alpha_j G(\cdot, x_j) \right), R\left(\sum_{k=1}^M \beta_k G(\cdot, y_k) \right) \right\rangle_{\mathcal{N}_\Phi(\Omega)} =$$
$$= \left\langle \sum_{j=1}^N \alpha_j G(\cdot, x_j), \sum_{k=1}^M \beta_k G(\cdot, y_k) \right\rangle_\Phi = \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \langle G(\cdot, x_j), G(\cdot, y_k) \rangle_\Phi =$$
$$= \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k R(G(\cdot, x_j))(y_k) = \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \rho(y_k, x_j),$$

which proves that $\langle R(f_1), R(f_2)\rangle_{\mathcal{G}} = \langle R(f_1), R(f_2)\rangle_{\mathcal{N}_\Phi(\Omega)}$ for $f_1, f_2 \in F_\Phi(\Omega)$.

For every $f \in R(\mathcal{F}_\Phi(\Omega))$ there exists a Cauchy sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq R(F_\Phi(\Omega))$ that satisfy

$$f(x) = \lim_{n\to\infty} f_n(x),$$

because $\overline{F_\Phi(\Omega)} = \mathcal{F}_\Phi(\Omega), R$ is an isometry and ρ is the reproducing kernel of $R(\mathcal{F}_\Phi(\Omega))$ (Theorem 2.4). But $\{f_n\}_{n\in\mathbb{N}}$ is also a Cauchy sequence in $\mathcal{G}_0$, so there exists $g \in \mathcal{G}_0$ such that

$$g(x) = \lim_{n\to\infty} f_n(x) = f(x),$$

because of Theorem 1.9 and the reproduction property of ρ in $\mathcal{G}$. The inclusion $R(\mathcal{F}_\Phi(\Omega)) \subseteq \mathcal{G}_0$ holds. Moreover,

$$\|f\|^2_{\mathcal{N}_\Phi(\Omega)} = \lim_{n\to\infty}\langle f_n, f_n\rangle_{\mathcal{N}_\Phi(\Omega)} = \lim_{n\to\infty}\langle f_n, f_n\rangle_{\mathcal{G}} = \|g\|^2_{\mathcal{G}} = \|f\|^2_{\mathcal{G}}$$

for every $f \in R(\mathcal{F}_\Phi(\Omega))$, which implies the equivalence of the inner product on $R(\mathcal{F}_\Phi(\Omega))$ (Theorem 1.21). To conlude the proof we suppose by contradiction that $R(\mathcal{F}_\Phi(\Omega)) \subsetneq \mathcal{G}_0$. Since $R(\mathcal{F}_\Phi(\Omega))$ is a Hilbert space with respcet to $\langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega)}$ (Theorem 2.4) then $R(\mathcal{F}_\Phi(\Omega))$ is closed in $\mathcal{N}_\Phi(\Omega)$ and in $\mathcal{G}$. We can find $g \in \mathcal{G}\setminus\{0\}$ that is orthogonal to $R(\mathcal{F}_\Phi(\Omega))$: if $R(\mathcal{F}_\Phi(\Omega))^{\perp_{\mathcal{G}}} = \langle 0\rangle$ then (Proposition 1.4) $R(\mathcal{F}_\Phi(\Omega)) = \overline{R(\mathcal{F}_\Phi(\Omega))} = \mathcal{G}$, which is a contradiction. Finally we obtain for $x \in \Omega$: $g(x) = \langle g, \rho(\cdot, x)\rangle_{\mathcal{G}} = \langle g, R(G(\cdot, x))\rangle_{\mathcal{G}} = 0$, that implies the contradiction $g = 0$.

□

2.2 Abstract characterization of native spaces

At the beginning of this section we fix a $\mathcal{P}$-unisolvent set $\Xi = \{\xi_1, \ldots, \xi_Q\} \subseteq \Omega$ and we used it to build the native space $\mathcal{N}_\Phi(\Omega)$ (Definition 2.2) and the reproducing kernel K (Theorem 2.5) for a conditionally positive definite kernel. A natural question should arise: is the space $\mathcal{N}_\Phi(\Omega)$ independent of the particular choice of Ξ?

To answer this question we need the following definition:

$$L_{\mathcal{P}}(\Omega) = \left\{ \lambda_{N,\alpha,X} = \sum_{j=1}^{N} \alpha_j \delta_{x_j} : N \in \mathbb{N}, X = \{x_1, \ldots, x_N\} \subseteq \Omega, \right. \\ \left. \lambda_{N,\alpha,X}(p) = 0 \text{ for all } p \in \mathcal{P} \right\}, \tag{2.12}$$

that is a set of finitely supported linear functionals on $\mathcal{C}(\Omega)$ that vanish on $\mathcal{P}$. The last condition can be read as

$$\sum_{j=1}^{N} \alpha_j p(x_j) = 0 \quad \text{for all } p \in \mathcal{P}.$$

If we apply $\lambda_{N,\alpha,X} \in L_{\mathcal{P}}(\Omega)$ to Φ we obtain

$$\lambda^x_{N,\alpha,X}(\Phi(\cdot,x)) = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j),$$

that with equation (2.3) permits us to define the inner product

$$\langle \lambda_{N,\alpha,X}, \lambda_{M,\beta,Y} \rangle_\Phi = \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(x_j, x_k).$$

We note that every $\{N, \alpha, X\}$ determines uniquely a functional of $L_{\mathcal{P}}(\Omega)$: we can define a bijective relation between $L_{\mathcal{P}}(\Omega)$ and $F_\Phi(\Omega)$.

$$\begin{aligned} L_{\mathcal{P}}(\Omega) &\longrightarrow F_\Phi(\Omega) \\ \lambda &\longmapsto \lambda^x(\Phi(\cdot, x)), \end{aligned} \tag{2.13}$$

where λ^x means action with respect to the variable x.

If $\lambda = \sum_{j=1}^{N} \alpha_j \delta_{x_j} \in L_{\mathcal{P}}(\Omega)$ and $f = \sum_{k=1}^{M} \beta_k \Phi(\cdot, y_k) \in F_\Phi(\Omega)$ then

$$\begin{aligned} \lambda(f) = \sum_{j=1}^{N} \alpha_j f(x_j) &= \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(x_j, y_k) = \\ &= \left\langle \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^{M} \beta_k \Phi(\cdot, y_k) \right\rangle_\Phi = \\ &= \langle \lambda^x(\Phi(\cdot, x)), f \rangle_\Phi, \end{aligned}$$

which gives

$$|\lambda(f)| = |\langle \lambda^x(\Phi(\cdot, x)), f \rangle_\Phi| \le \|\lambda^x(\Phi(\cdot, x))\|_\Phi \|f\|_\Phi = \|\lambda\|_\Phi \|f\|_\Phi,$$

with equality when $f = \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j)$. This let us to interpret the norm defined on $L_{\mathcal{P}}(\Omega)$ as the dual norm of $F_\Phi(\Omega)^*$.

We proved that the functionals in $L_{\mathcal{P}}(\Omega)$ are continuous in $F_\Phi(\Omega)$, now we want to study the boundness of these functionals from a different perspective.

Theorem 2.7 *Suppose that Φ is conditionally positive definite on Ω with respect to $\mathcal{P}$. Define*

$$\mathcal{G} = \{f \in \mathcal{C}(\Omega) : |\lambda(f)| \le C_f \|\lambda\|_\Phi \text{ for each } \lambda \in L_{\mathcal{P}}(\Omega)\}. \tag{2.14}$$

In this space it is defined the semi-norm

$$|f|_{\mathcal{G}} = \sup_{\lambda \in L_{\mathcal{P}}(\Omega) \setminus \{0\}} \frac{|\lambda(f)|}{\|\lambda\|_\Phi} \le C_f. \tag{2.15}$$

Then $\mathcal{N}_\Phi(\Omega) = \mathcal{G}$ and $|\cdot|_{\mathcal{G}} = |\cdot|_{\mathcal{N}_\Phi(\Omega)}$.

Proof
Suppose that $f \in \mathcal{N}_\Phi(\Omega)$. From equation (2.9) we have that $F_\Phi(\Omega) \subseteq \mathcal{N}_\Phi(\Omega)$.

If $\lambda = \sum_{j=1}^N \alpha_j \delta_{x_j} \in L_\mathcal{P}(\Omega)$ then

$$\begin{aligned} \sum_{j=1}^N \alpha_j G(\cdot, x_j) &= \sum_{j=1}^N \alpha_j \left(\Phi(\cdot, x_j) - \sum_{k=1}^Q p_k(x_j)\Phi(\cdot, \xi_k) \right) = \\ &= \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) - \sum_{k=1}^Q \Phi(\cdot, \xi_k) \underbrace{\sum_{j=1}^N \alpha_j p_k(x_j)}_{=0} = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j). \end{aligned} \tag{2.16}$$

We apply λ to the formula $f(x) = \Pi_\mathcal{P}(f)(x) + \langle f, G(\cdot, x)\rangle_{\mathcal{N}_\Phi(\Omega)}$ we proved in Theorem 2.3.

$$\begin{aligned} \lambda(f) &= \sum_{j=1}^N \alpha_j f(x_j) = \underbrace{\sum_{j=1}^N \alpha_j \Pi_\mathcal{P}(f)(x_j)}_{=\lambda(\Pi_\mathcal{P}(f))=0} + \sum_{j=1}^N \alpha_j \langle f, G(\cdot, x_j)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \left\langle f, \sum_{j=1}^N \alpha_j G(\cdot, x_j) \right\rangle_{\mathcal{N}_\Phi(\Omega)} = \left\langle f, \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \right\rangle_{\mathcal{N}_\Phi(\Omega)} \leq \\ &\leq |f|_{\mathcal{N}_\Phi(\Omega)} \left| \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \right|_{\mathcal{N}_\Phi(\Omega)} \overset{\text{eq. (2.11)}}{=} |f|_{\mathcal{N}_\Phi(\Omega)} \left\| \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \right\|_\Phi = \\ &= |f|_{\mathcal{N}_\Phi(\Omega)} \|\lambda^x(\Phi(\cdot, x))\|_\Phi = |f|_{\mathcal{N}_\Phi(\Omega)} \|\lambda\|_\Phi. \end{aligned} \tag{2.17}$$

This implies $|\lambda(f)| \leq |f|_{\mathcal{N}_\Phi(\Omega)} \|\lambda\|_\Phi$ for all $\lambda \in L_\mathcal{P}(\Omega)$, so $f \in \mathcal{G}$ and $|f|_\mathcal{G} \leq |f|_{\mathcal{N}_\Phi(\Omega)}$.

We prove the reverse inclusion. Suppose that $f \in \mathcal{G}$ so we can define a continuous linear functional

$$\begin{aligned} F_f : F_\Phi(\Omega) &\longrightarrow \mathbb{R} \\ \lambda^x(\Phi(\cdot, x)) &\longmapsto \lambda(f), \end{aligned}$$

which is well-defined because of the one-to-one correspondence between $F_\Phi(\Omega)$ and $L_\mathcal{P}(\Omega)$ (equation (2.13)). The continuity follows from the definition of $\mathcal{G}$: $|F_f(\lambda^x(\Phi(\cdot, x)))| = |\lambda(f)| \leq C_f \|\lambda\|_\Phi = C_f \|\lambda^x(\Phi(\cdot, x))\|_\Phi$.

We can extend F_f to a continuous functional of $\mathcal{F}_\Phi(\Omega)$ because $\overline{F_\Phi(\Omega)} = \mathcal{F}_\Phi(\Omega)$ and $\mathcal{F}_\Phi(\Omega)$ is a Hilbert space (Theorem 1.13). Using Theorem 1.6 there exists $S_f \in \mathcal{F}_\Phi(\Omega)$ such that

$$F_f(g) = \langle g, S_f\rangle_\Phi \quad \text{for all } g \in \mathcal{F}_\Phi(\Omega).$$

If we prove that $f - R(S_f) \in \mathcal{P}$ then we can conclude. If $\mu = \sum_{j=1}^{N} \alpha_j \delta_{x_j} \in L_{\mathcal{P}}(\Omega)$ then

$$\mu(R(S_f)) = \sum_{j=1}^{N} \alpha_j R(S_f)(x_j) = \sum_{j=1}^{N} \alpha_j \langle S_f, G(\cdot, x_j)\rangle_\Phi = \left\langle S_f, \sum_{j=1}^{N} \alpha_j G(\cdot, x_j) \right\rangle_\Phi \overset{\text{eq. (2.16)}}{=}$$

$$= \left\langle S_f, \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) \right\rangle_\Phi = \langle S_f, \mu^x(\Phi(\cdot, x))\rangle_\Phi = F_f(\mu^x(\Phi(\cdot, x))) = \mu(f),$$

so $\mu(f - R(S_f)) = \mu(f) - \mu(f) = 0$. Remembering the equation (2.4) and that Ξ is $\mathcal{P}$-unisolvent then $\delta_{(x)} \in L_{\mathcal{P}}(\Omega)$. We recall that $R(S_f)$ vanishes on Ξ (equation (2.8)), so choosing $\mu = \delta_{(x)}$ we obtain

$$f(x) - \sum_{k=1}^{Q} p_k(x) f(\xi_k) = R(S_f)(x) \quad \text{for all } x \in \Omega,$$

that implies $f \in \mathcal{N}_\Phi(\Omega)$.

We need to show that $|f|_{\mathcal{G}} \geq |f|_{\mathcal{N}_\Phi(\Omega)}$. We can find a sequence $\{\lambda_j\}_{j\in\mathbb{N}} \subseteq L_{\mathcal{P}}(\Omega)$ such that $\lambda_j^x(\Phi(\cdot, x)) \xrightarrow{n\to\infty} S_f \in \mathcal{F}_\Phi(\Omega)$, because $\mathcal{F}_\Phi(\Omega) = \overline{F_\Phi(\Omega)}$. We can write the following limit relation:

$$\lim_{j\to\infty} \lambda_j(f) = \lim_{j\to\infty} F_f(\lambda_j^x(\Phi(\cdot, x))) = \lim_{j\to\infty} \langle \lambda_j^x(\Phi(\cdot, x)), S_f\rangle_\Phi = \|S_f\|_\Phi^2,$$

$$\lim_{j\to\infty} \|\lambda_j\|_\Phi = \lim_{j\to\infty} \|\lambda_j^x(\Phi(\cdot, x))\|_\Phi = \|S_f\|_\Phi.$$

Before concluding we note that, since $\mathcal{P}$ is the null space of $\langle \cdot, \cdot \rangle_{\mathcal{N}_\Phi(\Omega)}$,

$$|f|^2_{\mathcal{N}_\Phi(\Omega)} = \langle f, f\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle \underbrace{f - R(S_f)}_{\in\mathcal{P}} + R(S_f), \underbrace{f - R(S_f)}_{\in\mathcal{P}} + R(S_f)\rangle_{\mathcal{N}_\Phi(\Omega)} =$$

$$= \langle R(S_f), R(S_f)\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle S_f, S_f\rangle_\Phi = \|S_f\|_\Phi^2.$$

If $\|S_f\|_\Phi \neq 0$, finally

$$|f|_{\mathcal{N}_\Phi(\Omega)} = \frac{\|S_f\|_\Phi^2}{\|S_f\|_\Phi} = \lim_{j\to\infty} \frac{\lambda_j(f)}{\|\lambda_j\|_\Phi} \leq \lim_{j\to\infty} \frac{|\lambda_j(f)|}{\|\lambda_j\|_\Phi} \leq |f|_{\mathcal{G}}.$$

□

As a consequence of Theorem 2.7 we can say that the native space $\mathcal{N}_\Phi(\Omega)$ of a conditionally positive definite kernel is independent of the choice of the $\mathcal{P}$-unisolvent set Ξ (also the semi-norm $|\cdot|_{\mathcal{N}_\Phi(\Omega)}$ is independent).

Now we will introduce some results that shows a sort of optimality of the interpolant $s_{f,X}$ (equation (2.1)) of a function f in the native space $\mathcal{N}_\Phi(\Omega)$.

Theorem 2.8 *Suppose that $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is $\mathcal{P}$-unisolvent. Denote the unique interpolant, based on a conditionally positive definite kernel Φ and the set X, of a function $f \in \mathcal{N}_\Phi(\Omega)$ by $s_{f,X}$ (equation (2.1)). Then we have*

$$\langle f - s_{f,X}, s \rangle_{\mathcal{N}_\Phi(\Omega)} = 0$$

for every $s \in (\langle \Phi(\cdot, x_j) : x_j \in X \rangle \cap F_\Phi(\Omega)) + \mathcal{P}$. In particular

$$\langle f - s_{f,X}, s_{f,X} \rangle_{\mathcal{N}_\Phi(\Omega)} = 0.$$

Proof
From equation (2.9) we know that $F_\Phi(\Omega) \subseteq \mathcal{N}_\Phi(\Omega)$. Any such function s can be written in the form $s = \lambda^x(\Phi(\cdot, x)) + \Pi_\mathcal{P}$ with a certain linear functional $\lambda = \sum_{j=1}^N \alpha_j \delta_{x_j} \in L_\mathcal{P}(\Omega)$ and $\Pi_\mathcal{P} \in \mathcal{P}$. From Theorem 2.3 follows that

$$(f - s_{f,X})(x) = \Pi_\mathcal{P}(f - s_{f,x})(x) + \langle f - s_{f,X}, G(\cdot, x) \rangle_{\mathcal{N}_\Phi(\Omega)} \qquad \text{for } x \in \Omega.$$

From equation (2.16) we get $\lambda^x(G(\cdot, x)) = \lambda^x(\Phi(\cdot, x))$ that combined with the fact that $\mathcal{P}$ in contained in $\ker(\lambda)$ and in the null space of $\langle \cdot, \cdot \rangle_{\mathcal{N}_\Phi(\Omega)}$ we can write

$$\begin{aligned}
0 = &\sum_{j=1}^N \alpha_j (f(x_j) - s_{f,X}(x_j)) = \lambda(f - s_{f,X}) = \\
=&\lambda(\Pi_\mathcal{P}(f - s_{f,X})) + \sum_{j=1}^N \alpha_j \langle f - s_{f,X}, G(\cdot, x_j) \rangle_{\mathcal{N}_\Phi(\Omega)} = \\
=&\left\langle f - s_{f,X}, \sum_{j=1}^N \alpha_j G(\cdot, x_j) \right\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle f - s_{f,X}, \lambda^x(G(\cdot, x)) \rangle_{\mathcal{N}_\Phi(\Omega)} = \\
=&\langle f - s_{f,X}, \lambda^x(\Phi(\cdot, x)) \rangle_{\mathcal{N}_\Phi(\Omega)} = \\
=&\langle f - s_{f,X}, \lambda^x(\Phi(\cdot, x)) + \Pi_\mathcal{P} \rangle_{\mathcal{N}_\Phi(\Omega)} - \underbrace{\langle f - s_{f,X}, \Pi_\mathcal{P} \rangle_{\mathcal{N}_\Phi(\Omega)}}_{=0} = \\
=&\langle f - s_{f,X}, s \rangle_{\mathcal{N}_\Phi(\Omega)}.
\end{aligned}$$

□

From this orthogonality theorem follows some useful property.

Theorem 2.9 *With the same hypothesis of Theorem 2.8 we have the estimates*

$$|s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \le |f|_{\mathcal{N}_\Phi(\Omega)} \quad \text{and} \quad |f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \le |f|_{\mathcal{N}_\Phi(\Omega)}.$$

Proof
Since $f = f - s_{f,X} + s_{f,X}$, from Theorem 2.8 and the Pythagorean law we obtain

$$|f - s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} + |s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} = |f|^2_{\mathcal{N}_\Phi(\Omega)}.$$

The statements of the theorem holds because $|f - s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} \leq |f - s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} + |s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)}$ and $|s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} \leq |f - s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} + |s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)}$.

□

A more practical characterization of the native space is the following.

Theorem 2.10 *Let Φ be a conditionally positive definite kernel on Ω with respect to $\mathcal{P}$. Denote by $s_{f,X}$ the interpolant to a function $f \in \mathcal{C}(\Omega)$ based on a $\mathcal{P}$-unisolvent set X (equation (2.1)). Then f belongs to the native space $\mathcal{N}_\Phi(\Omega)$ if and only if there exists a constant c_f such that $|s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \leq c_f$ for all $\mathcal{P}$-unisolvent set $X \subseteq \Omega$. Moreover, in the case $f \in \mathcal{N}_\Phi(\Omega)$ the smallest possible constant c_f is given by $|f|_{\mathcal{N}_\Phi(\Omega)}$.*

Proof
An implication follows from Theorem 2.9, which gives also $c_f \leq |f|_{\mathcal{N}_\Phi(\Omega)}$ if c_f is the minimal choice.

Let us prove the other implication. By hypothesis $|s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \leq c_f$ for all $\mathcal{P}$-unisolvent set $X \subseteq \Omega$. Fix an arbitrary

$$\lambda_{N,\alpha,X} = \sum_{j=1}^{N} \alpha_j \delta_{x_j} \in L_\mathcal{P}(\Omega),$$

we can choose a $\mathcal{P}$-unisolvent set Y with $X \subseteq Y$. If $s_{f,Y} \in \mathcal{N}_\Phi(\Omega)$ is the interpolant of f on the set Y then $\lambda_{N,\alpha,X}(f - s_{f,Y}) = 0$.

$$\begin{aligned} &|\lambda_{N,\alpha,X}(f)| \leq |\lambda_{N,\alpha,X}(f - s_{f,Y})| + |\lambda_{N,\alpha,X}(s_{f,Y})| = \\ =&|\lambda_{N,\alpha,X}(s_{f,Y})| \overset{\text{Theorem 2.7}}{\leq} |s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \|\lambda_{N,\alpha,X}\|_\Phi \leq \\ \leq& c_f \|\lambda_{N,\alpha,X}\|_\Phi. \end{aligned}$$

Since this is true for all $\lambda \in L_\mathcal{P}(\Omega)$ by Theorem 2.7 we conclude $f \in \mathcal{N}_\Phi(\Omega)$ and $|f|_{\mathcal{N}_\Phi(\Omega)} \leq c_f$ (equation (2.15)).

□

We briefly return to consider only positive definite kernels. With Theorem 1.23 we proved that $\mathcal{N}_\Phi(\mathbb{R}^d) \subseteq L^2(\mathbb{R}^d)$. If Ω is compact then we can state a more precise result.

Theorem 2.11 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is compact and Φ is a positive definite kernel on Ω. Then the native space $\mathcal{N}_\Phi(\Omega)$ has a continuous linear embedding into $L^2(\Omega)$.*

Proof
Since Φ is the reproducing kernel of $\mathcal{N}_\Phi(\Omega)$ (Definition 1.7) we get the following chain of inequalities for $f \in \mathcal{N}_\Phi(\Omega)$ and $x \in \Omega$.

$$|f(x)|^2 = |\langle f, \Phi(\cdot,x)\rangle_{\mathcal{N}_\Phi(\Omega)}|^2 \leq \|f\|^2_{\mathcal{N}_\Phi(\Omega)} \|\Phi(\cdot,x)\|^2_{\mathcal{N}_\Phi(\Omega)} \overset{\text{eq. (1.11)}}{=} \\ = \|f\|^2_{\mathcal{N}_\Phi(\Omega)} \|\Phi(\cdot,x)\|^2_\Phi = \|f\|^2_{\mathcal{N}_\Phi(\Omega)} \Phi(x,x).$$

We can conclude with

$$\|f\|_{L^2(\Omega)} \leq \sqrt{\int_\Omega \Phi(x,x) \mathrm{d}\, x}\, \|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

□

Now we introduce the integral operator $T : L^2(\Omega) \to L^2(\Omega)$ defined by

$$T(v)(x) = \int_\Omega \Phi(x,y) v(y) \mathrm{d}\, y. \tag{2.18}$$

First of all, T is well-defined and continuous because since Ω is compact $L^2(\Omega) \subseteq L^1(\Omega)$ (Hölder's inequality). Indeed,

$$\int_\Omega \left| \int_\Omega \Phi(x,y) v(y) \mathrm{d}\, y \right|^2 \mathrm{d}\, x \leq \int_\Omega \left(\int_\Omega |\Phi(x,y) v(y)| \mathrm{d}\, y \right)^2 \mathrm{d}\, x \leq \\ \leq \int_\Omega \|\Phi\|^2_{L^\infty(\Omega\times\Omega)} \left(\int_\Omega |v(y)| \mathrm{d}\, y \right)^2 \mathrm{d}\, x \leq \mathcal{L}(\Omega) \|\Phi\|^2_{L^\infty(\Omega\times\Omega)} \|v\|^2_{L^1(\Omega)} < +\infty,$$

where $\mathcal{L}$ is the Lebesgue measure on $\mathbb{R}^d$. Moreover, $T(v) : \Omega \to \mathbb{R}$ is continuous for every $v \in L^2(\Omega)$ because of the uniform continuity of Φ in $\Omega \times \Omega$.

$$|T(v)(x) - T(v)(z)| \leq \int_\Omega |\Phi(x,y) - \Phi(z,y)| |v(y)| \mathrm{d}\, y \leq \\ \leq \left(\int_\Omega |\Phi(x,y) - \Phi(z,y)|^2 \mathrm{d}\, y \right)^{\frac{1}{2}} \left(\int_\Omega |v(y)|^2 \mathrm{d}\, y \right)^{\frac{1}{2}}. \tag{2.19}$$

With Ascoli-Arzelà Theorem we can prove that T is a compact operator, i.e. for every bounded sequence $\{v_n\}_{n\in\mathbb{N}} \subseteq L^2(\Omega)$ then $\{T(v_n)\}_{n\in\mathbb{N}} \subseteq L^2(\Omega)$ admits a converging subsequence. If $\|v_n\|_{L^2(\Omega)} \leq M$ for all $n \in \mathbb{N}$ then

$$|T(v_n)(x)| = \left| \int_\Omega \Phi(x,y) v_n(y) \mathrm{d}\, y \right| \leq \int_\Omega |\Phi(x,y)| |v_n(y)| \mathrm{d}\, y \leq \\ \leq \left(\int_\Omega |\Phi(x,y)|^2 \mathrm{d}\, y \right)^{\frac{1}{2}} \left(\int_\Omega |v_n(y)|^2 \mathrm{d}\, y \right)^{\frac{1}{2}} \leq \mathcal{L}(\Omega) \|\Phi\|_{L^\infty(\Omega\times\Omega)} M \quad \text{for all } x \in \Omega,$$

which proves that $\{T(v_n)\}_{n\in\mathbb{N}}$ is uniformly bounded. Equicontinuity of $\{T(v_n)\}_{n\in\mathbb{N}}$ follows from equation (2.19). From Ascoli-Arzelà Theorem we get that $\overline{\{T(v_n)\}}_{n\in\mathbb{N}}$ is compact in $(\mathcal{C}(\Omega), \|\cdot\|_\infty)$. We can conclude by remarking that the uniform convergence in a bounded domain Ω implies $L^2(\Omega)$ convergence.

Theorem 2.12 *Suppose that Φ is a symmetric positive definite kernel of the compact set $\Omega \subseteq \mathbb{R}^d$. Then the integral operator T (equation (2.18)) maps $L^2(\Omega)$ continuously into the native space $\mathcal{N}_\Phi(\Omega)$. It is the adjoint of the embedding operator of the native space $\mathcal{N}_\Phi(\Omega)$ into $L^2(\Omega)$, i.e. satisfies*

$$\langle f, v\rangle_{L^2(\Omega)} = \langle f, T(v)\rangle_{\mathcal{N}_\Phi(\Omega)}, \quad f \in \mathcal{N}_\Phi(\Omega), v \in L^2(\Omega).$$

The range of T is dense in $\mathcal{N}_\Phi(\Omega)$.

Proof
To prove that $T(L^2(\Omega)) \subseteq \mathcal{N}_\Phi(\Omega)$ we use Theorem 2.7. Fix $\lambda = \sum_{j=1}^N \alpha_j \delta_{x_j} \in L(\Omega)$.

We study

$$\begin{aligned}\lambda(T(v)) &= \sum_{j=1}^N \alpha_j T(v)(x_j) = \sum_{j=1}^N \alpha_j \int_\Omega \Phi(x_j, y) v(y) \mathrm{d}\, y = \\ &= \int_\Omega \sum_{j=1}^N \alpha_j \Phi(x_j, y) v(y) \mathrm{d}\, y = \int_\Omega \lambda^x(\Phi(y, x)) v(y) \mathrm{d}\, y,\end{aligned}$$

for $v \in L^2(\Omega)$.

From Theorem 2.11 and equation (2.11) we get

$$\begin{aligned}|\lambda(T(v))| &\leq \|v\|_{L^2(\Omega)} \|\lambda^x(\Phi(\cdot, x))\|_{L^2(\Omega)} \leq \\ &\leq C\|v\|_{L^2(\Omega)} \|\lambda^x(\Phi(\cdot, x))\|_{\mathcal{N}_\Phi(\Omega)} = \\ &= \|v\|_{L^2(\Omega)} \|\lambda^x(\Phi(\cdot, x))\|_\Phi = C\|v\|_{L^2(\Omega)} \|\lambda\|_\Phi,\end{aligned}$$

where C is the operator norm of the inclusion. This gives $\|T(v)\|_{\mathcal{N}_\Phi(\Omega)} \leq C\|v\|_{L^2(\Omega)}$.

To continue the proof we use a density argument. Fix $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega)$.

$$\begin{aligned}\langle f, v\rangle_{L^2(\Omega)} &= \int_\Omega \left(\sum_{j=1}^N \alpha_j \Phi(y, x_j)\right) v(y) \mathrm{d}\, y = \sum_{j=1}^N \alpha_j \int_\Omega \Phi(y, x_j) v(y) \mathrm{d}\, y = \\ &= \sum_{j=1}^N \alpha_j \underbrace{T(v)}_{\in \mathcal{N}_\Phi(\Omega)} (x_j) = \sum_{j=1}^N \alpha_j \langle T(v), \Phi(\cdot, x_j)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \left\langle T(v), \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j)\right\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle T(v), f\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle f, T(v)\rangle_{\mathcal{N}_\Phi(\Omega)}.\end{aligned} \tag{2.20}$$

By density of $F_\Phi(\Omega)$ in $\mathcal{N}_\Phi(\Omega)$ (equation (1.10)) and Theorem 2.11 the equation (2.20) holds for $f \in \mathcal{N}_\Phi(\Omega)$.

To finish the proof we claim that $\ker(\iota) = \text{range}(T)^{\perp}$, where $\iota : \mathcal{N}_{\Phi}(\Omega) \to L^2(\Omega)$ is the inclusion of Theorem 2.11. From equation (2.20) we have

$$\langle \iota(f), v \rangle_{L^2(\Omega)} = \langle f, T(v) \rangle_{\mathcal{N}_\Phi},$$

that proves the claim. Using Proposition 1.4 and $\langle 0 \rangle = \ker(\iota) = \text{range}(T)^{\perp}$ we get

$$\mathcal{N}_{\Phi}(\Omega) = \overline{\text{range}(T)}.$$

□

From Theorem 2.12 we can derive

$$\langle T(v), v \rangle_{L_2(\Omega)} = \langle T(v), T(v) \rangle_{\mathcal{N}_\Phi(\Omega)} \geq 0,$$

for every $v \in L^2(\Omega)$. Because of the compactness of T, we can apply Mercer's theorem [28] and obtain a countable set of positive eigenvalues $\{\rho_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_{\geq 0}$ and a countable set of continuous eigenfunctions $\{\phi_n\}_{n\in\mathbb{N}} \subseteq L^2(\Omega)$ such that $\rho_n \geq \rho_{n+1}$ and $T(\varphi_n) = \rho_n \varphi_n$ for all $n \in \mathbb{N}$. In particular, $\{\varphi_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for $L^2(\Omega)$ and Φ can be written as an absolutely convergent series

$$\Phi(x, y) = \sum_{n=1}^{+\infty} \rho_n \varphi_n(x) \varphi_n(y), \quad x, y \in \Omega.$$

Before continuing it is important the following proposition.

Proposition 2.13 *Suppose* $\{a_n^{(j)}\}_{n\in\mathbb{N}}$ *is a convergent series for* $j = 1, \ldots, N$ *then*

$$\sum_{j=1}^{M} \sum_{n=1}^{+\infty} a_n^{(j)} = \sum_{n=1}^{+\infty} \sum_{j=1}^{M} a_n^{(j)}.$$

Proof

$$\left| \sum_{n=1}^{N} \sum_{j=1}^{M} a_n^{(j)} - \sum_{j=1}^{M} \sum_{n=1}^{+\infty} a_n^{(j)} \right| = \left| \sum_{j=1}^{M} \left(\sum_{n=1}^{N} a_n^{(j)} - \sum_{n=1}^{+\infty} a_n^{(j)} \right) \right| \leq$$

$$\leq \sum_{j=1}^{M} \left| \sum_{n=1}^{N} a_n^{(j)} - \sum_{n=1}^{+\infty} a_n^{(j)} \right| \xrightarrow[N\to+\infty]{} 0$$

□

Now we can give a final characterization of the native space for a positive definite kernel. It will be useful to improve the error estimates for radial basis function interpolant (equation (2.1)).

Theorem 2.14 *Suppose Φ is a symmetric positive definite kernel on a compact set $\Omega \subseteq \mathbb{R}^d$. Then its native space is given by*

$$\mathcal{N}_\Phi(\Omega) = \left\{ f \in L^2(\Omega) : \sum_{n=1}^{+\infty} \frac{1}{\rho_n} |\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2 < +\infty \right\}$$

and the inner product has the representation

$$\langle f, g \rangle_{\mathcal{N}_\Phi(\Omega)} = \sum_{n=1}^{+\infty} \frac{1}{\rho_n} \langle f, \varphi_n \rangle_{L^2(\Omega)} \langle g, \varphi_n \rangle_{L^2(\Omega)} \qquad \textit{for all } f, g \in \mathcal{N}_\Phi(\Omega).$$

Proof
Let us fix some notation. We denote

$$\mathcal{G} = \left\{ f \in L^2(\Omega) : \sum_{n=1}^{+\infty} \frac{1}{\rho_n} |\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2 < +\infty \right\}$$

and

$$\langle f, g \rangle_{\mathcal{G}} = \sum_{n=1}^{+\infty} \frac{1}{\rho_n} \langle f, \varphi_n \rangle_{L_2(\Omega)} \langle g, \varphi_n \rangle_{L_2(\Omega)} \qquad \text{for all } f, g \in \mathcal{G}.$$

Fix $f \in \mathcal{G}$. We claim that $f \in \mathcal{N}_\Phi(\Omega)$. Since $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq L^2(\Omega)$ is a set of continuous functions that in particular are an orthonormal basis for $L^2(\Omega)$, f admits the following form $f = \sum_{n=1}^{+\infty} \alpha_n \varphi_n$. For every $m \in \mathbb{N}$ we compute

$$\langle f, \varphi_m \rangle_{L^2(\Omega)} = \left\langle \sum_{n=1}^{+\infty} \alpha_n \varphi_n, \varphi_m \right\rangle_{L^2(\Omega)} = \sum_{n=1}^{+\infty} \alpha_n \langle \varphi_n, \varphi_m \rangle_{L^2(\Omega)} = \alpha_m.$$

At this point,

$$f(x) = \sum_{n=1}^{+\infty} \langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x) \qquad \text{for a.e. } x \in \Omega.$$

We study the absolute convergence of this series for every $x \in \Omega$.

Fix $M \in \mathbb{N}$ then

$$\begin{aligned}
\left| \sum_{n=M+1}^{+\infty} \langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x) \right| &\leq \sum_{n=M+1}^{+\infty} |\langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x)| = \\
&= \sum_{n=M+1}^{+\infty} \frac{|\langle f, \varphi_n \rangle_{L^2(\Omega)}|}{\sqrt{\rho_n}} |\varphi_n(x) \sqrt{\rho_n}| \leq \\
&\leq \left(\sum_{n=M+1}^{+\infty} \frac{|\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n} \right)^{\frac{1}{2}} \left(\sum_{n=M+1}^{+\infty} |\varphi_n(x)|^2 \rho_n \right)^{\frac{1}{2}} \leq \\
&\leq \left(\sum_{n=M+1}^{+\infty} \frac{|\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n} \right)^{\frac{1}{2}} \sqrt{\Phi(x,x)} \leq \\
&\leq \left(\underbrace{\sum_{n=M+1}^{+\infty} \frac{|\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n}}_{\text{convergent series}} \right)^{\frac{1}{2}} \|\Phi\|_{L^\infty(\Omega \times \Omega)} < \infty.
\end{aligned} \tag{2.21}$$

If $M = 0$ we proved that f is an absolute convergent series of continuous functions. Let us focus on the continuity of f.

$$\begin{aligned}
|f(x) - f(x_0)| \leq & \left| f(x) - \sum_{n=1}^{M} \langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x) \right| + \\
& + \left| \sum_{n=1}^{M} \langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x) - \sum_{n=1}^{M} \langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x_0) \right| + \\
& + \left| \sum_{n=1}^{M} \langle f, \varphi_n \rangle_{L^2(\Omega)} \varphi_n(x_0) - f(x_0) \right| < \varepsilon.
\end{aligned}$$

The first and third term of the right-hand side of the inequality can be made smaller than $\frac{\varepsilon}{3}$ choosing an appropriate $M \in \mathbb{N}$ (equation (2.21) is independent of $x \in \Omega$). When M is fixed the second term is a difference of continuous functions, so if x is close to x_0 we can conclude.

We want to use Theorem 2.7. Fix $\lambda = \sum_{j=1}^{N} \alpha_j \delta_{x_j} \in L(\Omega)$. With proposition 2.13 we have

$$\begin{aligned}
\lambda(f) = \sum_{j=1}^{N} \alpha_j f(x_j) = \sum_{j=1}^{N} \sum_{n=1}^{+\infty} \alpha_j \langle f, \varphi_n \rangle_{L_2(\Omega)} \varphi_n(x_j) = \\
= \sum_{n=1}^{+\infty} \langle f, \varphi_n \rangle_{L_2(\Omega)} \sum_{j=1}^{N} \alpha_j \varphi_n(x_j) = \sum_{n=1}^{+\infty} \langle f, \varphi_n \rangle_{L_2(\Omega)} \lambda(\varphi_n)
\end{aligned}$$

and

$$\sum_{n=1}^{+\infty} \rho_n |\lambda(\varphi_n)|^2 = \sum_{n=1}^{+\infty} \rho_n \left(\sum_{j,k=1}^{N} \alpha_j \alpha_k \varphi_n(x_j) \varphi_n(x_k) \right) =$$

$$= \sum_{j,k=1}^{N} \alpha_j \alpha_k \sum_{n=1}^{+\infty} \rho_n \varphi_n(x_j) \varphi_n(x_k) = \sum_{j,k=1}^{N} \alpha_j \alpha_k \Phi(x_j, x_k) = \|\lambda\|_\Phi.$$

So,

$$|\lambda(f)| = \left| \sum_{n=1}^{+\infty} \langle f, \varphi_n \rangle_{L^2(\Omega)} \lambda(\varphi_n) \right| \leq \left| \sum_{n=1}^{+\infty} \frac{\langle f, \varphi_n \rangle_{L^2(\Omega)}}{\sqrt{\rho_n}} \lambda(\varphi_n) \sqrt{\rho_n} \right| \leq$$

$$\leq \left(\sum_{n=1}^{+\infty} \frac{|\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n} \right)^{\frac{1}{2}} \left(\sum_{n=1}^{+\infty} |\lambda(\varphi_n)|^2 \rho_n \right)^{\frac{1}{2}} = \|f\|_{\mathcal{G}} \|\lambda\|_\Phi,$$

which proves that $f \in \mathcal{N}_\Phi(\Omega)$ and $\|f\|_{\mathcal{N}_\Phi(\Omega)} \leq \|f\|_{\mathcal{G}}$.

Now we will use the density of $T(L^2(\Omega))$ in $\mathcal{N}_\Phi(\Omega)$. If we fix $v \in L^2(\Omega)$ such that $v = \sum_{n=1}^{+\infty} \langle v, \varphi_n \rangle_{L^2(\Omega)} \varphi_n$ then

$$T(v) = \sum_{n=1}^{+\infty} \langle v, \varphi_n \rangle_{L^2(\Omega)} T(\varphi_n) = \sum_{n=1}^{+\infty} \langle v, \varphi_n \rangle_{L^2(\Omega)} \rho_n \varphi_n.$$

Since $\{\varphi_n\}_{n \in \mathbb{N}} \subseteq L^2(\Omega)$ is an orthonormal basis for $L^2(\Omega)$, with Theorem 2.12 we obtain

$$\begin{aligned} \|T(v)\|^2_{\mathcal{N}_\Phi(\Omega)} &= \langle T(v), T(v) \rangle_{\mathcal{N}_\Phi(\Omega)} = \langle T(v), v \rangle_{L^2(\Omega)} = \\ &= \sum_{n=1}^{+\infty} \langle v, \varphi_n \rangle_{L^2(\Omega)} \langle v, \varphi_n \rangle_{L^2(\Omega)} \rho_n = \\ &= \sum_{n=1}^{+\infty} \frac{|\langle T(v), \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n} = \|T(v)\|_{\mathcal{G}} \end{aligned} \tag{2.22}$$

This proves that $\| \cdot \|_{\mathcal{N}_\Phi(\Omega)}$ and $\| \cdot \|_{\mathcal{G}}$ coincide on $T(L^2(\Omega))$. From Theorem 2.12 we know that $T(L^2(\Omega))$ is dense in $\mathcal{N}_\Phi(\Omega)$, so if $f \in \mathcal{N}_\Phi(\Omega)$ then there exists a sequence $\{f_j\}_{j \in \mathbb{N}} \subseteq T(L^2(\Omega))$ such that $f_j \xrightarrow[j \to +\infty]{\| \cdot \|_{\mathcal{N}_\Phi(\Omega)}} f$. From equation (2.22) we get

$$\sum_{n=1}^{N} \frac{|\langle f_j, \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n} \leq \|f_j\|^2_{\mathcal{N}_\Phi(\Omega)},$$

for $N, j \in \mathbb{N}$. If we remark Theorem 2.11 and fix $N \in \mathbb{N}$, letting $j \to +\infty$ we obtain

$$\sum_{n=1}^{N} \frac{|\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2}{\rho_n} \leq \|f\|^2_{\mathcal{N}_\Phi(\Omega)}.$$

Now letting letting $N \to +\infty$, we can conclude $\|f\|_{\mathcal{G}} \leq \|f\|_{\mathcal{N}_\Phi(\Omega)}$.

□

Using Picard's theorem [29] on the range of a compact operator we can state

Theorem 2.15 *Suppose that Φ is a symmetric positive definite kernel on a compact set $\Omega \subseteq \mathbb{R}^d$. Then the range of the integral operator in equation 2.18 is given by*

$$T(L^2(\Omega)) = \left\{ f \in L^2(\Omega) : \sum_{n=1}^{+\infty} \frac{1}{\rho_n} |\langle f, \varphi_n \rangle_{L^2(\Omega)}|^2 < +\infty \right\}.$$

For the error analysis of the interpolation process in equation (2.1) will be useful the following theorems, that we will state without proofs. We see that the native space $\mathcal{N}_\Phi(\Omega)$ is not only constituted by continuous function if Φ is smooth.

Theorem 2.16 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is open and that $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ is a conditionally positive definite symmetric kernel with respect to $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$. The function $G(\cdot, \cdot)$ in equation (2.5) is k-times continuously differentiable with respect to the second argument, and for every $x \in \Omega$ and every $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ the function $D_2^\alpha(G(\cdot, x)) \in \mathcal{N}_\Phi(\Omega)$. D_2^α denotes that we differentiate with respect to the second argument.*

Similarly to Theorem 2.3 we have

Theorem 2.17 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is open and that $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ is a conditionally positive definite symmetric kernel with respect to $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$. Then $\mathcal{N}_\Phi(\Omega) \subseteq \mathcal{C}^k(\Omega)$ and for every $f \in \mathcal{N}_\Phi(\Omega), \alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ and $x \in \Omega$*

$$D^\alpha(f)(x) = D^\alpha(\Pi_{\mathcal{P}}(f))(x) + \langle f, D_2^\alpha(G(\cdot, x)) \rangle_{\mathcal{N}_\Phi(\Omega)}.$$

From Theorem 1.25 we know that if

$$c_1(1 + \|\omega\|_2^2)^{-s} \leq \widehat{\Phi}(\omega) \leq c_2(1 + \|\omega\|_2^2)^{-s} \qquad \omega \in \mathbb{R}^d$$

then $\mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d)$.

From Fourier inversion formula we obtain

$$\Phi(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{\Phi}(\omega) e^{i\langle x, \omega \rangle} \mathrm{d}\,\omega, \quad \text{ for each } x \in \Omega.$$

From this expression we get that if Φ is a positive definite function in $L^1(\mathbb{R}^d)$ that satisfies the condition of Theorem 1.25 with $s > k + \frac{d}{2}$ then Theorem 2.16 guarantees that $H^s(\mathbb{R}^d) \subseteq \mathcal{C}^k(\mathbb{R}^d)$, which is Sobolev's embedding theorem.

2.3 Extension of native spaces

If we decide to work with a domain Ω that is bounded and it has Lipschitz boundary then we can use an extension theorem for Sobolev spaces [30, 31].

Theorem 2.18 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is open and it has a Lipschitz boundary. If $s \geq 0$ then there exists a linear mapping $E : H^s(\Omega) \to H^s(\mathbb{R}^d)$ and a constant $C_s \in \mathbb{R}_{>0}$, such that*

- $E(f)_{|\Omega} = f$ *for* $f \in H^s(\Omega)$,
- $\|E(f)\|_{H^s(\mathbb{R}^d)} \leq C_s \|f\|_{H^s(\Omega)}$ *for* $f \in H^s(\Omega)$.

Since the extension operator in Theorem 2.18 is continuous then we can define an equivalent norm on $H^s(\Omega)$ by the position $\|f\| = \|E(f)\|_{H^s(\mathbb{R}^d)}$ because

$$\|f\|_{H^s(\Omega)} \leq \|E(f)\|_{H^s(\mathbb{R}^d)} \leq C_s \|f\|_{H^s(\Omega)}.$$

This result let us to give a characterization of the norm of the native space $\mathcal{N}_\Phi(\Omega)$ of a positive definite function when the domain $\Omega \subsetneq \mathbb{R}^d$.

Let us fix some assumptions. We work with two sets Ω_1 and Ω_2 such that $\Omega_1 \subseteq \Omega_2 \subseteq \mathbb{R}^d$. We consider a conditionally positive definite kernel $\Phi : \Omega_2 \times \Omega_2 \to \mathbb{R}$ with respect to $\mathcal{P} \subseteq \mathcal{C}(\Omega_2)$. To define the native spaces $\mathcal{N}_\Phi(\Omega_1)$ and $\mathcal{N}_\Phi(\Omega_2)$ we need a $\mathcal{P}$-unisolvent set Ξ that we suppose contained in Ω_1.

Theorem 2.19 *Each function $f \in \mathcal{N}_\Phi(\Omega_1)$ admits an extension $E(f) \in \mathcal{N}_\Phi(\Omega_2)$ such that $|E(f)|_{\mathcal{N}_\Phi(\Omega_2)} = |f|_{\mathcal{N}_\Phi(\Omega_1)}$.*

Proof
We begin by building an isometric embedding between $F_\Phi(\Omega_1)$ and $F_\Phi(\Omega_2)$. Since $\Omega_1 \subseteq \Omega_2$ we can define

$$\begin{aligned} \varepsilon : F_\Phi(\Omega_1) &\longrightarrow F_\Phi(\Omega_2) \subseteq \mathcal{F}_\Phi(\Omega_2) \\ \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) &\longmapsto \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j), \end{aligned}$$

with $\{x_1, \ldots, x_N\} \subseteq \Omega_1$ and $\sum_{j=1}^N \alpha_j p(x_j) = 0$ for $p \in \mathcal{P}$. ε is a linear isometric embedding because if $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega_1)$ then

$$\|f\|_{\Phi,\Omega_1} = \sum_{j,k=1}^N \alpha_j \alpha_k \Phi(x_j, x_k) = \|\varepsilon(f)\|_{\Phi,\Omega_2},$$

moreover if $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \in F_\Phi(\Omega_1)$ and $g = \sum_{k=1}^m \beta_k \Phi(\cdot, y_k) \in F_\Phi(\Omega_1)$ then

$$\langle f, g \rangle_{\Phi,\Omega_1} = \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \Phi(x_j, x_k) = \langle \varepsilon(f), \varepsilon(g) \rangle_{\Phi,\Omega_2}.$$

We can extend ε to $\mathcal{F}_\Phi(\Omega_1)$ in a isometric way: if $f \in \mathcal{F}_\Phi(\Omega_1) = \overline{F_\Phi(\Omega_1)}$ then we impose

$$\varepsilon(f) = \lim_{n \to +\infty} \varepsilon(f_n)$$

with $\{f_n\}_{n\in\mathbb{N}} \subseteq F_\Phi(\Omega_1)$ a Cauchy sequence that converges is $\mathcal{F}_\Phi(\Omega_1)$ to f. First of all the limit is well-defined because $\{\varepsilon(f_n)\}_{n\in\mathbb{N}} \subseteq \varepsilon(F_\Phi(\Omega_1))$ is a Cauchy sequence in the complete space $\mathcal{F}_\Phi(\Omega_2)$: $\|\varepsilon(f_n)-\varepsilon(f_m)\|_{\Phi,\Omega_2} = \|f_n - f_m\|_{\Phi,\Omega_1} \xrightarrow[n,m\to 0]{} 0$. Moreover the value of the limit does not depend on the sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq F_\Phi(\Omega_1)$ we choose, indeed if $\{f_n\}_{n\in\mathbb{N}}$ and $\{\widetilde{f}_n\}_{n\in\mathbb{N}}$ are two sequences in $F_\Phi(\Omega_1)$ that converge to $f \in \mathcal{F}_\Phi(\Omega_1)$ then

$$\|\varepsilon(f_n)-\varepsilon(\widetilde{f}_n)\|_{\Phi,\Omega_2} = \|f_n - \widetilde{f}_n\|_{\Phi,\Omega_1} \leq \|f_n - f\|_{\Phi,\Omega_1} + \|f - \widetilde{f}_n\|_{\Phi,\Omega_1} \xrightarrow{n\to+\infty} 0,$$

that implies if $\varepsilon(f)$ and $\varepsilon(\widetilde{f})$ are the limit of $\{f_n\}_{n\in\mathbb{N}}$ and $\{\widetilde{f}_n\}_{n\in\mathbb{N}}$ respectively

$$\|\varepsilon(f)-\varepsilon(\widetilde{f})\|_{\Phi,\Omega_2} \leq \|\varepsilon(f)-\varepsilon(f_n)\|_{\Phi,\Omega_2} + \|\varepsilon(f_n)-\varepsilon(\widetilde{f}_n)\|_{\Phi,\Omega_2} + \|\varepsilon(\widetilde{f}_n)-\varepsilon(\widetilde{f})\|_{\Phi,\Omega_2} \xrightarrow{n\to+\infty} 0.$$

The extension is continuous because

$$\|\varepsilon(f)\|_{\Phi,\Omega_2} = \lim_{n\to+\infty} \|\varepsilon(f_n)\|_{\Phi,\Omega_2} = \lim_{n\to+\infty} \|f_n\|_{\Phi,\Omega_1} = \|f\|_{\Phi,\Omega_1}$$

and it is isometric by

$$\langle \varepsilon(f), \varepsilon(g)\rangle_{\Phi,\Omega_2} = \lim_{n\to+\infty} \langle \varepsilon(f_n), \varepsilon(g_n)\rangle_{\Phi,\Omega_2} = \lim_{n\to+\infty} \langle f_n, g_n\rangle_{\Phi,\Omega_1} = \langle f, g\rangle_{\Phi,\Omega_1},$$

with $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ in $F_\Phi(\Omega_1)$ converging to $f, g \in \mathcal{F}_\Phi(\Omega_1)$ respectively. If $f \in \mathcal{N}_\Phi(\Omega_1) = R_{\Omega_1}(\mathcal{F}_\Phi(\Omega_1)) \oplus \mathcal{P}$ then $f = R_{\Omega_1}(\widetilde{f}) + p$ with $\widetilde{f} \in \mathcal{F}_\Phi(\Omega_1)$ and $p \in \mathcal{P}$. The decomposition is unique because R_{Ω_1} is injective (Theorem 2.2), Ω_1 contains a $\mathcal{P}$-unisolvent set and $R_{\Omega_1}(\mathcal{F}_\Phi(\Omega_1)) \cap \mathcal{P} = \langle 0\rangle$. We can define an extension operator as

$$\begin{aligned} E : \mathcal{N}_\Phi(\Omega_1) &\longrightarrow \mathcal{N}_\Phi(\Omega_2) \\ R_{\Omega_1}(\widetilde{f}) + p &\longmapsto R_{\Omega_2}(\varepsilon(\widetilde{f})) + p, \end{aligned}$$

with $\widetilde{f} \in \mathcal{F}_\Phi(\Omega_1)$ and $p \in \mathcal{P}$. We claim that $E(f)_{|\Omega_1} = f$ for $f \in \mathcal{N}_\Phi(\Omega_1)$ because if $x \in \Omega_1$

$$\begin{aligned} R_{\Omega_2}(\varepsilon(\widetilde{f}))(x) = \langle \varepsilon(\widetilde{f}), G_{\Omega_2}(\cdot, x)\rangle_{\Phi,\Omega_2} = \langle \varepsilon(\widetilde{f}), \varepsilon(G_{\Omega_1}(\cdot, x))\rangle_{\Phi,\Omega_2} = \\ = \langle \widetilde{f}, G_{\Omega_1}(\cdot, x)\rangle_{\Phi,\Omega_1} = R_{\Omega_1}(\widetilde{f})(x). \end{aligned}$$

To finish the proof we show that $E : \mathcal{N}_\Phi(\Omega_1) \to \mathcal{N}_\Phi(\Omega_2)$ is an isometric embedding. If $f = R_{\Omega_1}(\widetilde{f}) + p_f$ and $g = R_{\Omega_1}(\widetilde{g}) + p_g$ with $\widetilde{f}, \widetilde{g} \in \mathcal{F}_\Phi(\Omega_1)$ and $p_f, p_g \in \mathcal{P}$ then

$$\begin{aligned} &\langle E(f), E(g)\rangle_{\mathcal{N}_\Phi(\Omega_2)} = \langle R_{\Omega_2}(\varepsilon(\widetilde{f})) + p_f, R_{\Omega_2}(\varepsilon(\widetilde{g})) + p_g\rangle_{\mathcal{N}_\Phi(\Omega_2)} = \\ =&\langle R_{\Omega_2}(\varepsilon(\widetilde{f})), R_{\Omega_2}(\varepsilon(\widetilde{g}))\rangle_{\mathcal{N}_\Phi(\Omega_2)} = \langle \varepsilon(\widetilde{f}), \varepsilon(\widetilde{g})\rangle_{\Phi,\Omega_2} = \langle \widetilde{f}, \widetilde{g}\rangle_{\Phi,\Omega_1} = \\ =&\langle R_{\Omega_1}(\widetilde{f}), R_{\Omega_1}(\widetilde{g})\rangle_{\mathcal{N}_\Phi(\Omega_1)} = \langle R_{\Omega_1}(\widetilde{f}) + p_f, R_{\Omega_1}(\widetilde{g}) + p_g\rangle_{\mathcal{N}_\Phi(\Omega_1)} = \langle f, g\rangle_{\mathcal{N}_\Phi(\Omega_1)}, \end{aligned}$$

because $\mathcal{P}$ is in the null space of $\langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega_1)}$ and$\langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega_)}$. We remark that R_{Ω_1} and R_{Ω_2} are isometry (Theorem 2.4).

□

From Theorem 1.25 we know that a Sobolev space can be thought as a native space, in the same direction Theorem 2.19 proves that native spaces are a generalization of Sobolev spaces because the extension operator for native spaces works for general domains and not only for bounded lipschitz domains (Theorem 2.18).

Theorem 2.20 *The restriction $f_{|\Omega_1}$ of any function $f \in \mathcal{N}_\Phi(\Omega_2)$ is contained in $\mathcal{N}_\Phi(\Omega_1)$ and $|f_{|\Omega_1}|_{\mathcal{N}_\Phi(\Omega_1)} \leq |f|_{\mathcal{N}_\Phi(\Omega_2)}$.*

Proof
To prove the statement we use Theorem 2.7. If $\lambda = \sum_{j=1}^N \alpha_j \delta_{x_j} \in L_\mathcal{P}(\Omega_1)$ then $\lambda \in L_\mathcal{P}(\Omega_2)$ because $\{x_1, \ldots, x_N\} \subseteq \Omega_1 \subseteq \Omega_2$ and $\sum_{j=1}^N \alpha_j p(x_j) = 0$ for each $p \in \mathcal{P}$.

Since

$$|\lambda(f_{|\Omega_1})| = \left| \sum_{j=1}^N \alpha_j f_{|\Omega_1}(x_j) \right| = \left| \sum_{j=1}^N \alpha_j f(x_j) \right| = |\lambda(f)|$$

and

$$\|\lambda\|_{\Phi,\Omega_1} = \sum_{j,k=1}^N \alpha_j \alpha_k \Phi(x_j, x_k) = \|\lambda\|_{\Phi,\Omega_2}$$

we have

$$|f_{|\Omega_1}|_{\mathcal{N}_\Phi(\Omega_1)} = \sup_{\substack{\lambda \in L_\mathcal{P}(\Omega_1) \\ \lambda \neq 0}} \frac{|\lambda(f_{|\Omega_1})|}{\|\lambda\|_{\Phi,\Omega_1}} \leq \sup_{\substack{\lambda \in L_\mathcal{P}(\Omega_2) \\ \lambda \neq 0}} \frac{|\lambda(f)|}{\|\lambda\|_{\Phi,\Omega_2}} = |f|_{\mathcal{N}_\Phi(\Omega_2)}.$$

□

Theorem 2.21 *Suppose that $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a positive definite symmetric kernel with $\mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ and that $\|\cdot\|_{\mathcal{N}_\Phi(\mathbb{R}^d)}$ and $\|\cdot\|_{H^s(\mathbb{R}^d)}$ are equivalent. If $\Omega \subseteq \mathbb{R}^d$ is a bounded Lipschitz domain then $\mathcal{N}_\Phi(\Omega) = H^s(\Omega)$ with equivalent norms $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$, $\|\cdot\|_{H^s(\Omega)}$.*

Proof
If $f \in \mathcal{N}_\Phi(\Omega)$ then by Theorem 2.19 we have that there exists $E(f) \in \mathcal{N}_\Phi(\mathbb{R}^d) = H^s(\mathbb{R}^d)$ such that $f = E(f)_{|\Omega} \in H^s(\mathbb{R}^d)_{|\Omega} \subseteq H^s(\Omega)$. By recalling the property of the extension $E(f)$ given in Theorem 2.19 and the equivalence of $\|\cdot\|_{\mathcal{N}_\Phi(\mathbb{R}^d)}$ and $\|\cdot\|_{H^s(\mathbb{R}^d)}$ we obtain

$$\|f\|_{H^s(\Omega)} = \|E(f)_{|\Omega}\|_{H^s(\Omega)} \leq \|E(f)\|_{H^s(\mathbb{R}^d)} \leq c\|E(f)\|_{\mathcal{N}_\Phi(\mathbb{R}^d)} = c\|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

If $f \in H^s(\Omega)$ then by Theorem 2.18 there exists $E(f) \in H^s(\mathbb{R}^d) = \mathcal{N}_\Phi(\mathbb{R}^d)$ such that $f = E(f)_{|\Omega} \in \mathcal{N}_\Phi(\Omega)$ (Theorem 2.20). By remarking the property of the restriction of Theorem 2.20, the boundness of the extension operator of Sobolev spaces (Theorem 2.18) and the equivalence between $\|\cdot\|_{\mathcal{N}_\Phi(\mathbb{R}^d)}$ and $\|\cdot\|_{H^s(\mathbb{R}^d)}$ we can get

$$\|f\|_{\mathcal{N}_\Phi(\Omega)} = \|E(f)_{|\Omega}\|_{\mathcal{N}_\Phi(\Omega)} \leq \|E(f)\|_{\mathcal{N}_\Phi(\mathbb{R}^d)} \leq c\|E(f)\|_{H^s(\mathbb{R}^d)} \leq cC_s\|f\|_{H^s(\Omega)}.$$

□

If we are in the hypothesis of Theorem 1.26 by carefully analysing the proof of Theorem 2.21 we obtain that $\mathcal{N}_{\Phi_\delta}(\Omega) = H^s(\Omega)$ with

$$\|f\|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq (2\pi)^{-\frac{d}{4}} C_s c_1^{-\frac{1}{2}} \delta^{-\frac{d}{2}} \|f\|_{H^s(\Omega)} \tag{2.23}$$

and

$$\|f\|_{H^s(\Omega)} \leq (2\pi)^{\frac{d}{4}} c_2^{\frac{1}{2}} \delta^{\frac{d}{2}-s} \|f\|_{\mathcal{N}_{\Phi_\delta}(\Omega)}, \tag{2.24}$$

with C_s the norm of the extension operator in Theorem (2.18).

Chapter 3

Approximation methods with polynomial reproduction

3.1 Local polynomial reproduction

This section follows the presentation of [12, 18]. In general, if an approximation process recover polynomials at least locally then it inherits the local approximation order. To make this more clear we consider the univariate case: if f has k continuous derivatives around a point $x_0 \in \mathbb{R}$ then the Taylor polynomial

$$p(x) = \sum_{j=0}^{k-1} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j$$

has the following local approximation error for $|x - x_0| \leq h$

$$|f(x) - p(x)| = \frac{|f^{(k)}(\xi)|}{k!}|x - x_0|^k \leq Ch^k$$

with ξ between x and x_0.

We consider a set $X = \{x_1, \ldots, x_N\}$ of pairwise distinct points in $\Omega \subseteq \mathbb{R}^d$ and function values $f(x_1), \ldots, f(x_N)$. Our goal is to find and approximant s to the unknown function f.

We decide to study an approximant with the following form:

$$s(x) = \sum_{j=1}^{N} f(x_j)u_j(x),$$

where $u_j : \Omega \to \mathbb{R}$ for $j = 1, \ldots, N$. If the function $\{u_j\}_{j=1,\ldots,N}$ are cardinal with respect to X, i.e. $u_j(x_k) = \delta_{j,k}$ for $1 \leq j, k \leq N$ we have an interpolant. If the functions $\{u_j\}_{j=1,\ldots,N}$ are not cardinal then the approximant is called a *quasi-interpolant*.

For the error analysis of an approximation process it is important the fill distance, which is defined in this way: for a set of points $X = \{x_1, \dots, x_N\}$ in a bounded domain $\Omega \subseteq \mathbb{R}^d$ we call fill distance

$$h_{X,\Omega} = \sup_{x\in\Omega} \min_{1\leq j\leq N} \|x - x_j\|_2.$$

The fill distance measures how well a set of data sites covers the domain Ω. Each point $x \in \Omega$ has a distance from a data site not greater than $h_{X,\Omega}$. This means that each ball centered in a point of Ω of radius greater than $h_{X,\Omega}$ intersects X in at least one point. We can define more precisely what is a local reproduction property for an approximation process.

Definition 3.1 *A process that defines for every set $X = \{x_1, \dots, x_N\} \subseteq \Omega$ a family of functions $u_j = u_j^X : \Omega \to \mathbb{R}$ for $1 \leq j \leq N$ provides a local polynomial reproduction of degree ℓ on Ω if there exists constants h_0, C_1, C_2 such that*

- $\sum_{j=1}^N p(x_j)u_j = p$ *for each* $p \in \pi_\ell(\mathbb{R}^d)$,
- $\sum_{j=1}^N |u_j(x)| \leq C_1$ *for all* $x \in \Omega$,
- $u_j(x) = 0$ *if* $\|x - x_j\|_2 > C_2 h_{X,\Omega}$

are satisfied for all X with $h_{X_\Omega} \leq h_0$.

In the Definition 3.1 $\pi_\ell(\mathbb{R}^d)$ is the space of polynomials of degree at most ℓ in $\mathbb{R}^d$.

The most important observation is that the constants involved are independent of the data sites. The second condition guarantees that the process is stable, that is the Lebesgue constant is bounded. Indeed,

$$\begin{aligned}\left|\sum_{j=1}^N f(x_j)u_j(x) - \sum_{j=1}^N \widetilde{f}(x_j)u_j(x)\right| &\leq \sum_{j=1}^N |f(x_j) - \widetilde{f}(x_j)||u_j(x)| \leq \\ &\leq \max_{1\leq j\leq N} |f(x_j) - \widetilde{f}(x_j)| \sum_{j=1}^N |u_j(x)| \leq \\ &\leq \max_{1\leq j\leq N} |f(x_j) - \widetilde{f}(x_j)| C_1.\end{aligned}$$

Instead, the first and third conditions concern local polynomial reproduction. In this context the functions $\{u_j\}_{j=1,\dots,N}$ are arbitrary (it is not important the smoothness in our analysis).

Theorem 3.1 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is bounded. Define Ω^* to be the closure of $\bigcup_{x\in\Omega} B(x, C_2 h_0)$. Define $s_{f,X} = \sum_{j=1}^N f(x_j)u_j$, where $\{u_j\}_{j=1,\dots,N}$ are generated by a process with local polynomial reproduction of order m on Ω (Definition 3.1). If $f \in \mathcal{C}^{m+1}(\Omega^*)$ then there exists a constant $C = C(C_1, C_2, m)$ such that*

$$\|f - s_{f,X}\|_{L^\infty(\Omega)} \leq C h_{X,\Omega}^{m+1} |f|_{\mathcal{C}^{m+1}(\Omega^*)}$$

for all X with $h_{X,\Omega} \leq h_0$. The semi-norm in the inequality is defined by $|f|_{\mathcal{C}^{m+1}(\Omega^*)} := \max_{|\alpha|=m+1} \|D^\alpha f\|_{L^\infty(\Omega^*)}$.

Proof
Let p be an arbitrary polynomial from $\pi_m(\mathbb{R}^d)$. Using Definition 3.1 we obtain

$$\begin{aligned}
|f(x)-s_{f,X}(x)| &\leq |f(x)-p(x)| + \left| p(x) - \sum_{j=1}^{N} f(x_j)u_j(x) \right| \leq \\
&\leq |f(x)-p(x)| + \sum_{j=1}^{N} |p(x_j)-f(x_j)||u_j(x)| \leq \\
&\leq |f(x)-p(x)| + \sum_{j:\|x-x_j\|_2\leq C_2 h_{X,\Omega}} |p(x_j)-f(x_j)||u_j(x)| \leq \\
&\leq \|f-p\|_{L^\infty(B(x,C_2 h_{X,\Omega}))} \left(1 + \sum_{j:\|x-x_j\|_2\leq C_2 h_{X,\Omega}} |u_j(x)| \right) \leq \\
&\leq (1+C_1)\|f-p\|_{L^\infty(B(x,C_2 h_{X,\Omega}))}.
\end{aligned}$$

To end the proof we choose p to be the Taylor polynomial of f around x of order m. For $y \in B(x, C_2 h_{X,\Omega})$ there exists $\xi \in \Omega^*$ such that

$$f(y) - \sum_{|\alpha|\leq m} \frac{D^\alpha f(x)}{\alpha!}(y-x)^\alpha = \sum_{|\alpha|=m+1} \frac{D^\alpha f(\xi)}{\alpha!}(y-x)^\alpha.$$

Remarking that $|(y-x)^\alpha| = \prod_{i=1}^d |y_i - x_i|^{\alpha_i} \leq \prod_{i=1}^d \|y-x\|^{\alpha_i} \leq \|y-x\|^{|\alpha|}$ we can conclude

$$\|f-p\|_{L^\infty(B(x,C_2 h_{X,\Omega}))} \leq \sum_{|\alpha|=m+1} \frac{\|D^\alpha f\|_{L^\infty(\Omega^*)}}{\alpha!}(C_2 h_{X,\Omega})^{m+1} \leq C(C_2,m) h_{X,\Omega}^{m+1} |f|_{C^{m+1}(\Omega^*)}.$$

□

The approximation result given above is local: if f is less smooth in a subregion of Ω, i.e. it has only $\ell \leq m$ continuous derivatives then the approximant has $\mathcal{O}(h_{X,\Omega}^\ell)$ as global convergence rate.

Before introducing Moving least squares we briefly exploit the existence of approximation schemes with local polynomial reproduction. As we pointed out before, it is not important that the functions $\{u_j\}_{j=1,\dots,N}$ are smooth to get convergence. For example if $m = 0$ we can define $\{u_j\}_{j=1,\dots,N}$ in this way: for each $x \in \Omega$ choose j such that $\|x - x_j\|_2$ is minimal and define $u_j(x) = 1$ and $u_k(x) = 0$ for $k \neq j$. This approximation process satisfy Definition 3.1 for constants with arbitrary h_0 ($C_1 = C_2 = 1$).

For our purposes it is not restrictive to study domains satisfying an interior cone condition.

Definition 3.2 *A set $\Omega \subseteq \mathbb{R}^d$ is said to satisfy an interior cone condition if there exists an angle $\vartheta \in (0, \pi/2)$ and a radius $r > 0$ such that for every $x \in \Omega$ a unit vector $\xi(x)$ exists such that the cone*

$$C(x, \xi(x), \vartheta, r) := \{x + \lambda y : y \in \mathbb{R}^d,\ \|y\|_2 = 1,\ \langle y, \xi(x)\rangle \geq \cos(\vartheta),\ \lambda \in [0,r]\} \subseteq \Omega.$$

Theorem 3.2 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is compact and satisfies an interior cone condition with angle $\vartheta \in]0, \pi/2[$ and radius $r > 0$. Fix $m \in \mathbb{N}$. Then there exists constants h_0, C_1, C_2 depending on m, ϑ, r such that for every $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ with $h_{X,\Omega} \leq h_0$ and every $x \in \Omega$ we can find real numbers $\{\widetilde{u}_j(x)\}_{j=1,\ldots,N}$ with*

- $\sum_{j=1}^{N} p(x_j)\widetilde{u}_j(x) = p(x)$ *for each* $p \in \pi_\ell(\mathbb{R}^d)$,
- $\sum_{j=1}^{N} |\widetilde{u}_j(x)| \leq C_1$,
- $\widetilde{u}_j(x) = 0$ *if* $\|x - x_j\|_2 > C_2 h_{X,\Omega}$.

We can show explicit values for the constants involved (without restriction we can suppose $\vartheta \leq \pi/5$):

$$C_1 = 2, \qquad C_2 = \frac{16(1 + \sin(\vartheta))^2 m^2}{3\sin(\vartheta)^2}, \qquad h_0 = \frac{r}{C_2}.$$

3.2 Moving least squares

To introduce this concept we follow [19, 20, 21, 22]. The crucial point in local polynomial reproduction (Theorem 3.1) is the compact support of the basis functions $\{u_j\}_{j=1,\ldots,N}$. The diameters of the supports of the functions $\{u_j\}_{j=1,\ldots,N}$ are bounded by a constant proportional to $h_{X,\Omega}$. Compact support means that data points far away from the current point of interest $x \in \Omega$ have no influence on the function value at x. The moving least squares method provides an efficient method to build families of functions with local polynomial reproduction.

For our work is not restrictive to consider a domain Ω that respect the interior cone condition with angle ϑ and radius r (Definition 3.2) The idea of the moving least square approximation is to solve for every point $x \in \Omega$ a locally weighted least squares problem. The influence of the data points is governed by a weight function $w : \Omega \times \Omega \to \mathbb{R}$, which becomes smaller the greater is the norm of the difference between its arguments. For example w can vanish if computed at $x, y \in \Omega$ with $\|x - y\|_2$ greater than a threshold. We will use a weight function with the following form:

$$w(x, y) = \Phi_\delta(x - y) = \Phi\left(\frac{x - y}{\delta}\right),$$

where $\Phi : \mathbb{R}^d \to \mathbb{R}$ is a compactly supported function.

Definition 3.3 *For $x \in \Omega$, the value $s_{f,X}(x)$ of the moving least squares approximant is given by $s_{f,X}(x) = p^*(x)$ where p^* is the solution of*

$$\min\left\{\sum_{i=1}^{N}(f(x_i) - p(x_i))^2 w(x, x_i) : p \in \pi_m(\mathbb{R}^d)\right\}. \tag{3.1}$$

As anticipated we will use a weight function induced by a compactly supported function Φ, which is supported in the closed ball $\overline{B(0,1)}$ and strictly positive on $B(0,1)$. With this properties we can rewrite equation (3.1) as

$$\min\left\{\sum_{i\in I(x)}(f(x_i)-p(x_i))^2\Phi_\delta(x-x_i):p\in\pi_m(\mathbb{R}^d)\right\}, \tag{3.2}$$

where $I(x)=I(x,\delta,X)=\{j\in\{1,\dots,N\}:\|x-x_j\|_2<\delta\}$. So far it is not clear at all why moving least squares provides local polynomial reproduction: we need some results on quadratic optimization.

Theorem 3.3 *Let $a\in\mathbb{R}, b\in\mathbb{R}^n, A\in M_n(\mathbb{R})$ and $P\in M_{n,m}(\mathbb{R})$ be given. For $v\in\mathbb{R}^m$ define $M_v:=\{x\in\mathbb{R}^n:P^\top x=v\}$. Suppose $A=A^\top$ is positive definite on M_0. If M_v is not empty the the quadratic function*

$$f(x):=a+b^\top x+x^\top Ax$$

has a unique minimum on M_v.

Proof
We observe that M_v is defined by linear constraints, so it is convex and closed. We claim that if f is strictly convex on M_v then we obtain uniqueness of the solution: suppose that there exists $x^*\in M_v$ such that $f(x^*)\le f(x)$ for all $x\in M_v$. Fix $x\in M_v\setminus\{x^*\}$, then

$$f(x^*)\le f(\alpha x^*+(1-\alpha)x)<\alpha f(x^*)+(1-\alpha)f(x)\Rightarrow(1-\alpha)f(x^*)<(1-\alpha)f(x),$$

which implies uniqueness if $\alpha\in]0,1[$.

We claim that since A is positive definite on the vector subspace M_0 then there exists $\widetilde{\lambda}>0$ such that $x^\top Ax\ge\widetilde{\lambda}\|x\|_2^2$.

Suppose $M_0=\langle v_1,\dots,v_t\rangle\le\mathbb{R}^n$ is generated by an orthonormal basis. There exists

$$\begin{aligned}T:\mathbb{R}^t&\to M_0\\ e_i&\mapsto v_i\qquad\text{for each } i=1,\dots t,\end{aligned} \tag{3.3}$$

which is an isometry. Through T we can define a positive definite bilinear form on $\mathbb{R}^t$.

$$h(x,y)=x^\top By\qquad\text{for } x,y\in\mathbb{R}^t, \tag{3.4}$$

where $B\in M_t(\mathbb{R})$ with $B_{ij}=T(e_i)^\top AT(e_j)$ for $i,j=1,\dots,N$.

We say that h is a symmetric positive definite bilinear form because A is symmetric and positive definite on M_0.

$$x^\top Bx = \sum_{i=1}^t \sum_{j=1}^t x_i x_j B_{ij} = \sum_{i=1}^t \sum_{j=1}^t x_i x_j T(e_i)^\top AT(e_j) = \sum_{i=1}^t \sum_{j=1}^t T(x_i e_i)^\top AT(x_j e_j)$$

$$= \sum_{i=1}^t T(x_i e_i)^\top AT\left(\sum_{j=1}^t x_j e_j\right) = T\left(\sum_{i=1}^t x_i e_i\right)^\top AT\left(\sum_{j=1}^t x_j e_j\right) = \tag{3.5}$$

$$= T(x)^\top AT(x) > 0.$$

With equation (3.5) and taking an orthonormal basis of $\mathbb{R}^t$ with respect to h we can write that

$$T(x)^\top AT(x) = x^\top Bx \geq \widetilde{\lambda}\|x\|_2^2 = \widetilde{\lambda}\|T(x)\|_2^2,$$

where $\widetilde{\lambda} > 0$ is the smallest eigenvalue of B (the last equality holds because T is an isometry).

A straightforward computation gives us, for different $x, y \in M_v$ and $\lambda \in]0,1[$

$$\begin{aligned}
&(1-\lambda)f(x) + \lambda f(y) - f((1-\lambda)x + \lambda y) = \\
=&(1-\lambda)x^\top Ax + \lambda y^\top Ay - (1-\lambda)^2 x^\top Ax - \lambda^2 y^\top Ay - 2\lambda(1-\lambda)x^\top Ay = \\
=&\lambda(1-\lambda)(x^\top Ax - 2x^\top Ay + y^\top Ay) = \lambda(1-\lambda)(x-y)^\top A(x-y) \geq \\
\geq&\lambda(1-\lambda)\widetilde{\lambda}\|x-y\|_2^2 > 0,
\end{aligned}$$

which implies that f is strictly convex on M_v.

To show the existence of the minimum we will use Weierstrass Theorem. Since M_v is not empty we can fix an $x_0 \in M_v$ and restric the minimization problem to $\widetilde{M_v} := \{x \in M_v : f(x) \leq f(x_0)\}$. From the continuity of f we obtain that $\widetilde{M_v}$ is closed. From the definition of M_v and M_0 we get that every element $x \in \widetilde{M_v}$ has the form $x = x_0 + z$, with $z \in M_0$.

$$\begin{aligned}
0 \geq f(x) - f(x_0) &= b^\top x + x^\top Ax - b^\top x_0 - x_0^\top Ax_0 = \\
&= (b + 2Ax_0)^\top (x - x_0) + (x - x_0)^\top A(x - x_0) \geq \\
&\geq -\|b + 2Ax_0\|_2 \|x - x_0\|_2 + \widetilde{\lambda}\|x - x_0\|_2^2.
\end{aligned}$$

explains that $\|x\|_2 \leq \|x - x_0\|_2 + \|x_0\|_2$ is bounded for each $x \in \widetilde{M_v}$. There exists a minimun $x^* \in \widetilde{M_v}$ for f in $\widetilde{M_v}$. The identified point is also a minimum for f in M_v.

□

To make the following statement more clear we introduce a definition.

Definition 3.4 *The points $X = \{x_1, \ldots, x_N\} \subseteq \mathbb{R}^d$ with $N \geq Q = dim(\pi_m(\mathbb{R}^d))$ are called $\pi_m(\mathbb{R}^d)$-unisolvent if the zero polynomial is the only polynomial from $\pi_m(\mathbb{R}^d)$ that vanishes on X.*

Theorem 3.4 *Suppose that for every $x \in \Omega$ the set $\{x_j : j \in I(x)\}$ is $\pi_m(\mathbb{R}^d)$-unisolvent. In this situation, the problem stated in equation (3.2) is uniquely solvable and the solution $s_{f,X}(x) = p^*(x)$ can be represented as*

$$s_{f,X}(x) = \sum_{i \in I(x)} f(x_i) a_i^*(x),$$

where the coefficients $a_i^*(x)$ are determined by minimizing the quadratic form

$$\sum_{i \in I(x)} \frac{a_i(x)^2}{\Phi_\delta(x - x_i)} \tag{3.6}$$

under the constraints

$$\sum_{i \in I(x)} p(x_i) a_i(x) = p(x), \qquad p \in \pi_m(\mathbb{R}^d). \tag{3.7}$$

Proof
Denote a basis of $\pi_m(\mathbb{R}^d)$ by $p_1, \ldots, p_Q$. Suppose our polynomial has the form $p = \sum_{j=1}^{Q} b_j p_j$. This rewrite the minimization problem in the equation (3.2) to find the optimal coefficient vector b^*. We will use the following notation:

$$\begin{aligned}
b &= (b_1, \ldots, b_Q)^\top \in \mathbb{R}^Q, \\
\widetilde{f} &= (f(x_i) : i \in I(x))^\top \in \mathbb{R}^{\# I(x)}, \\
P &= (p_j(x_i))_{\substack{i \in I(x) \\ 1 \le j \le Q}} \in M_{\# I(x), Q}(\mathbb{R}), \\
D &= D(x) = \operatorname{diag}(\Phi_\delta(x - x_i) : i \in I(x)) \in M_{\# I(x), \# I(x)}(\mathbb{R}), \\
R(x) &= (p_1(x), \ldots, p_Q(x))^\top \in \mathbb{R}^Q.
\end{aligned}$$

With this notation we can rewrite the function to minimized in equation (3.2) as

$$\begin{aligned}
C(b) = \sum_{i \in I(x)} &\left(f(x_i) - \sum_{j=1}^{Q} b_j p_j(x_i) \right)^2 \Phi_\delta(x - x_i) = \\
= \sum_{i \in I(x)} &(\widetilde{f}_i - (Pb)_i)^2 \Phi_\delta(x - x_i) = \\
=&(\widetilde{f} - Pb)^\top D (\widetilde{f} - Pb) = \\
=&\widetilde{f}^\top D \widetilde{f} - 2 \widetilde{f}^\top D P b + b^\top P^\top D P b.
\end{aligned} \tag{3.8}$$

Since $C(b)$ is a quadratic form in b we can apply Theorem 3.3 if $P^\top D P$ is positive definite on $\mathbb{R}^Q$. For clarity of exposition, the instance of Theorem 3.3, that we use, has the matrix that defines M_v and M_0 equals to 0.

If we analyse

$$b^\top P^\top DPb = b^\top P^\top D^{\frac{1}{2}} D^{\frac{1}{2}} Pb = \|D^{\frac{1}{2}} Pb\|_2^2 \tag{3.9}$$

and suppose that $b^\top P^\top DPb = 0$, we get that $\sum_{i\in I(x)} \Phi_\delta(x - x_i)(Pb)_i^2 = 0 \Rightarrow Pb = 0$ since $\Phi_\delta(x - x_i) > 0$ for each $i \in I(x)$. In this situation the polynomial $p = \sum_{j=1}^{Q} b_j p_j$ vanishes on $\{x_i : i \in I(x)\}$, which is unisolvent (Definition 3.4). So b has to be 0.

Since we are minimizing a differentiable function on a open set with a unique minimun we determine $b^* \in \mathbb{R}^d$ such that $\nabla C(b^*) = 0$.

$$0 = \nabla C(b^*) = (-2\widetilde{f}^\top DP)^\top + 2P^\top DPb^* \Rightarrow (b^*)^\top = \widetilde{f}^\top DP(P^\top DP)^{-1},$$

from this we obtain the solution to the minimization problem in equation (3.2):

$$p^*(x) = (b^*)^\top R(x) = \widetilde{f}^\top DP(P^\top DP)^{-1} R(x).$$

To conclude the proof of the statement we study the minimization problem in equation (3.6) with constraints (3.7). In other words we minimize the function

$$C(a) = \sum_{i\in I(x)} \frac{a_i^2}{\Phi_\delta(x - x_i)} = a^\top D^{-1} a$$

subjects to

$$\begin{aligned} M := &= \left\{ a \in \mathbb{R}^{\#I(x)} : \sum_{i\in I(x)} a_i p(x_i) = p(x) \text{ for } p \in \pi_m(\mathbb{R}^d) \right\} = \\ &= \left\{ a \in \mathbb{R}^{\#I(x)} : \sum_{i\in I(x)} a_i p_j(x_i) = p_j(x), 1 \le j \le Q \right\} = \\ &= \{a \in \mathbb{R}^{\#I(x)} : P^\top a = R(x)\}. \end{aligned}$$

We supposed that $\{x_i : i \in I(x)\}$ is $\pi_m(\mathbb{R}^d)$-unisolvent so the matrix $P \in M_{\#I(x),Q}(\mathbb{R})$ is injective with full column rank. This proves that $P^\top$ is surjective and $M \neq \emptyset$. Since D^{-1} is positive definite we can obtain a unique solution of the problem in equation (3.6) subjects to the constraints described in equation (3.7) with Theorem 3.3.

To find the solution of a constrained optimization problem we use Lagrange multipliers.

Consider the Lagrange function defined by

$$\mathcal{L}(a, \lambda) = a^\top D^{-1} a - \lambda^\top (P^\top a - R(x))$$

and impose

$$\nabla_a \mathcal{L}(a^*, \lambda) = 0 \Rightarrow 2D^{-1}a^* - P\lambda = 0 \Rightarrow 2D^{-1}a^* = P\lambda \Rightarrow a^* = DP\lambda/2. \tag{3.10}$$

Since $a^* \in M$ we get $R(x) = P^\top a^* = P^\top DP\lambda/2$. From the positive definiteness of $P^\top DP$ (equation (3.9)) we obtain an explicit expression for λ

$$\lambda = 2(P^\top DP)^{-1}R(x). \tag{3.11}$$

In the end we can write

$$\sum_{i\in I(x)} a_i^* f(x_i) = \widetilde{f}^\top a^* = \widetilde{f}^\top DP(P^\top DP)^{-1}R(x) = p^*(x).$$

□

We notice that for Theorem 3.1 the smoothness of the functions $\{u_j\}_{j=1,\dots,N}$ is not involved in the proof. For some practical application it is necessary to have available a smooth approximant. The proof of Theorem 3.4 let us to characterize the functions $\{a_j^*\}_{j=1,\dots,N}$.

Theorem 3.5 *The functions $\{a_j^*\}_{j=1,\dots,N}$ of Theorem 3.4 have the form*

$$a_j^*(x) = \Phi_\delta(x - x_j)\sum_{k=1}^{Q} \lambda_k(x)p_k(x_j)$$

where $\{\lambda_k(x)\}_{k=1,\dots,Q}$ are the unique solutions of

$$\sum_{k=1}^{Q} \lambda_k(x) \sum_{j\in I(x)} \Phi_\delta(x - x_j)p_k(x_j)p_\ell(x_j) = p_\ell(x), \qquad 0 \le \ell \le Q. \tag{3.12}$$

Moreover, if $\Phi \in \mathcal{C}^k(\Omega)$ then the approximant $s_{f,X} \in \mathcal{C}^k(\Omega)$.

Proof
Let us fix $x \in \Omega$ and $j \in \{1,\dots,N\}$ such that $\|x - x_j\|_2 < \delta$. With notation of Theorem 3.4 and equation (3.10),

$$a^*(x) = D(x)P\lambda/2,$$

becomes, since $D(x)$ is diagonal,

$$a_j^*(x) = \Phi_\delta(x - x_j)(P\lambda/2)_j = \Phi_\delta(x - x_j)\sum_{k=1}^{Q} \lambda_k(x)p_k(x_j).$$

For sake of clarity we have identified $\frac{\lambda}{2} \sim \lambda$.

We can end the proof by recalling equation (3.11), which gives us

$$P^\top DP\lambda = R(x).$$

Making the matrices involved explicit we obtain

$$DP = \begin{pmatrix} \Phi_\delta(x - x_{w_1}) & & \\ & \ddots & \\ & & \Phi_\delta(x - x_{w_{\#I(x)}}) \end{pmatrix} \begin{pmatrix} p_1(x_{w_1}) & \cdots & p_Q(x_{w_1}) \\ \vdots & & \vdots \\ p_1(x_{w_{\#I(x)}}) & \cdots & p_Q(x_{w_{\#I(x)}}) \end{pmatrix} =$$

$$= \begin{pmatrix} \Phi_\delta(x - x_{w_1})p_1(x_{w_1}) & \cdots & \Phi_\delta(x - x_{w_1})p_Q(x_{w_1}) \\ \vdots & & \vdots \\ \Phi_\delta(x - x_{w_{\#I(x)}})p_1(x_{w_{\#I(x)}}) & \cdots & \Phi_\delta(x - x_{w_{\#I(x)}})p_Q(x_{w_{\#I(x)}}) \end{pmatrix} =$$

$$= \big(\Phi_\delta(x - x_i)p_j(x_i)\big)_{i \in I(x), j=1,\dots,Q}.$$

From this computation, we can write for $\ell \in \{1, \dots, Q\}$

$$\begin{aligned} p_\ell(x) &= \sum_{k=1}^{Q} \lambda_k(x)(P^\top DP)_{\ell,k} = \\ &= \sum_{k=1}^{Q} \lambda_k(x) \sum_{j \in I(x)} p_\ell(x_j)\Phi_\delta(x - x_j)p_k(x_j). \end{aligned}$$

We just proved that the smoothness of $s_{f,X}$ depends on the smoothness of Φ and the smoothness of $\{\lambda_j\}_{j=1,\dots,Q}$. Since Φ is supported in $\overline{B(0,1)}$ equation (3.12) becomes

$$\sum_{k=1}^{Q} \lambda_k(x) \sum_{j=1}^{N} \Phi_\delta(x - x_j)p_k(x_j)p_\ell(x_j) = p_\ell(x), \qquad 0 \le \ell \le Q.$$

That is, with an extended version of the notation of Theorem 3.4,

$$P^\top DP\lambda(x) = R(x),$$

where $P \in M_{N,Q}(\mathbb{R})$ and $D = D(x) \in M_N(\mathbb{R})$. The matrix $P^\top DP$ remains positive definite and invertible (compare equation (3.9), only some components of Pb are zero, but this is sufficient to guarantee that $b = 0$ because of the unisolvent property of $\{x_j : j \in I(x)\}$). $\lambda(\cdot) \in \mathbb{R}^Q$ is continuous on Ω because each components is a continuous function of the components of $P^\top D(x)P$ and $R(x)$ divided by $\det(P^\top DP)$. We remark that P is independent of $x \in \Omega$.

□

With this results we can compute the moving least square approximation method for $m = 0$. If we choose $\{1\}$ as basis of $\pi_0(\mathbb{R}^d)$ equation (3.11) gives us

$$\frac{\lambda}{2} = (P^\top DP)^{-1}R(x) = (P^\top DP)^{-1} = \frac{1}{\sum_{i \in I(x)} \Phi_\delta(x - x_i)},$$

because $R(x) = 1$ and $P = (1, \dots, 1)^\top \in M_{\#I(x),1}$. Theorem 3.5 let us to conclude

$$a_j^*(x) = \frac{\Phi_\delta(x - x_j)}{\sum_{i \in I(x)} \Phi_\delta(x - x_i)} \qquad \text{for each } j \in \{1, \dots, N\}.$$

For $m = 0$ the minimization problem described in equation (3.2) has the solution

$$s_{f,X}(x) = \sum_{j=1}^{N} f(x_j) \underbrace{\frac{\Phi_\delta(x - x_j)}{\sum_{i=1}^{N} \Phi_\delta(x - x_i)}}_{a_j^*(x)},$$

which is a particular instance of the Shepard approximation method [22] and it reproduces constants exactly. If we choose δ to be $Ch_{X,\Omega}$ with $C > 1$, then the moving least squares approximant with $m = 0$ respect the Definition 3.1 with $\ell = 0, C_1 = 1$ and $C_2 = C$. The method is well-defined because the denominantor is different from 0. For each $x \in \Omega$ there exists $k \in \{1, \dots, N\}$ such that $\|x - x_k\|_2$ is minimal, which provides $\|x - x_k\|_2 \leq h_{X,\Omega} < \delta$. This proves that $x_k \in I(x) \neq \emptyset$ and

$$\sum_{i \in I(x)} \Phi_\delta(x - x_i) > 0$$

With Theorem 3.1 we can prove the existence of a stable quasi-interpolant method with convergence rate $\mathcal{O}(h_{X,\Omega})$. The mentioned method can provide a smooth interpolant as smooth as Φ.

Before continuing, Theorem 3.5 let us to analyse the computational complexity of the method for the evaluation of the approximant at a point $x \in \Omega$. To compute $\{\lambda_1(x), \dots, \lambda_Q(x)\}$ we need to solve a $Q \times Q$ linear system, so the computational cost is $\mathcal{O}(Q^3)$. The cost to build the matrix of the linear system is $\mathcal{O}(\#I(x)Q^2)$ because for each component of the matrix we perform a number of multiplications and sums proportional to $\#I(x)$ (we assume that the cost for polynomial evaluation is constant). Since the basis functions $\{a_1^*, \dots, a_N^*\}$ are compactly supported we need to calculate only $\#I(x)$ of them, so since we perform a number of multiplications and sums proportional to Q the cost is $\mathcal{O}(\#I(x)Q)$. After all to build the basis functions at a point $x \in \Omega$ the computational cost is

$$\mathcal{O}(Q^3 + \#I(x)Q^2 + \#I(x)Q)$$

and we have to add $\mathcal{O}(\#I(x))$ to compute the value of the approximant. Since we are working in a quasi-uniform setting Theorem 7.3 gives as a uniform upper bound for $\#I(x)$ if δ is proportional to $h_{X,\Omega}$. Under these constraints the cost for the evaluation of the approximant at a point $x \in \Omega$ is constant. For each $x \in \Omega$ we can build $\#I(x)$ in $\mathcal{O}(N)$ comparisons, where N is the number of data sites. If we have to compute the value of the approximant at M points a naive implementation of the algorithm leads to a computational cost of $\mathcal{O}(NM)$. A more refined stratey, that consists to divide the domain in boxes of side length h and to consider for each $x \in \Omega$ only the relevant ones, provides a final computational cost of $\mathcal{O}(N + M)$.

In Theorem 3.5 appears a basis for the polynomial space and, even if it is not relevant from a theoretical point of view, a careful choice of the basis can lead to a more stable implementation.

Now we want to prove that the moving least squares approximation method provides local polynomial reproduction with $\ell = m$ (Definition 3.1). We will need some new definitions.

Definition 3.5 *The separation distance of $X = \{x_1, \ldots, x_N\}$ is defined by*

$$q_X := \frac{1}{2} \min_{i \neq j} \|x_i - x_j\|_2.$$

We note that if $i \neq j$ then $B(x_i, q_X) \cap B(x_j, q_X) = \emptyset$. If by contradiction we suppose that $B(x_i, q_X) \cap B(x_j, q_X) \neq \emptyset$ then there exists $z \in B(x_i, q_X) \cap B(x_j, q_X)$, which gives us:

$$2q_X \leq \|x_i - x_j\|_2 \leq \|x_i - z\|_2 + \|z - x_j\|_2 < 2q_X.$$

It is also important to note that q_X is the largest possible radius with the property above.

Definition 3.6 *A set $X = \{x_1, \ldots, x_N\}$ is said to be quasi-uniform with respect to a constant $c_{qu} > 0$ if*

$$q_X \leq h_{X,\Omega} \leq c_{qu} q_X.$$

To require $q_X \leq h_{X,\Omega}$ instead of $cq_X \leq h_{X,\Omega}$ with $c < 1$ is not a restriction if Ω satisfies an interior cone condition with radius $r > 0$ and angle $\vartheta > 0$ (Definition 3.2). If $q_X \leq r$ then for each $x_i \in X$ we can find $y \in \Omega$ with $\|y - x_i\|_2 = q_X$, because if $\|z\|_2 = 1$ and $\langle z, \xi(x) \rangle \geq \cos(\vartheta)$ then

$$\|x_i + \lambda z - x_i\|_2 = \lambda \in [0, r]$$

and $x_i + \lambda z \in \Omega$. For any other $x_j \in \Omega$

$$\|y - x_j\|_2 \geq \|x_j - x_i\|_2 - \|y - x_i\|_2 \geq 2q_X - q_X = q_X,$$

that gives $h_{X,\Omega} \geq q_X$.

The Definition 3.6 is useful when we consider a sequence of data sets that have the same constant c_{qu} and the fill distance becomes smaller and smaller.

For quasi-uniform sets we can bound from above and below q_X and $h_{X,\Omega}$ with functions of $\frac{1}{N}$, where $N = \#X$.

Proposition 3.6 *Let $\Omega \subseteq \mathbb{R}^d$ be bounded and measurable. Suppose that $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is quasi-uniform with respect to $c_{qu} > 0$. Then there exists constants $c_1, c_2, c_3, c_4 > 0$ depending only on the space dimension d, on Ω and on c_{qu} such that*

$$c_1 \left(\frac{1}{N}\right)^{\frac{1}{d}} \leq h_{X,\Omega} \leq c_2 \left(\frac{1}{N}\right)^{\frac{1}{d}} \quad \textit{and} \quad c_3 \left(\frac{1}{N}\right)^{\frac{1}{d}} \leq q_X \leq c_4 \left(\frac{1}{N}\right)^{\frac{1}{d}}$$

Proof
From the definition of $h_{X,\Omega}$ we have

$$\Omega \subseteq \bigcup_{j=1}^{N} \overline{B(x_j, h_{X,\Omega})}.$$

By the monotonicity and the subadditivity of the Lebesgue measure we have

$$\mathcal{L}(\Omega) \leq N h_{X,\Omega}^d \mathcal{L}(B(0,1)),$$

which gives $c_1 = \left(\frac{\mathcal{L}(\Omega)}{\mathcal{L}(B(0,1))}\right)^{\frac{1}{d}}$. This proves the first inequality.

Since Ω is bounded there exists a ball $B(x_0, R)$ such that $\Omega \subseteq B(x_0, R)$. From triangular inequality we obtain

$$\bigcup_{j=1}^{N} B(x_j, q_X) \subseteq B(x_0, R + q_X),$$

because if $y \in B(x_j, q_X)$ then $\|y - x_0\|_2 \leq \|y - x_j\|_2 + \|x_j - x_0\|_2 < q_X + R$.

We claim that $q_X \leq R$ because if $x_i, x_j \in X$ are distinct then

$$2q_X \leq \|x_i - x_j\|_2 \leq \|x_i - x_0\|_2 + \|x_0 - x_j\|_2 \leq 2R.$$

Since the balls $\{B(x_j, q_X)\}_{j=1,\dots,N}$ are pairwise disjoint the additivity of the Lebesgue measure on disjoint sets gives us

$$N q_X^d \mathcal{L}(B(0,1)) \leq (R + q_X)^d \mathcal{L}(B(0,1)) \Rightarrow N q_X^d \leq (R + q_X)^d \Rightarrow$$
$$\Rightarrow N \leq \left(1 + \frac{R}{q_X}\right)^d \leq (2R)^d \left(\frac{1}{q_X}\right)^d.$$

The last implication is true because $1 \leq \frac{R}{q_X}$ and it gives $c_4 = 2R$. From the quasi-uniformity of X (Definition 3.6) we obtain

$$N \leq (2Rc_{qu})^d \left(\frac{1}{h_{X,\Omega}}\right)^d$$

and $c_2 = 2Rc_{qu}$. The first inequality for q_X follows from the first inequality for $h_{X,\Omega}$ and it produces $c_3 = \frac{c_1}{c_{qu}}$.

□

A careful analysis of the proof of Proposition 3.6 highlights that to prove the first inequality for $h_{X,\Omega}$ and the second inequality for q_X we can consider arbitrary $X \subseteq \Omega$. It is not necessary that X is a quasi-uniform date set.

Let us return to the local reproduction property of moving least squares method.

Theorem 3.7 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is compact a satisfies an interior cone condition with angle $\vartheta \in]0, \pi/2[$ and radius $r > 0$. Fix $m \in \mathbb{N}$. Let h_0, C_1 and C_2 denote the constants of Theorem 3.2. Suppose that $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is a quasi-uniform data sets with respect to $c_{qu} > 0$ and $h_{X,\Omega} \leq h_0$. Let $\delta = 2C_2 h_{X,\Omega}$. Then the basis functions $\{a_j^*(x)\}_{j=1,\dots,N}$ of Theorem 3.5 provide local polynomial reproduction, i.e.*

- $\sum_{j=1}^N p(x_j) a_j^*(x) = p(x)$ *for all* $p \in \pi_m(\mathbb{R}^d), x \in \Omega$,
- $\sum_{j=1}^N |a_j^*(x)| \leq \widetilde{C}_1$,
- $a_j^*(x) = 0$ *if* $\|x - x_j\|_2 > \widetilde{C}_2 h_{X,\Omega}$,

with certain constants $\widetilde{C}_1, \widetilde{C}_2$ that can be derived explicitly.

Proof
The first property is a consequence of equation (3.2) and Theorem 3.4 that define the moving least squares method. We proved in Theorem 3.5 the a_j^* is supported in $\overline{B(x_j, \delta)}$ for $j = 1, \dots, N$. With our choice of δ the third property holds for $\widetilde{C}_2 = 2C_2$. To prove the second property we bound two quantities separately.

$$\sum_{i=1}^N |a_i^*(x)| = \sum_{i \in I(x)} |a_i^*(x)| = \sum_{i \in I(x)} \frac{|a_i^*(x)|}{\sqrt{\Phi_\delta(x - x_i)}} \sqrt{\Phi_\delta(x - x_i)} \leq$$
$$\leq \left(\sum_{i \in I(x)} \frac{|a_i^*(x)|^2}{\Phi_\delta(x - x_i)} \right)^{\frac{1}{2}} \left(\sum_{i \in I(x)} \Phi_\delta(x - x_i) \right)^{\frac{1}{2}}.$$

There exists $\{\widetilde{u}_j(x)\}_{j=1,\dots,N}$ providing local polynomial reproduction (Theorem 3.2):

$$p(x) = \sum_{j=1}^N p(x_j) \widetilde{u}_j(x) = \sum_{j : \|x - x_j\| \leq C_2 h_{X,\Omega}} p(x_j) \widetilde{u}_j(x) = \sum_{j \in I(x)} p(x_j) \widetilde{u}_j(x).$$

The minimal property stated in Theorem 3.4 gives

$$\sum_{i\in I(x)} \frac{|a_i^*(x)|^2}{\Phi_\delta(x-x_i)} \le \sum_{i\in I(x)} \frac{|\widetilde{u}_i(x)|^2}{\Phi_\delta(x-x_i)} = \sum_{i\in \widetilde{I}(x)} \frac{|\widetilde{u}_i(x)|^2}{\Phi_\delta(x-x_i)} \le$$
$$\le \frac{1}{\min_{y\in\overline{B(0,1/2)}}\Phi(y)} \sum_{i\in\widetilde{I}(x)} |\widetilde{u}_i(x)|^2 \le$$
$$\le \frac{1}{\min_{y\in\overline{B(0,1/2)}}\Phi(y)} \left(\sum_{i\in\widetilde{I}(x)} |\widetilde{u}_i(x)|\right)^2 \le$$
$$\le \frac{C_1^2}{\min_{y\in\overline{B(0,1/2)}}\Phi(y)},$$

where $\widetilde{I}(x) = \left\{j \in \{1,\ldots,N\} : x_j \in \overline{B(x,\frac{\delta}{2})}\right\}$. For a better understanding of the inequalities, it is useful to remember that $\widetilde{u}_j(x) = 0$ if $\|x - x_j\| > C_2 h_{X,\Omega} = \frac{\delta}{2}$ and that $(\cdot)^2$ is superadditive in $\mathbb{R}_{\ge 0}$.

We can bound the second factor with a volume comparison.

$$\sum_{i\in I(x)} \Phi_\delta(x-x_i) \le \#I(x)\|\Phi\|_{L^\infty(\mathbb{R}^d)}.$$

We have that

$$\bigcup_{i\in I(x)} B(x_i, q_X) \subseteq B(x, q_X+\delta)$$

because if $y \in B(x_i, q_X)$ for $i \in I(x)$ then $\|y-x\|_2 \le \|y-x_i\|_2 + \|x_i - x\|_2 < q_X + \delta$.

The additivity of Lebesgue measure on disjoint sets let us to write

$$\#I(x)\mathcal{L}(B(0,1))q_X^d \le \mathcal{L}(B(0,1))(q_X+\delta)^d$$

because $\{B(x_i,q_X)\}_{i\in I(x)}$ are disjoint. Since X is quasi-uniform with respect to c_{qu} we can conlude that

$$\#I(x) \le \left(1+\frac{\delta}{q_X}\right)^d = \left(1+\frac{2C_2h_{X,\Omega}}{q_X}\right)^d \le (1+2C_2c_{qu})^d.$$

□

Under the hypothesis of Theorem 3.7 we can obtain an error estimate for the moving least squares method with Theorem 3.1.

Chapter 4

Error estimates for RBF interpolation

The most significant results of this section can be found in [8, 9, 10, 11, 27, 32, 33].

4.1 Power function

Our goal in this section is to study the difference between the derivatives of a function $f \in \mathcal{N}_\Phi(\Omega)$ and the derivatives of its interpolant $s_{f,X}$, built with a (conditionally) positive definite symmetric kernel Φ (equation (2.1)).

The first step in many error estimates techniques is to write the interpolant $s_{f,X}$ of f in its Lagrangian form.

Choose $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ to be a $\mathcal{P}$-unisolvent set. Let us fix some notation: $A_{\Phi,X} = (\Phi(x_i, x_j))_{i,j=1,\ldots,N} \in M_N(\mathbb{R})$ and $P = (p_j(x_i))_{i=1,\ldots,N, j=1,\ldots,Q} \in M_{N,Q}(\mathbb{R})$, where $\{p_1, \ldots, p_Q\}$ is a basis of $\mathcal{P}$. We will denote $R(x) = (\Phi(x, x_1), \ldots, \Phi(x, x_N))^\top \in \mathbb{R}^N$ and $S(x) = (p_1(x), \ldots, p_Q(x))^\top \in \mathbb{R}^Q$.

If $(\alpha^{(j)}, \beta^{(j)}) \in \mathbb{R}^{N+Q}$ solves the system

$$\widetilde{A}_{\Phi,X} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} A_{\Phi,X} & P \\ P^\top & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} e^{(j)} \\ 0 \end{pmatrix} \tag{4.1}$$

where $e^{(j)}$ is the j-th unit vector, then

$$u_j^*(x) = \sum_{i=1}^{N} \alpha_i^{(j)} \Phi(x, x_i) + \sum_{k=1}^{Q} \beta_k^{(j)} p_k(x) \tag{4.2}$$

satisfies $u_j^*(x_i) = \delta_{i,j}$ for $i, j = 1, \ldots, N$.

We claim that, for each $j = 1, \dots, N$

$$u_j^\star \in (\langle \Phi(\cdot, x_j) : x_j \in X \rangle \cap F_\Phi(\Omega)) + \mathcal{P} = V_X. \tag{4.3}$$

From Theorem 2.1 we know that every $f \in V_X$ is uniquely determined by its values on X so

$$f = \sum_{j=1}^{N} f(x_j) u_j^* \tag{4.4}$$

and for $f \in \mathcal{N}_\Phi(\Omega)$ we have

$$s_{f,X} = \sum_{j=1}^{N} \underbrace{f(x_j)}_{s_{f,X}(x_j)} u_j^*. \tag{4.5}$$

The values of the cardinal functions $\{u_j^*\}_{j=1,\dots,N}$ at $x \in \Omega$ can be computed simultaneously by solving a linear system.

Theorem 4.1 *Suppose that Φ is a conditionally positive definite kernel with respect to $\mathcal{P}$ on $\Omega \subseteq \mathbb{R}^d$. Suppose that $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is $\mathcal{P}$-unisolvent. Then there exist functions $\{u_1^*, \dots, u_N^*\} \subseteq V_X$ such that $u_j^*(x_i) = \delta_{i,j}$ for $i, j = 1, \dots, N$. Moreover, for each $x \in \Omega$, there exists functions $\{v_1^*, \dots, v_Q^*\}$ such that*

$$\begin{pmatrix} A_{\Phi,X} & P \\ P^\top & 0 \end{pmatrix} \begin{pmatrix} u^*(x) \\ v^*(x) \end{pmatrix} = \begin{pmatrix} R(x) \\ S(x) \end{pmatrix}.$$

Proof
For brevity we will use A in place of $A_{\Phi,X}$. The functions $\{u_1^*, \dots, u_N^*\} \subseteq V_X$ are the cardinal functions with respect to X. Since $\mathcal{P} \subseteq V_X$ from equation (4.4) we obtain $p = \sum_{j=1}^{N} p(x_j) u_j^*$ for each $p \in \mathcal{P}$, that can be written as $P^\top u^*(x) = S(x)$.

If $P = (P_1 \cdots P_Q)$, where $P_i \in \mathbb{R}^N$ for $i = 1, \dots, Q$, then

$$P(\mathbb{R}^Q) = \left\{ \sum_{k=1}^{Q} v_k P_k : v_k \in \mathbb{R} \text{ for } k = 1, \dots, Q \right\} \subseteq \mathbb{R}^N$$

and

$$P(\mathbb{R}^Q)^\perp = \{\gamma \in \mathbb{R}^N : P^\top \gamma = 0\}.$$

We can conclude the proof of the theorem if we can prove that $Au^*(x) - R(x) \in P(\mathbb{R}^Q)$, because in this case there exists $v^*(x) \in \mathbb{R}^Q$ such that

$$Au^*(x) - R(x) = \sum_{k=1}^{Q} v_k^*(x) P_k.$$

To achieve this we will prove that $Au^*(x) - R(x) \in (P(\mathbb{R}^Q)^\perp)^\perp = P(\mathbb{R}^Q)$. If we fix $\gamma \in \mathbb{R}^N$ such that $P^\top \gamma = 0$ the $\gamma^\top R(x) \in V_X$. So, by equation (4.4),

$$\gamma^\top R(x) = \sum_{j=1}^{N} u_j^*(x)\gamma^\top R(x_j) = \sum_{j=1}^{N} u_j^*(x) \sum_{i=1}^{N} \gamma_i \underbrace{\Phi(x_j, x_i)}_{\Phi(x_i,x_j)} = \gamma^\top A u^*(x).$$

We have that $\gamma^\top(Au^*(x) - R(x)) = 0$ for each $\gamma \in P(\mathbb{R}^Q)^\perp$, thus $Au^*(x) - R(x) \in P(\mathbb{R}^Q)$.

□

From Theorem 2.1 we remark that the solution of the system in Theorem 4.1 is unique. So if we write the i-th equation of the system in Theorem 4.1 with $x = x_i$ we obtain

$$\sum_{j=1}^{N} \Phi(x_i, x_j) u_j^*(x_i) + \sum_{k=1}^{Q} p_k(x_i) v_k^*(x_i) = \Phi(x_i, x_i),$$

which gives us $u_j^*(x_i) = \delta_{i,j}$ for $j = 1, \dots, N$ and $v_k^*(x_i) = 0$ for $k = 1, \dots, Q$.

With equation (4.5) we have that the smoothness of $s_{f,X}$ depends on the smoothness of the cardinal functions $\{u_1^*, \dots, u_N^*\}$, that inherit the smoothness of Φ with respect to the first argument and that of $\mathcal{P}$. So $s_{f,X} \in \mathcal{C}^k(\Omega)$ if $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$ and Φ with respect to the first arguments admits k continuous derivatives.

From Theorem 4.1 we obtain

$$A_{\Phi,X} u^*(x) + P v^*(x) = R(x) \Rightarrow P^\top P v^*(x) = P^\top(A_{\Phi,X} u^*(x) - R(x)).$$

Since $P \in M_{N,Q}(\mathbb{R})$ has full column rank ($Q \leq N$) then $P^\top P \in M_Q(\mathbb{R})$ is invertible. We proved that also the smoothness of $\{v_1^*, \dots, v_Q^*\}$ depends on smoothness of Φ with respect to the first argument and that of $\mathcal{P}$.

If we denote with $D^\alpha(R)(x) = (D_1^\alpha(\Phi(x, x_1)), \dots, D_1^\alpha(\Phi(x, x_N)))^\top \in \mathbb{R}^N$ then we obtain, with Theorem 4.1,

$$\begin{pmatrix} A_{\Phi,X} & P \\ P^\top & 0 \end{pmatrix} \begin{pmatrix} D^\alpha(u^*)(x) \\ D^\alpha(v^*)(x) \end{pmatrix} = \begin{pmatrix} D^\alpha(R)(x) \\ D^\alpha(S)(x) \end{pmatrix}. \tag{4.6}$$

Under the assumption of Theorem 2.17 we achieve that $\mathcal{N}_\Phi(\Omega) \subseteq \mathcal{C}^k(\Omega)$ and $V_X \subseteq \mathcal{C}^k(\Omega)$, so it makes sense to study

$$D^\alpha(f - s_{f,X}) \quad \text{for } |\alpha| \leq k.$$

Definition 4.1 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is open and that $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ is a conditionally positive definite kernel on Ω with respect to $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$. If $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is $\mathcal{P}$-unisolvent then for every $x \in \Omega$ and $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ the power function is defined* *by*

$$(P^{\alpha}_{\Phi,X}(x))^2 = D_1^{\alpha} D_2^{\alpha} \Phi(x,x) - 2\sum_{j=1}^{N} D^{\alpha} u_j^{\star}(x) D_1^{\alpha}\Phi(x,x_j) + \\ + \sum_{i,j=1}^{N} D^{\alpha} u_i^{*}(x) D^{\alpha} u_j^{*}(x) \Phi(x_i,x_j),$$

where $\{u_1^*, \dots, u_N^*\}$ *are the cardinal functions with respect to* X *(equation* (4.2)*).*

To better understand this definition let us fix x, X, Φ and α and replace the constant vector $D^{\alpha}u^*(x) \in \mathbb{R}^N$ by an arbitrary vector $u \in \mathbb{R}^N$. We can define the quadratic form $\mathcal{Q} : \mathbb{R}^N \to \mathbb{R}$ by

$$\mathcal{Q}(u) = D_1^{\alpha} D_2^{\alpha} \Phi(x,x) - 2\sum_{j=1}^{N} u_j D_1^{\alpha}\Phi(x,x_j) + \\ + \sum_{i,j=1}^{N} u_i u_j \Phi(x_i,x_j) \quad u \in \mathbb{R}^N.$$

With this notation the power function becomes

$$(P^{\alpha}_{\Phi,X}(x))^2 = \mathcal{Q}(D^{\alpha}(u^*)(x)). \tag{4.7}$$

The next theorem shows that the Definition 4.1 is well-posed and provides a useful representation of the power function.

Theorem 4.2 *Suppose that* $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ *is a conditionally positive definite kernel with respect to* $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$. *Now suppose that* $u^{(\alpha)} \in \mathbb{R}^N$ *is a vector that satisfies* $\sum_{j=1}^{N} u_j^{(\alpha)} p(x_j) = D^{\alpha}p(x)$ *for all* $p \in \mathcal{P}$. *Then the quadratic form* $\mathcal{Q}$ *has the representation*

$$\mathcal{Q}(u^{(\alpha)}) = \left| D_2^{\alpha} G(\cdot,x) - \sum_{j=1}^{N} u_j^{(\alpha)} G(\cdot,x_j) \right|^2_{\mathcal{N}_{\Phi}(\Omega)},$$

where G *appears in equation* (2.5)*.*

Proof
From Theorem 2.16 we have that $D_2^{\alpha}G(\cdot,x) \in \mathcal{N}_{\Phi}(\Omega)$. We compute

$$\left| D_2^{\alpha} G(\cdot,x) - \sum_{j=1}^{N} u_j^{(\alpha)} G(\cdot,x_j) \right|^2_{\mathcal{N}_{\Phi}(\Omega)} = \\ = |D_2^{\alpha}G(\cdot,x)|^2_{\mathcal{N}_{\Phi}(\Omega)} - 2\sum_{j=1}^{N} u_j^{(\alpha)} \langle D_2^{\alpha}G(\cdot,x), G(\cdot,x_j)\rangle_{\mathcal{N}_{\Phi}(\Omega)} + \\ + \sum_{i,j=1}^{N} u_i^{(\alpha)} u_j^{(\alpha)} \langle G(\cdot,x_i), G(\cdot,x_j)\rangle_{\mathcal{N}_{\Phi}(\Omega)}. \tag{4.8}$$

From $G(\cdot,x)\in F_\Phi(\Omega)$ and equation (2.11) we get

$$\langle G(\cdot,x_i),G(\cdot,x_j)\rangle_{\mathcal{N}_\Phi(\Omega)}=\langle G(\cdot,x_i),G(\cdot,x_j)\rangle_\Phi,$$

that with equation 2.5 gives us

$$\begin{aligned}\langle G(\cdot,x_i),G(\cdot,x_j)\rangle_{\mathcal{N}_\Phi(\Omega)}=&\Phi(x_i,x_j)+\sum_{n,\ell=1}^{Q}p_n(x_i)p_\ell(x_j)\Phi(\xi_\ell,\xi_n)+\\&-\sum_{n=1}^{Q}p_n(x_i)\Phi(x_j,\xi_n)-\sum_{\ell=1}^{Q}p_\ell(x_j)\Phi(x_i,\xi_\ell).\end{aligned}$$

If we use the hypothesis $\sum_{j=1}^{N}u_j^{(\alpha)}p(x_j)=D^\alpha p(x)$ for all $p\in\mathcal{P}$ we can compute

$$\sum_{i,j=1}^{N}u_i^{(\alpha)}u_j^{(\alpha)}\sum_{n,\ell=1}^{Q}p_n(x_i)p_\ell(x_j)=\sum_{n,\ell=1}^{Q}\sum_{i,j=1}^{N}u_i^{(\alpha)}u_j^{(\alpha)}p_n(x_i)p_\ell(x_j)=\sum_{n,\ell=1}^{Q}D^\alpha p_n(x)D^\alpha p_\ell(x)$$

and

$$\sum_{i,j=1}^{N}u_i^{(\alpha)}u_j^{(\alpha)}\sum_{n=1}^{Q}p_n(x_i)=\sum_{n=1}^{Q}\sum_{i,j=1}^{N}u_i^{(\alpha)}u_j^{(\alpha)}p_n(x_i)=\sum_{n=1}^{Q}\sum_{j=1}^{N}u_j^{(\alpha)}D^\alpha p_n(x).$$

We can study the third term in equation 4.8

$$\begin{aligned}\sum_{i,j=1}^{N}u_i^{(\alpha)}u_j^{(\alpha)}\langle G(\cdot,x_i),G(\cdot,x_j)\rangle_{\mathcal{N}_\Phi(\Omega)}=&\boxed{\sum_{i,j=1}^{N}u_i^{(\alpha)}u_j^{(\alpha)}\Phi(x_i,x_j)}+\\&+\sum_{n,\ell=1}^{Q}D^\alpha p_n(x)D^\alpha p_\ell(x)\Phi(\xi_\ell,\xi_n)+\\&-2\sum_{n=1}^{Q}\sum_{j=1}^{N}u_j^{(\alpha)}D^\alpha p_n(x)\Phi(x_j,\xi_n).\end{aligned}$$

From Theorem 2.17 we know that $D_2^\alpha G(\cdot,x)\in\mathcal{N}_\Phi(\Omega)$ and it can be written as (Theorem 2.3)

$$D_2^\alpha G(y,x)-\sum_{\ell=1}^{Q}D_2^\alpha G(\xi_\ell,x)p_\ell(y)=\langle D_2^\alpha G(\cdot,x),G(\cdot,y)\rangle_{\mathcal{N}_\Phi(\Omega)}.$$

We can compute the second term in equation (4.8) as

$$\sum_{j=1}^{N}u_j^{(\alpha)}\langle D_2^\alpha G(\cdot,x),G(\cdot,x_j)\rangle_{\mathcal{N}_\Phi(\Omega)}=\sum_{j=1}^{N}u_j^{(\alpha)}\left(D_2^\alpha G(x_j,x)-\sum_{\ell=1}^{Q}D_2^\alpha G(\xi_\ell,x)p_\ell(x_j)\right).$$

If we apply Definition 2.5 to the previous equation we obtain

$$\sum_{j=1}^{N} u_j^{(\alpha)} D_2^\alpha G(x_j, x) = \boxed{\sum_{j=1}^{N} u_j^{(\alpha)} \underbrace{D_2^\alpha \Phi(x_j, x)}_{D_1^\alpha \Phi(x, x_j)}} - \sum_{j=1}^{N}\sum_{n=1}^{Q} u_j^{(\alpha)} D^\alpha p_n(x) \Phi(x_j, \xi_n)$$

and

$$\sum_{j=1}^{N}\sum_{\ell=1}^{Q} u_j^{(\alpha)} D_2^\alpha G(\xi_\ell, x) p_\ell(x_j) = \sum_{j=1}^{N}\sum_{\ell=1}^{Q} u_j^{(\alpha)} D_2^\alpha \Phi(\xi_\ell, x) p_\ell(x_j) +$$
$$- \sum_{j=1}^{N}\sum_{\ell=1}^{Q}\sum_{n=1}^{Q} u_j^{(\alpha)} D^\alpha p_n(x) \Phi(\xi_\ell, \xi_n) p_\ell(x_j) =$$
$$= \sum_{\ell=1}^{Q} u_j^{(\alpha)} D_2^\alpha \Phi(\xi_\ell, x) D^\alpha p_\ell(x) - \sum_{\ell=1}^{Q}\sum_{n=1}^{Q} u_j^{(\alpha)} D^\alpha p_n(x) \Phi(\xi_\ell, \xi_n) D^\alpha p_\ell(x),$$

where we used the property of $u^{(\alpha)} \in \mathbb{R}^N$ stated in the hypothesis.

Recalling Theorem 2.17 we have

$$D_1^\alpha D_2^\alpha G(y, x) - \sum_{\ell=1}^{Q} D_2^\alpha G(\xi_\ell, x) D^\alpha p_\ell(y) = \langle D_2^\alpha G(\cdot, x), D_2^\alpha G(\cdot, y) \rangle_{\mathcal{N}_\Phi(\Omega)}.$$

The first term of equation 4.8 becomes

$$\langle D_2^\alpha G(\cdot, x), D_2^\alpha G(\cdot, x) \rangle_{\mathcal{N}_\Phi(\Omega)} = D_1^\alpha D_2^\alpha G(x, x) - \sum_{\ell=1}^{Q} D_2^\alpha G(\xi_\ell, x) D^\alpha p_\ell(x).$$

If we apply Definition 2.5 to the previous equation we obtain

$$D_1^\alpha D_2^\alpha G(x, x) = \boxed{D_1^\alpha D_2^\alpha \Phi(x, x)} - \sum_{n=1}^{Q} D^\alpha p_n(x) \underbrace{D_1^\alpha \Phi(x, \xi_n)}_{= D_2^\alpha \Phi(\xi_n, x)}$$

and

$$\sum_{\ell=1}^{Q} D_2^\alpha G(\xi_\ell, x) D^\alpha p_\ell(x) = \sum_{\ell=1}^{Q} D_2^\alpha \Phi(\xi_\ell, x) D^\alpha p_\ell(x) - \sum_{\ell=1}^{Q}\sum_{n=1}^{Q} D^\alpha p_n(x) D^\alpha p_\ell(x) \Phi(\xi_\ell, \xi_n)$$

Remembering that $\Phi(x + he_i, y)/h = \Phi(y, x + he_i)/h$, we can prove by induction on $|\alpha|$ that $D_1^\alpha \Phi(x, y) = D_2^\alpha \Phi(y, x)$ holds.

We can achieve the result in the statement by summing up all the expressions we found.

□

If we restrict to positive definite kernels and $\alpha = 0$ then we can read Theorem 4.2 as

$$\mathcal{Q}(u) = \Phi(x,x) - 2\sum_{j=1}^{N} u_j \Phi(x,x_j) + \sum_{i,j=1}^{N} u_i u_j \Phi(x_i,x_j) =$$
$$= \left\| \Phi(\cdot,x) - \sum_{j=1}^{N} u_j \Phi(\cdot,x_j) \right\|^2_{\mathcal{N}_\Phi(\Omega)}.$$

Finally we can state an important fact that links the power function and the error.

Theorem 4.3 *Let $\Omega \subseteq \mathbb{R}^d$ be open. Suppose that $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ is a conditionally positive definite kernel on Ω with respect to $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$. Suppose that $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is $\mathcal{P}$-unisolvent. Denote the interpolant of $f \in \mathcal{N}_\Phi(\Omega)$ on X by $s_{f,X}$. Then for every $x \in \Omega$ and every $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ the error between f and its interpolant can be bounded by*

$$|D^\alpha(f)(x) - D^\alpha(s_{f,X})(x)| \leq P^\alpha_{\Phi,X}(x)|f|_{\mathcal{N}_\Phi(\Omega)}.$$

Proof
From equation (4.4) we have that $\sum_{j=1}^{N} p(x_j)u_j^*(x) = p(x)$ for each $x \in \Omega$. If we use equation (4.5) and we apply Theorem 2.3 to $f \in \mathcal{N}_\Phi(\Omega)$ then

$$D^\alpha(s_{f,X})(x) = \sum_{j=1}^{N} f(x_j) D^\alpha(u_j^*)(x) =$$
$$= \sum_{j=1}^{N} (\Pi_\mathcal{P}(f)(x_j) + \langle f, G(\cdot,x_j)\rangle_{\mathcal{N}_\Phi(\Omega)}) D^\alpha(u_j^*)(x) =$$
$$= D^\alpha(\Pi_\mathcal{P}(f))(x) + \left\langle f, \sum_{j=1}^{N} D^\alpha(u_j^*)(x) G(\cdot,x_j) \right\rangle_{\mathcal{N}_\Phi(\Omega)}.$$

If we apply Theorem 2.17 to $f \in \mathcal{N}_\Phi(\Omega)$ then

$$D^\alpha(f)(x) = D^\alpha(\Pi_\mathcal{P}(f))(x) + \langle f, D_2^\alpha(G(\cdot,x))\rangle_{\mathcal{N}_\Phi(\Omega)}.$$

Combining these two equations we obtain

$$|D^\alpha(f)(x) - D^\alpha(s_{f,X})(x)| = \left| \left\langle f, D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} D^\alpha(u_j^*)(x) G(\cdot,x_j) \right\rangle_{\mathcal{N}_\Phi(\Omega)} \right| \leq$$
$$\leq |f|_{\mathcal{N}_\Phi(\Omega)} \left| D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} D^\alpha(u_j^*)(x) G(\cdot,x_j) \right|_{\mathcal{N}_\Phi(\Omega)}.$$

We can conclude with Theorem 4.2 and equation (4.7).

□

The interpolation error depends on f through $|f|_{\mathcal{N}_\Phi(\Omega)}$ (independent of X) and on X through $P^\alpha_{\Phi,X}$ (independent of f). If we can bound the power function $P^\alpha_{\Phi,X}$ then we have an explicit error estimates for $f \in \mathcal{N}_\Phi(\Omega)$.

We show that the cardinal functions $\{u_1^*, \ldots, u_N^*\}$ (equation (4.2)) have a minimality property.

Theorem 4.4 *Let $\Omega \subseteq \mathbb{R}^d$ be open. Suppose that $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ is a conditionally positive definite kernel on Ω with respect to $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$. Suppose that $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ is $\mathcal{P}$-unisolvent. Define for $x \in \Omega$ and $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ the function $\mathcal{Q} : \mathbb{R}^d \to \mathbb{R}$ by*

$$\mathcal{Q}(u) = D_1^\alpha D_2^\alpha \Phi(x, x) - 2 \sum_{j=1}^N u_j D_1^\alpha \Phi(x, x_j) + \sum_{i,j=1}^N u_i u_j \Phi(x_i, x_j) \quad u \in \mathbb{R}^N.$$

The minimum of the function $\mathcal{Q}$ on the set

$$M = \left\{ u \in \mathbb{R}^N : \sum_{j=1}^N u_j p(x_j) = D^\alpha(p)(x) \text{ for all } p \in \mathcal{P} \right\}$$

is given by the vector $D^\alpha(u^*)(x)$. In other words

$$(P^\alpha_{\Phi,X}(x))^2 = \mathcal{Q}(D^\alpha(u^*)(x)) \leq \mathcal{Q}(u) \quad \text{for each } u \in M.$$

Proof
Using the notation of Theorem 4.1 we can rewrite $\mathcal{Q}$ as

$$\mathcal{Q}(u) = D_1^\alpha D_2^\alpha \Phi(x, x) - 2u^\top D^\alpha(R)(x) + u^\top A_{\Phi,X} u$$

and M becomes

$$M = \{u \in \mathbb{R}^N : P^\top u = D^\alpha(S)(x)\}.$$

From equation (4.6) we state that $M \neq \emptyset$. Since Φ is conditionally positive definite on Ω then $A_{\Phi,X}$ is positive definite on $M_0 = \{u \in \mathbb{R}^N : P^\top u = 0\}$. Theorem 3.3 ensures the existence and uniqueness of a minimum.

Define the Lagrange function $\mathcal{L}$ as

$$\mathcal{L}(u, \lambda) = D_1^\alpha D_2^\alpha \Phi(x, x) - 2u^\top D^\alpha(R)(x) + u^\top A_{\Phi,X} u + \lambda^\top (P^\top u - D^\alpha(S)(x))$$

and impose

$$\nabla_u \mathcal{L}(u^*, \lambda^*) = 0 \quad \text{and} \quad \nabla_\lambda \mathcal{L}(u^*, \lambda^*) = 0.$$

These imposed conditions give us

$$\begin{aligned} 0 &= -2D^\alpha(R)(x) + 2A_{\Phi,X} u^* + P\lambda^* \\ 0 &= P^\top u^* - D^\alpha(S)(x), \end{aligned}$$

that shows with $\lambda^* = 2v^*$

$$A_{\Phi,X}u^* + Pv^* = D^\alpha(R)(x)$$
$$P^\top u^* = D^\alpha(S)(x),$$

By Theorem 2.1 and equation (4.6) we obtain that the unique solution is $u^* = D^\alpha(u^*)(x)$ and $v^* = D^\alpha(v^*)(x)$.

□

As a consequence of Theorem 4.4 we can bound the power function $P^\alpha_{\Phi,X}(x)$ (Definition 4.1) applying to $\mathcal{Q}$ (equation (4.7)) a vector $\widetilde{u}^{(\alpha)}(x) \in \mathbb{R}^N$ such that $P^\top \widetilde{u}^{(\alpha)}(x) = D^\alpha(S)(x)$.

4.2 Scaling and power function

For our purposes it is useful to scale the (conditionally) positive definite functions we are considering (equation (1.14)).

We state that

$$P_{\Phi_\delta,X}(x) = P_{\Phi,X/\delta}(x/\delta). \tag{4.9}$$

If Φ_δ is conditionally positive definite with respect to $\mathcal{P}$ on Ω then Φ is conditionally positive definite with respect to $\mathcal{P}_\delta = \{p(\delta\cdot) : p \in \mathcal{P}\}$ on Ω/δ.

Suppose that $X = \{x_1, \dots, x_N\} \subseteq \Omega$ are pairwise distinct points such that

$$\sum_{j=1}^N \alpha_j p\left(\delta \frac{x_j}{\delta}\right) = 0 \Leftrightarrow \sum_{j=1}^N \alpha_j p(x_j) = 0$$

for all $p \in \mathcal{P}$ then, since Φ_δ is conditionally positive definite,

$$\sum_{j,k=1}^N \alpha_j\alpha_k \Phi\left(\frac{x_j - x_k}{\delta}\right) > 0 \Leftrightarrow \sum_{j,k=1}^N \alpha_j\alpha_k \Phi\left(\frac{x_j}{\delta} - \frac{x_k}{\delta}\right) > 0.$$

We also remark that if $X = \{x_1, \dots, x_N\} \subseteq \Omega$ is $\mathcal{P}$-unisolvent then $X/\delta = \{x_1/\delta, \dots, x_N/\delta\} \subseteq \Omega/\delta$ is $\mathcal{P}_\delta$-unisolvent. Fix $p(\delta\cdot) \in \mathcal{P}_\delta$ with $p \in \mathcal{P}$ such that

$$p\left(\delta\frac{x_i}{\delta}\right) = 0 \quad i = 1, \dots, N \Rightarrow p(x_i) = 0 \quad i = 1, \dots, N \Rightarrow p = 0 \Rightarrow p(\delta\cdot) = 0.$$

Let us fix some notation. From equation (4.2) the cardinal function of Φ_δ with respect to X are

$$u_j^*(x) = \sum_{i=1}^N \alpha_i^{(j)} \Phi_\delta(x - x_i) + \sum_{k=1}^Q \beta_k^{(j)} p_k(x) = \sum_{i=1}^N \alpha_i^{(j)} \Phi\left(\frac{x - x_i}{\delta}\right) + \sum_{k=1}^Q \beta_k^{(j)} p_k(x)$$

for $j = 1, \dots, N$ where $(\alpha^{(j)}, \beta^{(j)}) \in \mathbb{R}^{N+Q}$ solves the system in equation (4.1).

We state that $\{u_1^*(\delta\cdot), \dots, u_N^*(\delta\cdot)\}$ are the cardinal function of Φ with respect to X/δ because

$$u_j^*(\delta\cdot) = \sum_{i=1}^{N} \alpha_i^{(j)} \Phi\left(\cdot - \frac{x_i}{\delta}\right) + \underbrace{\sum_{k=1}^{Q} \beta_k^{(j)} p_k(\delta\cdot)}_{\in \mathcal{P}_\delta} \quad \text{for } j = 1, \dots, N$$

and

$$u_j^*\left(\delta \frac{x_i}{\delta}\right) = u_j^*(x_i) = \delta_{i,j}.$$

Finally, from Definition (4.1) we get

$$\begin{aligned}
(P_{\Phi_\delta,X}(x))^2 &= \Phi_\delta(0) - 2\sum_{j=1}^{N} u_j^\star(x)\Phi_\delta(x - x_j) + \sum_{i,j=1}^{N} u_i^*(x)u_j^*(x)\Phi_\delta(x_i - x_j) = \\
&= \Phi\left(\frac{0}{\delta}\right) - 2\sum_{j=1}^{N} u_j^\star(x)\Phi\left(\frac{x - x_j}{\delta}\right) + \sum_{i,j=1}^{N} u_i^*(x)u_j^*(x)\Phi\left(\frac{x_i - x_j}{\delta}\right) = \\
&= \Phi(0) - 2\sum_{j=1}^{N} u_j^\star\left(\delta\frac{x}{\delta}\right)\Phi\left(\frac{x}{\delta} - \frac{x_j}{\delta}\right) + \sum_{i,j=1}^{N} u_i^\star\left(\delta\frac{x}{\delta}\right)u_j^\star\left(\delta\frac{x}{\delta}\right)\Phi\left(\frac{x_i}{\delta} - \frac{x_j}{\delta}\right) = \\
&= \left(P_{\Phi,\frac{X}{\delta}}\left(\frac{x}{\delta}\right)\right)^2
\end{aligned}$$

From Theorem 1.26 we have to take into account that the native space norm of $\mathcal{N}_{\Phi_\delta}(\Omega)$ varies with δ.

4.3 Improved error estimates

In Theorem 4.3 the difference between f and its interpolant $s_{f,X}$ at a point $x \in \Omega$ is bounded by the power function $P_{\Phi_\delta,X}(x)$ and the semi-norm $|f|_{\mathcal{N}_\Phi(\Omega)}$. We want to improve the error estimates by getting rid of $|f|_{\mathcal{N}_\Phi(\Omega)}$.

If $f \in \mathcal{N}_\Phi(\Omega)$ then $f - s_{f,X} \in \mathcal{N}_\Phi(\Omega)$ and it vanishes on X. So, $s_{f-s_{f,X},X} = 0$ with Theorem 4.3 gives us

$$|f(x) - s_{f,X}(x)| \le P_{\Phi,X}(x)|f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)}. \tag{4.10}$$

We restrict our analysis for positive definite kernel on a compact set $\Omega \subseteq \mathbb{R}^d$.

Theorem 4.5 *Suppose that Φ is a symmetric positive definite kernel on a compact set $\Omega \subseteq \mathbb{R}^d$. Then for every $f \in T(L^2(\Omega))$ we have*

$$|f(x) - s_{f,X}(x)| \le P_{\Phi,X}(x)\|P_{\Phi,X}\|_{L^2(\Omega)}\|T^{-1}(f)\|_{L^2(\Omega)} \quad \textit{for } x \in \Omega,$$

where T is defined in equation (2.18).

Proof
Fix $f = T(v)$ with $v \in L^2(\Omega)$. By integrating the square of equation (4.10) we obtain

$$\|f - s_{f,X}\|_{L^2(\Omega)} \le |f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \|P_{\Phi,X}\|_{L^2(\Omega)}.$$

We remark that Theorem 2.8 and Theorem 2.12 give us

$$\begin{aligned} |f - s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} &= \langle f - s_{f,X}, f - s_{f,X}\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle f - s_{f,X}, f\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \langle f - s_{f,X}, T(v)\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle f - s_{f,X}, v\rangle_{L^2(\Omega)} \le \\ &\le \|f - s_{f,X}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \le |f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \|P_{\Phi,X}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}, \end{aligned}$$

that proves

$$|f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \le \|P_{\Phi,X}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

If we apply this inequality to equation (4.10) we can conclude.

□

Suppose that we have an estimate for the power function $P_{\Phi,X}$ of the form

$$P_{\Phi,X}(x) \le r(h_{X,\Omega}),$$

then Theorem 4.3 gives us

$$\|f - s_{f,X}\|_{L^\infty(\Omega)} \le r(h_{X,\Omega}) |f|_{\mathcal{N}_\Phi(\Omega)}$$

instead Theorem 4.5 gives us

$$\|f - s_{f,X}\|_{L^\infty(\Omega)} \le r(h_{X,\Omega})^2 \|T^{-1}(f)\|_{L^2(\Omega)} \sqrt{\mathcal{L}(\Omega)}, \tag{4.11}$$

where $\mathcal{L}$ is the Lebesgue measure on $\mathbb{R}^d$. We note that if we scale the function Φ with a parameter δ (equation (1.14)) then the integral operator T (equation (2.18)) depends on δ.

Chapter 5

Stability

Now we will briefly discuss the stability of the interpolation process with radial basis functions [34, 35, 36, 37, 38, 39].

In this section we will work with a conditionally positive definite kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ with respect to $\mathcal{P}$. We will use the notation introduced in equation 2.2.

To study the stability of the interpolation process we introduce the following quantity:

$$\lambda_{\min}(A_{\Phi,X}) = \inf_{\alpha \in \mathbb{R}^d, P^\top \alpha = 0} \frac{\alpha^\top A_{\Phi,X} \alpha}{\alpha^\top \alpha} \tag{5.1}$$

Suppose we have two functions $f, \widetilde{f} : \Omega \to \mathbb{R}^d$ and consider $(\alpha, \beta), (\widetilde{\alpha}, \widetilde{\beta})$ the solutions of the system in equation (2.2) with f and $\widetilde{f}$ respectively. Then we have that

$$A_{\Phi,X}(\alpha - \widetilde{\alpha}) + P(\beta - \widetilde{\beta}) = f_{|X} - \widetilde{f}_{|X}$$

with

$$P^\top(\alpha - \widetilde{\alpha}) = 0.$$

We can write

$$(\alpha - \widetilde{\alpha})^\top A_{\Phi,X}(\alpha - \widetilde{\alpha}) + \underbrace{(\alpha - \widetilde{\alpha})^\top P(\beta - \widetilde{\beta})}_{(\beta - \widetilde{\beta})^\top P^\top (\alpha - \widetilde{\alpha}) = 0} = (\alpha - \widetilde{\alpha})^\top (f_{|X} - \widetilde{f}_{|X}),$$

that implies

$$\begin{aligned} \|\alpha - \widetilde{\alpha}\|_2^2 \lambda_{\min}(A_{\Phi,X}) &\le (\alpha - \widetilde{\alpha})^\top A_{\Phi,X}(\alpha - \widetilde{\alpha}) = \\ &= (\alpha - \widetilde{\alpha})^\top (f_{|X} - \widetilde{f}_{|X}) \le \|\alpha - \widetilde{\alpha}\|_2 \|f_{|X} - \widetilde{f}_{|X}\|_2. \end{aligned}$$

It is possible to conclude

$$\|\alpha - \widetilde{\alpha}\|_2 \le \frac{1}{\lambda_{\min}(A_{\Phi,X})} \|f_{|X} - \widetilde{f}_{|X}\|_2, \tag{5.2}$$

which explains why it is worth to study a lower bound of $\lambda_{\min}(A_{\Phi,X})$ to understand the stability of the interpolation scheme. We note that in Theorem 3.3 we proved that $\lambda_{\min}(A_{\Phi,X}) > 0$ since $A_{\Phi,X}$ is positive definite on $\{x \in \mathbb{R}^N : P^\top x = 0\}$.

For now restrict our work for symmetric positive definite kernels. In this case the condition number of the interpolation matrix $A_{\Phi,X}$ is

$$\operatorname{cond}(A_{\Phi,X}) = \|A_{\Phi,X}\|_2 \|A_{\Phi,X}^{-1}\|_2 = \frac{\lambda_{\max}(A_{\Phi,X})}{\lambda_{\min}(A_{\Phi,X})},$$

where $\lambda_{\max}(A_{\Phi,X})$ and $\lambda_{\min}(A_{\Phi,X})$ denotes respectively the maximum and the minimum eigenvalue of the matrix $A_{\Phi,X}$.

Suppose that $v \in \mathbb{R}^N \setminus \{0\}$ is an eigenvector for the eigenvalue $\lambda_{\max}(A_{\Phi,X})$ then

$$\sum_{j=1}^{N} \Phi(x_i, x_j) v_j = \lambda_{\max}(A_{\Phi,X}) v_i \quad \text{for } i = 1, \ldots, N$$

and

$$\sum_{\substack{j=1 \\ j \neq i}}^{N} \Phi(x_i, x_j) v_j = (\lambda_{\max}(A_{\Phi,X}) - \Phi(x_i, x_i)) v_i \quad \text{for } i = 1, \ldots, N.$$

There exists $i \in \{1, \ldots, N\}$ such that $|v_i|$ is maximal and strictly positive. We can write

$$|\lambda_{\max}(A_{\Phi,X}) - \Phi(x_i, x_i)| \leq \sum_{\substack{j=1 \\ j \neq i}}^{N} |\Phi(x_i, x_j)| \underbrace{\frac{|v_j|}{|v_i|}}_{\leq 1} \leq \sum_{\substack{j=1 \\ j \neq i}}^{N} |\Phi(x_i, x_j)| \leq N \|\Phi\|_{L^\infty(\Omega \times \Omega)}. \tag{5.3}$$

The condition number of the interpolation matrix $A_{\Phi,X}$ gives an intuition on the numerical stability of the interpolation scheme.

If we consider $X \subseteq \mathbb{R}^d$ to be quasi-uniform (Definition 3.6) then from Proposition 3.6 we obtain that

$$\lambda_{\max}(A_{\Phi,X}) = \mathcal{O}\left(\frac{1}{h_{X,\Omega}^d}\right).$$

5.1 Trade-off principle

In this section we will build a link between the smallest eigenvalue of the main part of the interpolation matrix and the power function.

Theorem 5.1 *If $u_1^*, \ldots, u_N^*$ are the cardinal functions defined in Theorem 4.1 then for all $x \in \Omega \setminus X$*

$$\frac{P_{\Phi,X}^2(x)}{\lambda_{\min}(A_{\Phi,X \cup \{x\}})} \geq 1 + \sum_{j=1}^{N} |u_j^*(x)|^2.$$

Proof
Let us fix some notation: $u_0^*(x) = -1$ and $x_0 = x$. Definition 4.1 gives

$$P_{\Phi,X}^2(x) = \sum_{j,k=0}^{N} u_j^*(x) u_k^*(x) \Phi(x_j, x_k)$$

with

$$p(x) = \sum_{j=1}^{N} p(x_j) u_j^*(x) \text{ for all } p \in \mathcal{P} \Rightarrow \sum_{j=0}^{N} p(x_j) u_j^*(x) = 0 \text{ for all } p \in \mathcal{P}.$$

From equation 5.1 we obtain

$$P_{\Phi,X}^2(x) \geq \lambda_{\min}(A_{\Phi,X\cup\{x\}}) \sum_{j=0}^{N} |u_j^*(x)|^2 = \lambda_{\min}(A_{\Phi,X\cup\{x\}}) \left(1 + \sum_{j=1}^{N} |u_j^*(x)|^2\right)$$

□

The first consequence of Theorem 5.1 is

$$\lambda_{\min}(A_{\Phi,X}) \leq \min_{1\leq k\leq N} P_{\Phi,X\setminus\{x_k\}}^2(x_k),$$

that with a scaling of Φ (equation (4.9)) holds

$$\lambda_{\min}(A_{\Phi_\delta,X}) \leq \min_{1\leq k\leq N} P_{\Phi,\frac{\Omega}{\delta}\setminus\{\frac{x_k}{\delta}\}}^2 \left(\frac{x_k}{\delta}\right).$$

We can read these inequalities as: an upper-bound for $\lambda_{\min}(A_{\Phi_\delta,X})$ is the power function and a lower bound for the power function is $\lambda_{\min}(A_{\Phi_\delta,X})$. Now we can understand the trade-off principle because to guarantee a stable interpolation process we need $\lambda_{\min}(A_{\Phi_\delta,X})$ to be large but this implies that the power function is large so the error estimates in Theorem 4.3 is bad. Conversely, a small power function leads to an unstable interpolation scheme.

Theorem 5.1 gives us a bound for the Lebesgue function of the interpolation process with conditionally positive definite kernels, indeed

$$\left|\sum_{j=1}^{N} f(x_j) u_j^*(x)\right| \leq \sum_{j=1}^{N} |f(x_j)||u_j^*(x)| \leq \|f_{|X}\|_2 \|u^*(x)\|_2,$$

where we can bound $\|u^*(x)\|_2$ with

$$\|u^*(x)\|_2 \leq \frac{P_{\Phi,X}^2(x)}{\lambda_{\min}(A_{\Phi,X\cup\{x\}})} - 1.$$

5.2 Lower bounds for minimum eigenvalue

From numerical experiments it can be seen that the fill distance $h_{X,\Omega}$ is not a good index to measure stability. A point set X might have a big fill distance and the interpolation process can be not well-conditioned: if we cover densely only a portion of Ω then the fill distance can be large and a lot of points are close to each other.

It is more natural to bound $\lambda_{\min}(A_{\Phi,X})$ with the separation distance q_X (Definition 3.5).

The idea to bound $\lambda_{\min}(A_{\Phi,X})$ follows from the following inequality:

$$\sum_{j,k=1}^{N} \alpha_j \alpha_k \Phi(x_j, x_k) \geq \sum_{j,k=1}^{N} \alpha_j \alpha_k \Psi(x_j, x_k) \geq \lambda \|\alpha\|_2^2, \tag{5.4}$$

indeed $\lambda_{\min}(A_{\Phi,X}) \leq \lambda$. Our goal is to find an appropriate positive definite kernel $\Psi : \Omega \times \Omega \to \mathbb{R}$.

We focus on conditionally positive definite functions and we will see equation (5.4) from a different perspective, so we need a different representation for Φ. It is necessary to introduce some preliminary results.

Definition 5.1 *The Schwartz space $\mathcal{S}$ [40] consists of all functions $\gamma \in \mathcal{C}^\infty(\mathbb{R}^d)$ that satisfy*

$$|x^\alpha D^\beta(\gamma)(x)| \leq C_{\alpha,\beta,\gamma}, \quad x \in \mathbb{R}^d,$$

for all multi-index $\alpha, \beta \in \mathbb{N}^d$ with a constant $C_{\alpha,\beta,\gamma}$ independent of $x \in \mathbb{R}^d$.

Definition 5.2 *We say that a function f is slowly increasing if there exists a constant $m \in \mathbb{N}$ such that $f(x) = \mathcal{O}_\infty(\|x\|_2^m)$.*

Definition 5.3 *For $m \in \mathbb{N}$ the set of all functions $\gamma \in \mathcal{S}$ (Definition 5.1) that satisfy $\gamma(\omega) = \mathcal{O}_0(\|\omega\|_2^m)$ will be denoted with $\mathcal{S}_m$.*

Definition 5.4 *Suppose that $\Phi : \mathbb{R}^d \to \mathbb{C}$ is continuous and slowly increasing. A measure able function $\widehat{\Phi} \in L^2_{loc}(\mathbb{R}^d \setminus \{0\})$ is called the generalized Fourier transform of Φ of order $m \in \mathbb{N}$ if*

$$\int_{\mathbb{R}^d} \Phi(x)\widehat{\gamma}(x)\,dx = \int_{\mathbb{R}^d} \widehat{\Phi}(\omega)\gamma(\omega)\,d\omega$$

holds for all $\gamma \in \mathcal{S}_{2m}$.

Since $\mathcal{O}_0(\|\omega\|_2^\ell) \subseteq \mathcal{O}_0(\|\omega\|_2^m)$ for $\ell \geq m$ then if $\widehat{\Phi}$ has order m then it has also order ℓ.

Theorem 5.2 *Suppose that $\Phi : \mathbb{R}^d \to \mathbb{C}$ is continuous, slowly increasing and admits a generalized Fourier transform $\widehat{\Phi} \in \mathcal{C}(\mathbb{R}^d \setminus \{0\})$ of order $m \in \mathbb{N}$. Then Φ is conditionally positive definite with respect to $\Pi_{m-1}(\mathbb{R}^d)$ if and only if $\widehat{\Phi}$ is non-negative and non-vanishing.*

Theorem 5.3 *Suppose that $\Phi : \mathbb{R}^d \to \mathbb{C}$ is continuous and slowly increasing. Φ admits a non-negative and non-vanishing generalized Fourier transform $\widehat{\Phi} \in \mathcal{C}(\mathbb{R}^d \setminus \{0\})$ of order m. Then for all pairwise distinct points $x_1, \ldots, x_N \in \mathbb{R}^d$ we have*

$$\sum_{j,k=1}^{N} \alpha_j \overline{\alpha_k} \Phi(x_j - x_k) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \left| \sum_{j=1}^{N} \alpha_j e^{i\langle \omega, x_j \rangle} \right|^2 \widehat{\Phi}(\omega) d\omega,$$

if $\sum_{j=1}^{N} \alpha_j p(x_j) = 0$ for all $p \in \Pi_{m-1}(\mathbb{R}^d)$.

With Theorem 5.3 if $\widehat{\Phi}(\omega) \geq \widehat{\Psi}(\omega)$ for $\omega \in \mathbb{R}^d$ and the order of $\widehat{\Psi}$ is smaller than the order of $\widehat{\Phi}$ then equation (5.4) holds. To build Ψ we need some result on Fourier transform of radial functions.

Definition 5.5 *The Bessel function [41] of the first kind of order $\nu \in \mathbb{C}$ is defined by*

$$J_\nu(z) = \sum_{m=0}^{+\infty} \frac{(-1)^m (z/2)^{2m+\nu}}{m! \Gamma(\nu + m + 1)}$$

for $z \in \mathbb{C} \setminus \{0\}$. Γ is the usual gamma function [42].

Theorem 5.4 *If we denote for $d \geq 2$ the unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$ then we have, for $x \in \mathbb{R}^d$*

$$\int_{\mathbb{S}^{d-1}} e^{i\langle x, y \rangle} dS(y) = (2\pi)^{\frac{d}{2}} \|x\|_2^{-\frac{d-2}{2}} J_{\frac{d-2}{2}}(\|x\|_2).$$

Proof

Suppose that $x \in \mathbb{R}^d$ then there exists an orthogonal transformation $R \in O_d(\mathbb{R})$ such that $R(x) = \|x\|_2 e_1$. By recalling the spherical coordinates $y = \sigma(\theta_1, \ldots, \theta_{d-1})$

$$\begin{aligned}
y_1 &= \cos(\theta_1) \\
y_2 &= \sin(\theta_1) \cos(\theta_2) \\
y_3 &= \sin(\theta_1) \sin(\theta_2) \cos(\theta_3) \\
&\vdots \\
y_{d-1} &= \sin(\theta_1) \cdots \sin(\theta_{d-2}) \cos(\theta_{d-1}) \\
y_d &= \sin(\theta_1) \cdots \sin(\theta_{d-2}) \sin(\theta_{d-1})
\end{aligned}$$

with $\theta_1, \ldots, \theta_{d-1} \in [0, \pi]$ and $\theta_d \in [0, 2\pi[$ we obtain

$$\mathrm{d}_{\mathbb{S}^{d-1}}(\theta_1, \ldots, \theta_{d-1}) = \sin(\theta_1)^{d-2} \sin(\theta_2)^{d-3} \cdots \sin(\theta_{d-2}) \mathrm{d}\theta_1 \mathrm{d}\theta_2 \cdots \mathrm{d}\theta_{d-1}.$$

By recalling that $\langle x, R(y) \rangle = \langle R(x), y \rangle$ and $R(v_1 \times \cdots \times v_{n-1}) = R(v_1) \times \cdots \times R(v_{n-1})$ since $R \in O_d(\mathbb{R})$ if we parameterize $\mathbb{S}^{d-1}$ with $R \circ \sigma$ then we obtain

$$\int_{\mathbb{S}^{d-1}} e^{i\langle x, y \rangle} \mathrm{d}\, S(y) = |\mathbb{S}^{d-2}| \int_0^\pi e^{i\|x\|_2 \cos(\theta)} \sin(\theta)^{d-2} \mathrm{d}\,\theta,$$

where

$$|\mathbb{S}^{d-2}| = \frac{2\pi^{\frac{d-1}{2}}}{\Gamma\left(\frac{d-1}{2}\right)}.$$

We can continue with

$$\int_0^{\pi} e^{i\|x\|_2 \cos(\theta)} \sin(\theta)^{d-2} \mathrm{d}\,\theta = \int_0^{\pi} \left(\sum_{k=0}^{+\infty} \frac{(i\|x\|_2 \cos(\theta))^k}{k!}\right) \sin(\theta)^{d-2} \mathrm{d}\,\theta =$$

$$= \sum_{k=0}^{+\infty} \frac{i^k \|x\|_2^k}{k!} \underbrace{\int_0^{\pi} \cos(\theta)^k \sin(\theta)^{d-2} \mathrm{d}\,\theta}_{a_k}.$$

By induction we can prove that $a_{2k+1} = 0$ for $k \in \mathbb{N}$ and

$$a_{2k} = \frac{(2k)!\Gamma((d-1)/2)\Gamma(1/2)}{2^{2k}k!\Gamma((2k+d)/2)}.$$

We can conlude with Definition 5.5,

$$\frac{2\pi^{(d-1)/2}}{\Gamma((d-1)/2)} \sum_{k=0}^{+\infty} \frac{i^{2k}\|x\|_2^{2k}}{(2k)!} \frac{(2k)!\Gamma((d-1)/2)\overbrace{\Gamma(1/2)}^{\pi^{1/2}}}{2^{2k}k!\Gamma((2k+d)/2)} =$$

$$=2\pi^{d/2} \sum_{k=0}^{+\infty} (-1)^k \left(\frac{\|x\|_2}{2}\right)^{2k} \frac{1}{k!\Gamma((2k+d)/2)} =$$

$$=2\pi^{d/2} \left(\frac{\|x\|_2}{2}\right)^{-(d-2)/2} \sum_{k=0}^{+\infty} (-1)^k \left(\frac{\|x\|_2}{2}\right)^{2k+(d-2)/2} \frac{1}{k!\Gamma(k+(d-2)/2+1)} =$$

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

□

Theorem 5.5 *Suppose that $\Phi \in L^1(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d)$ is radial, i.e. $\Phi(x) = \varphi(\|x\|_2)$ for each $x \in \mathbb{R}^d$. Then its Fourier transform $\widehat{\Phi}$ is also radial, i.e. $\widehat{\Phi}(\omega) = \mathcal{F}_d(\varphi)(\|\omega\|_2)$ for $\omega \in \mathbb{R}^d$ with*

$$\mathcal{F}_d(\varphi)(r) = r^{-(d-2)/2} \int_0^{\infty} \varphi(t) t^{d/2} J_{(d-2)/2}(rt)\, dt.$$

Proof
Suppose $d \geq 2$. Using Theorem 5.4 we have

$$\begin{aligned}
\widehat{\Phi}(x) &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(\omega) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \varphi(\|\omega\|_2) e^{-i\langle x,\omega\rangle} \mathrm{d}\,\omega = \\
&= (2\pi)^{-d/2} \int_0^{+\infty} \left(\int_{\partial B(0,t)} \varphi(\|\omega\|_2) e^{-i\langle x,\omega\rangle} \mathrm{d}S(\omega) \right) \mathrm{d}\,t = \\
&= (2\pi)^{-d/2} \int_0^{+\infty} t^{d-1} \left(\int_{\partial B(0,1)} \varphi(t\|\omega\|_2) e^{-i\langle x,t\omega\rangle} \mathrm{d}S(\omega) \right) \mathrm{d}\,t = \\
&= (2\pi)^{-d/2} \int_0^{+\infty} t^{d-1} \varphi(t) \left(\int_{\partial B(0,1)} e^{-i\langle tx,\omega\rangle} \mathrm{d}S(\omega) \right) \mathrm{d}\,t = \\
&= (2\pi)^{-d/2} \int_0^{+\infty} t^{d-1} \varphi(t) (2\pi)^{\frac{d}{2}} \|tx\|_2^{-\frac{d-2}{2}} J_{\frac{d-2}{2}}(\|tx\|_2) \mathrm{d}\,t = \\
&= \|x\|_2^{-\frac{d-2}{2}} \int_0^{+\infty} t^{\frac{d}{2}} \varphi(t) J_{\frac{d-2}{2}}(t\|x\|_2) \mathrm{d}\,t.
\end{aligned}$$

Instead for $d = 1$ we have

$$\begin{aligned}
\widehat{\Phi}(x) &= (2\pi)^{-1/2} \int_{\mathbb{R}} \Phi(\omega) e^{-ix\omega} \mathrm{d}\,\omega = (2\pi)^{-1/2} \int_{\mathbb{R}} \varphi(|\omega|) e^{-ix\omega} \mathrm{d}\,\omega = \\
&= (2\pi)^{-1/2} \left(\int_{\mathbb{R}_{>0}} \varphi(|\omega|) e^{-ix\omega} \mathrm{d}\,\omega + \underbrace{\int_{\mathbb{R}_{<0}} \varphi(|\omega|) e^{-ix\omega} \mathrm{d}\,\omega}_{\substack{\text{change of variable:} \\ \omega \mapsto -\omega}} \right) = \\
&= (2\pi)^{-1/2} \left(\int_{\mathbb{R}_{>0}} \varphi(|\omega|) e^{-ix\omega} \mathrm{d}\,\omega + \int_{\mathbb{R}_{>0}} \varphi(|\omega|) e^{ix\omega} \mathrm{d}\,\omega \right) = \\
&= (2\pi)^{-1/2} \int_0^{\infty} \varphi(\omega)(e^{-ix\omega} + e^{ix\omega}) \mathrm{d}\,\omega = (2\pi)^{-1/2} \int_0^{\infty} \varphi(\omega) 2\cos(x\omega) \mathrm{d}\,\omega = \\
&= \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_0^{\infty} \varphi(t) \cos(t|x|) \mathrm{d}\,t = |x|^{1/2} \int_0^{\infty} \varphi(t) t^{1/2} \left(\frac{2}{\pi t|x|}\right)^{\frac{1}{2}} \cos(t|x|) \mathrm{d}\,t.
\end{aligned}$$

We can conclude if

$$J_{-\frac{1}{2}}(t) = \left(\frac{2}{\pi t}\right)^{\frac{1}{2}} \cos(t).$$

By Definiton 5.5 we have

$$J_{-\frac{1}{2}}(t) = \sum_{m=0}^{+\infty} \frac{(-1)^m (t/2)^{2m-1/2}}{m!\Gamma(m+\frac{1}{2})} = \left(\frac{2}{t}\right)^{\frac{1}{2}} \sum_{m=0}^{+\infty} \frac{(-1)^m (t/2)^{2m}}{m! \underbrace{\Gamma\left(m+\frac{1}{2}\right)}_{\frac{(2m)!}{4^m m!}\pi^{1/2}}} =$$

$$= \left(\frac{2}{t\pi}\right)^{\frac{1}{2}} \sum_{m=0}^{+\infty} \frac{(-1)^m t^{2m}}{(2m)!} = \left(\frac{2}{t\pi}\right)^{\frac{1}{2}} \cos(t).$$

□

Now we can start to build Ψ.

Theorem 5.6 *Let χ_M be the characteristic function of $B(0, M)$ with $M > 0$, i.e. $\chi_M(x) = 1$ if $\|x\| \leq M$ and $\chi_M(x) = 0$ otherwise. Then*

$$\widehat{\chi_M}(x) = (\chi_M)^\vee(x) = M^{\frac{d}{2}} \|x\|_2^{-\frac{d}{2}} J_{\frac{d}{2}}(M\|x\|_2),$$

where J_ν is a Bessel function of the first kind (Definition 5.5).

Proof

We note that χ_M is radial. First we prove that $\widehat{\chi_M}(x) = (\chi_M)^\vee(x)$:

$$\widehat{\chi_M}(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \chi_M(\omega) e^{-ix\omega} \mathrm{d}\,\omega =$$

$$= (2\pi)^{-d/2} \underbrace{\int_{\mathbb{R}^d} \chi_M(-\omega) e^{ix\omega} \mathrm{d}\,\omega}_{\substack{\text{change of variable:} \\ \omega \mapsto -\omega}} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \chi_M(\omega) e^{ix\omega} \mathrm{d}\,\omega = (\chi_M)^\vee(x).$$

With Theorem 5.5 we can obtain

$$\widehat{\chi_M}(x) = \|x\|_2^{-(d-2)/2} \int_0^M t^{d/2} J_{(d-2)/2}(\|x\|_2 t) \mathrm{d}\,t.$$

Using Definition 5.5 we get

$$\widehat{\chi_M}(x) = \|x\|_2^{-(d-2)/2} \int_0^M t^{d/2} \left(\sum_{m=0}^{+\infty} \frac{(-1)^m (\|x\|_2 t/2)^{2m+(d-2)/2}}{m!\Gamma((d-2)/2+m+1)} \right) \mathrm{d}\,t =$$

$$= \|x\|_2^{-d/2+1} \sum_{m=0}^{+\infty} \int_0^M \frac{(-1)^m (\|x\|_2 t/2)^{2m+(d-2)/2}}{m!\Gamma(d/2+m)} t^{d/2} \mathrm{d}\,t =$$

$$= \|x\|_2^{-d/2+1} \sum_{m=0}^{+\infty} \frac{(-1)^m (\|x\|_2/2)^{2m+(d-2)/2}}{m!\Gamma(d/2+m)} \int_0^M t^{2m+d-1} \mathrm{d}\,t =$$

$$= \|x\|_2^{-d/2} \sum_{m=0}^{+\infty} \frac{(-1)^m (\|x\|_2/2)^{2m+d/2} (1/2)^{-1}}{m!\Gamma(d/2+m)} \frac{M^{2m+d}}{2m+d} =$$

$$= M^{d/2} \|x\|_2^{-d/2} \sum_{m=0}^{+\infty} \frac{(-1)^m (M\|x\|_2/2)^{2m+d/2}}{m!\Gamma(d/2+m)(m+d/2)} = M^{\frac{d}{2}} \|x\|_2^{-\frac{d}{2}} J_{\frac{d}{2}}(M\|x\|_2),$$

because $\Gamma(d/2+m+1) = \Gamma(d/2+m)(d/2+m)$.

□

Finaly we can state and prove the main result of this section.

Theorem 5.7 *Let Φ be an even conditionally positive definite function that possesses a positive Fourier transform $\widehat{\Phi} \in \mathcal{C}(\mathbb{R}^d \setminus 0)$. With the function*

$$\varphi_0(M) = \inf_{\|\omega\|_2 \le 2M} \widehat{\Phi}(\omega)$$

a lower bound on $\lambda_{\min}(A_{\Phi,X})$ is given by

$$\lambda_{\min}(A_{\Phi,X}) \ge \frac{\varphi_0(M)}{2\Gamma(d/2+1)} \left(\frac{M}{2^{3/2}}\right)^d$$

for any $M > 0$ satisfying

$$M \ge \frac{12}{q_X} \left(\frac{\pi\Gamma(d/2+1)^2}{9}\right)^{1/(d+1)}$$

or, a fortiori,

$$M \ge \frac{6.38d}{q_X}.$$

Proof
Let us define Ψ in equation (5.4) through its Fourier transform in the following way

$$\widehat{\Psi}(\omega) = \widehat{\Psi_M}(\omega) = \frac{\varphi_0(M)\Gamma(d/2+1)}{2^d M^d \pi^{d/2}} (\chi_M * \chi_M)(\omega).$$

We recall that

$$(\chi_M * \chi_M)(\omega) = \int_{\mathbb{R}^d} \chi_M(y)\chi_M(\omega - y) \mathrm{d}\, y = \int_{B(\omega,M)} \chi_M(y) \mathrm{d}\, y,$$

so if $\|\omega\|_2 > 2M$ and $y \in B(\omega, M)$ then $\|y\|_2 \ge \underbrace{\|\omega\|_2}_{\omega=\omega-y+y} - \|\omega - y\|_2 > 2M - M = M$.

We proved that $\widehat{\Psi}$ has support in $\overline{B(0,2M)}$ and that

$$0 \le (\chi_M * \chi_M)(\omega) \le \mathcal{L}(B(0,M)) \le \mathcal{L}(B(0,2M)),$$

where $\mathcal{L}$ is the Lebesgue measure on $\mathbb{R}^d$, because $0 \le \chi_M \le 1$.

For ω in the support of $\widehat{\Psi}$, i.e. $\|\omega\|_2 \le 2M$ we can write

$$\widehat{\Psi}(\omega) \le \frac{\varphi_0(M)\Gamma(d/2+1)}{2^d M^d \pi^{d/2}} \underbrace{\mathcal{L}(B(0,2M))}_{\frac{\pi^{d/2}}{\Gamma(d/2+1)}(2M)^d} = \varphi_0(M) \le \widehat{\Phi}(\omega).$$

It is time to write the radial function Ψ in an explicit form with Theorem 5.6:

$$\begin{aligned}\Psi_M(x) &= \frac{\varphi_0(M)\Gamma(d/2+1)}{2^d M^d \pi^{d/2}}(\chi_M * \chi_M)^\vee(x) = \\ &= \frac{\varphi_0(M)\Gamma(d/2+1)}{2^{d/2} M^d}|(\chi_M)^\vee|^2(x) = \\ &= \frac{\varphi_0(M)\Gamma(d/2+1)}{2^{d/2} M^d} M^d \|x\|_2^{-d} J_{\frac{d}{2}}(M\|x\|_2)^2 = \\ &= \frac{\varphi_0(M)\Gamma(d/2+1)}{2^{d/2}} \|x\|_2^{-d} J_{\frac{d}{2}}(M\|x\|_2)^2.\end{aligned}$$

We call $B \in M_N(\mathbb{R})$ the matrix whose components $B_{jk} = \Psi_M(x_j - x_k)$ for $j, k = 1, \ldots, N$. By Gershgorin's Theorem [43] there exists $j \in \{1, \ldots, N\}$ such that

$$\begin{aligned}\Psi_M(0) - \lambda_{\min}(B) \le |\Psi_M(0) - \lambda_{\min}(B)| \le \\ \le \sum_{\substack{k=1\\k\neq j}}^{N} |\Psi_M(x_j - x_k)| \le \max_{1\le j\le N} \sum_{\substack{k=1\\k\neq j}}^{N} |\Psi_M(x_j - x_k)|.\end{aligned}$$

We can obtain for $\alpha \in \mathbb{R}^N$ that

$$\frac{\alpha^\top B \alpha}{\alpha^\top \alpha} \ge \lambda_{\min}(B) \ge \Psi_M(0) - \max_{1\le j\le N} \sum_{\substack{k=1\\k\neq j}}^{N} |\Psi_M(x_j - x_k)|,$$

that is

$$\sum_{j,k=1}^{N} \alpha_j \alpha_k \Psi_M(x_j - x_k) \ge \|\alpha\|_2^2 \left(\Psi_M(0) - \max_{1\le j\le N} \sum_{\substack{k=1\\k\neq j}}^{N} |\Psi_M(x_j - x_k)| \right).$$

Since

$$\lim_{r\to 0^+} r^{-d} J_{d/2}^2(r) = \frac{1}{2^d \Gamma(d/2+1)^2},$$

then

$$\Psi_M(0) = \frac{\varphi_0(M)}{\Gamma(d/2+1)2^{d/2}} \left(\frac{M}{2}\right)^d = \frac{\varphi_0(M)}{\Gamma(d/2+1)} \left(\frac{M}{2^{3/2}}\right)^d.$$

If we prove that

$$\max_{1\le j\le N} \sum_{\substack{k=1\\k\neq j}}^{N} |\Psi_M(x_j - x_k)| \le \frac{1}{2}\Psi_M(0) \tag{5.5}$$

then

$$\sum_{j,k=1}^{N} \alpha_j \alpha_k \Psi_M(x_j - x_k) \ge \|\alpha\|_2^2 \frac{1}{2} \Psi_M(0),$$

that is the thesis in our statement.

Without loss of generality we can suppose that the maximum in equation (5.5) is reached in $x_1 = 0$, i.e.

$$\max_{1\le j\le N}\sum_{\substack{k=1\\k\neq j}}^{N}|\Psi_M(x_j-x_k)| = \sum_{k=2}^{N}|\Psi_M(x_k)|.$$

We define a sequence of disjoint sets that covers $\mathbb{R}^d$ as

$$E_n = \{x\in\mathbb{R}^d : nq_X \le \|x\|_2 < (n+1)q_X\}.$$

for $n\in\mathbb{N}$. Each x_j for $j = 2,\ldots,N$ is contained in exactly one E_n with $n\ge 1$, because $x_j\notin B(0,q_X) = B(x_1,q_X)$ for $j = 2,\ldots,N$ since $\{B(x_j,q_X)\}_{j=1,\ldots,N}$ are pairwise disjoint.

We claim that if $x_j\in E_n$ then

$$B(x_j,q_X)\subseteq\{x\in\mathbb{R}^d : (n-1)q_X\le\|x\|_2<(n+2)q_X\}.$$

If $y\in B(x_j,q_X)$ then $\|y\|_2\le\|y-x_j\|_2+\|x_j\|_2 < q_X+(n+1)q_X$ and $nq_X - q_X\le \|x_j\|_2-\|x_j-y\|_2\le\|y\|_2$. We have the following inequality with the Lebesgue measure $\mathcal{L}$

$$\#\{x_j\in E_n\}\mathcal{L}(B(0,q_X))\le\mathcal{L}(B(0,(n+2)q_X))-\mathcal{L}(B(0,(n-1)q_X)),$$

that let us to achieve

$$\#\{x_j\in E_n\}\le(n+2)^d-(n-1)^d\le 3^d n^{d-1}.$$

The last inequality holds for induction on d:

For $d = 1$ we have $n+2-n+1 = 3$. Let us suppose that it is true for d, we will prove the inequality for $d+1$.

$$\begin{aligned}
&(n+2)^{d+1}-(n-1)^{d+1} = (n+2)(n+2)^d-(n-1)(n-1)^d =\\
=&n((n+2)^d-(n-1)^d)+2(n+2)^d+(n-1)^d =\\
=&n((n+2)^d-(n-1)^d)+2((n+2)^d-(n-1)^d)+3(n-1)^d\le\\
\le&3^dn^d+2\cdot 3^dn^{d-1}+3(n-1)^d\le 3^{d+1}n^d,
\end{aligned}$$

where the last inequality holds because by dividing for 3^dn^d we have

$$1+\frac{2}{n}+\frac{3}{3^d}\left(\frac{n-1}{n}\right)^d\le 3,$$

that for $n = 1$ becomes $3\le 3$, instead for $n\ge 2$ is

$$\frac{2}{n}+\frac{3}{3^d}\left(\frac{n-1}{n}\right)^d\le\frac{2}{n}+\left(\frac{n-1}{n}\right)^d\le\frac{2}{n}+1\le 1+1 = 2.$$

Since

$$J_{d/2}(r)^2 \le \frac{2^{d+2}}{r\pi} \qquad \text{for } r > 0 \text{ and } d \in \mathbb{N},$$

we have

$$\begin{aligned}
|\Psi_M(x)| &\le \frac{\varphi_0(M)\Gamma(d/2+1)}{2^{d/2}}\|x\|_2^{-d}\frac{2^{d+2}}{M\|x\|_2\pi} = \\
&= \frac{\varphi_0(M)\Gamma(d/2+1)2^{d/2+2}}{M\pi}\|x\|_2^{-(d+1)} = \\
&= \frac{\varphi_0(M)}{\Gamma(d/2+1)}\left(\frac{M}{2^{3/2}}\right)^d\frac{\Gamma(d/2+1)^2}{\pi}4\left(\frac{1}{M\|x\|_2}\right)^{d+1}2^{(1/2+3/2)d} = \\
&= \frac{\varphi_0(M)}{\Gamma(d/2+1)}\left(\frac{M}{2^{3/2}}\right)^d\frac{\Gamma(d/2+1)^2}{\pi}\left(\frac{4}{M\|x\|_2}\right)^{d+1} = \\
&= \Psi_M(0)\frac{\Gamma(d/2+1)^2}{\pi}\left(\frac{4}{M\|x\|_2}\right)^{d+1}.
\end{aligned}$$

If $x \in E_n$ with $n \ge 1$ then $nq_X \le \|x\|_2$ gives us

$$|\Psi_M(x)| \le \Psi_M(0)\frac{\Gamma(d/2+1)^2}{\pi}\left(\frac{4}{Mnq_X}\right)^{d+1},$$

that implies with $\sum_{n=1}^{\infty} n^{-2} = \pi^2/6$

$$\begin{aligned}
\sum_{k=2}^{N}|\Psi_M(x_k)| &\le \sum_{n=1}^{+\infty}\#\{x_j \in E_n\}\sup_{x\in E_n}|\Psi_M(x)| \le \\
&\le \sum_{n=1}^{+\infty}3^d n^{d-1}\Psi_M(0)\frac{\Gamma(d/2+1)^2}{\pi}\left(\frac{4}{Mnq_X}\right)^{d+1} \le \\
&\le \sum_{n=1}^{+\infty}n^{-2}\Psi_M(0)\frac{\Gamma(d/2+1)^2}{3\pi}\left(\frac{12}{Mq_X}\right)^{d+1} \le \\
&\le \Psi_M(0)\frac{\pi\Gamma(d/2+1)^2}{18}\left(\frac{12}{Mq_X}\right)^{d+1} \le \\
&\le \frac{1}{2}\Psi_M(0),
\end{aligned}$$

the last inequality holds because

$$\frac{\pi\Gamma(d/2+1)^2}{18}\left(\frac{12}{Mq_X}\right)^{d+1} \le \frac{1}{2} \Leftrightarrow (Mq_X)^{d+1} \ge \frac{\pi\Gamma(d/2+1)^2}{9}(12)^{d+1}$$

is our hypothesis on M.

We can conclude with Stirling's formula

$$1 \le \frac{\Gamma(x+1)}{\sqrt{2\pi x}x^x e^{-x}} \le e^{\frac{1}{12x}},$$

which gives us

$$\frac{\pi}{9}\Gamma\left(\frac{d}{2}+1\right)^2 \leq \frac{\pi}{9}2\pi\frac{d}{2}\left(\frac{d}{2}\right)^d e^{-d}e^{\frac{1}{3d}} \Leftrightarrow \frac{\pi}{9}\Gamma\left(\frac{d}{2}+1\right)^2 \leq \frac{\pi^2}{9}d^{d+1}(2e)^{-d}e^{\frac{1}{3d}}.$$

If we apply $(\cdot)^{\frac{1}{d+1}}$ to the inequality we get

$$\left(\frac{\pi}{9}\Gamma\left(\frac{d}{2}+1\right)^2\right)^{\frac{1}{d+1}} \leq d\left(\frac{\pi^2}{9}\right)^{\frac{1}{d+1}}(2e)^{-\frac{d}{d+1}}e^{\frac{1}{3d(d+1)}}$$

Since $d \geq 1$ we can write $\frac{1}{d+1} \leq \frac{1}{2} \leq \frac{d}{d+1}$, $\frac{\pi^2}{9} > 1$, $2e > 1$ and $6 \leq (3d(d+1))$ from which we obtain

$$\left(\frac{\pi^2}{9}\right)^{\frac{1}{d+1}} \leq \left(\frac{\pi^2}{9}\right)^{\frac{1}{2}} = \frac{\pi}{3}, \quad \sqrt{2e} \leq (2e)^{\frac{d}{d+1}} \quad \text{and} \quad e^{\frac{1}{3d(d+1)}} \leq e^{\frac{1}{6}}$$

that gives us

$$\frac{12}{q_X}\left(\frac{\pi}{9}\Gamma\left(\frac{d}{2}+1\right)^2\right)^{\frac{1}{d+1}} \leq \frac{d}{q_X}12\frac{\pi}{3\sqrt{2e}}e^{\frac{1}{6}} \leq 6.38\frac{d}{q_X}.$$

□

If we define some constants that will depend on the space dimension d, Theorem 5.7 can be read as

$$\lambda_{\min}(A_{\Phi,X}) \geq C_d\varphi_0\left(\frac{M_d}{q_X}\right)\frac{1}{q_X^d}$$

with

$$M_d = 12\left(\frac{\pi\Gamma(d/2+1)^2}{9}\right)^{1/(d+1)} \quad \text{and} \quad C_d = \frac{1}{2\Gamma(d/2+1)}\left(\frac{M_d}{2^{3/2}}\right)^d,$$

but M_d can also be

$$M_d = 6.38d.$$

Let us apply Theorem 5.7 to a conditionally positive definite function that satisfies the hypothesis of Theorem 1.26. In the proof of Theorem 1.26 we proved that

$$\widehat{\Phi_\delta}(\omega) \geq \frac{c_1\delta^d}{(1+\delta^2\|\omega\|_2^2)^s},$$

so

$$\varphi_0^\delta\left(\frac{M_d}{q_X}\right)\frac{1}{q_X^d} \geq \frac{c_1\left(\frac{\delta}{q_X}\right)^d}{\left(1+\delta^2\frac{4M_d^2}{q_X^2}\right)^s} = \frac{c_1\left(\frac{\delta}{q_X}\right)^d}{\left(1+\overline{M_d}\left(\frac{\delta}{q_X}\right)^2\right)^s}$$

with $\overline{M_d} = 4M_d^2$. If we impose $K_d = \overline{M_d}+1$ then since $(\overline{M_d}+1)x^2 \geq 1+\overline{M_d}x^2$ for $|x| \geq 1$ we obtain

$$\varphi_0^\delta\left(\frac{M_d}{q_X}\right)\frac{1}{q_X^d} \geq \frac{c_1}{K_d^s}\left(\frac{\delta}{q_X}\right)^{d-2s}$$

if $\delta \geq q_X$. We can conclude

$$\lambda_{\min}(A_{\Phi_\delta,X}) \geq \frac{C_d c_1}{K_d^s}\left(\frac{\delta}{q_X}\right)^{d-2s}, \tag{5.6}$$

where the constants involved depend only on the space dimension d and on Φ, in particular they do not depend on δ. We note that if Φ has compact support contained in $\overline{B(0,1)}$ then Φ_δ has support in $\overline{B(0,\delta)}$. Under the condition $\delta < q_X$ then, since $B(x_i, q_X) \cap B(x_j, q_X) = \emptyset$ for $i \neq j$, the matrix $A_{\Phi_\delta,X}$ is diagonal with $\lambda_{\min}(A_{\Phi_\delta,X}) = \Phi(0)$.

Chapter 6

Optimal recovery

The importance of radial basis function methods for interpolation arise quite naturally from Mairhuber-Curtis theorem because with RBFs the interpolation space depends on the interpolation points.

Definition 6.1 *Supose that $\Omega \subseteq \mathbb{R}^d$ contains at least N points. Let $V \subseteq \mathcal{C}(\Omega)$ be an N-dimensional linear space. Then V is called Haar space of dimension N on Ω if for arbitrary distinct points $x_1, \ldots, x_N \in \Omega$ and arbitrary $f_1, \ldots, f_N \in \mathbb{R}$ there exists exactly one function $s \in V$ with $s(x_i) = f_i$ for $i = 1, \ldots, N$.*

Theorem 6.1 *(Mairhuber-Curtis [44, 45]) Suppose that $\Omega \subseteq \mathbb{R}^d, d \geq 2$, contains an interior point. Then there exists no Haar space on Ω of dimension $N \geq 2$.*

Moreover, we studied in Theorem 1.10 that positive definite kernels appear when we deal with reproducing-kernel Hilbert spaces. We will show that interpolants built with conditionally positive definite functions are optimal. The goal of this section is to explain the word "optimal". In equation (4.3) we defined the interpolation space

$$V_X = \left\{ \sum_{j=1}^{N} \alpha_j \Phi(\cdot, x_j) + p : p \in \mathcal{P} \text{ and } \sum_{j=1}^{Q} \alpha_j q(x_j) = 0 \text{ for all } q \in \mathcal{P} \right\}.$$

Theorem 6.2 *Suppose that $\Phi \in \mathcal{C}(\Omega \times \Omega)$ is a conditionally positive definite kernel with respect to the finite-dimensional space $\mathcal{P} \subseteq \mathcal{C}(\Omega)$. Suppose that $X = \{x_1, \ldots, x_N\}$ is $\mathcal{P}$-unisolvent and that $f \in \mathcal{N}_\Phi(\Omega)$. Then the interpolant $s_{f,X}$ is the best approximation to f in V_X with respect to the native space (semi-)norm, i.e.*

$$|f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \leq |f - s|_{\mathcal{N}_\Phi(\Omega)}$$

for each $s \in V_X$. Hence, $s_{f,X}$ is the orthogonal projection of f onto V_X (Definition 1.2).

Proof
From Theorem 2.8 we get

$$\begin{aligned}|f-s|^2_{\mathcal{N}_\Phi(\Omega)} &= |f-s_{f,X}+s_{f,X}-s|^2_{\mathcal{N}_\Phi(\Omega)} = \\ &= |f-s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} + |\underbrace{s_{f,X}-s}_{\in V_X}|^2_{\mathcal{N}_\Phi(\Omega)} \geq |f-s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)}.\end{aligned}$$

□

Note that for conditionally positive definite function if we add to $s_{f,X}$ and element of $\mathcal{P}$ then

$$|f-s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} = |f-s_{f,X}-p|_{\mathcal{N}_\Phi(\Omega)} \quad \text{for each } p \in \mathcal{P},$$

because $\mathcal{P}$ is the null space of $\langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega)}$. We can avoid the non-uniqueness of the orthogonal projection by changing the inner-product (Theorem 2.5). The interpolant is not only an orthogonal projection, it minimizes also the native space (semi-)norm.

Theorem 6.3 *Suppose that $\Phi \in \mathcal{C}(\Omega\times\Omega)$ is a conditionally positive definite kernel with respect to the finite-dimensional space $\mathcal{P} \subseteq \mathcal{C}(\Omega)$ and that $X=\{x_1,\dots,x_N\}$ is $\mathcal{P}$-unisolvent. Fix $f_1,\dots,f_N \in \mathbb{R}$ then the interpolant $s_{f,X}$ has minimal (semi-)norm $|\cdot|_{\mathcal{N}_\Phi(\Omega)}$ under all functions $s \in \mathcal{N}_\Phi(\Omega)$ that interpolates $\{f_1,\dots,f_N\}$ on X, i.e.*

$$|s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} = \min\{|s|_{\mathcal{N}_\Phi(\Omega)} : s \in \mathcal{N}_\Phi(\Omega) \text{ such that } s(x_j)=f_j \text{ for } j=1,\dots,N\}.$$

Proof
The iterpolant $s_{f,X} \in V_X$, so it admits the form $s_{f,X} = \lambda^x(\Phi(\cdot,x))+q$ with $\lambda = \sum_{j=1}^N \alpha_j\delta_{x_j} \in L_{\mathcal{P}}(\Omega)$ and $q \in \mathcal{P}$ (equation 2.12).

From equation 2.17 we obtain for $v \in \mathcal{N}_\Phi(\Omega)$ such that $v(x_j)=0$ for $j=1,\dots,N$

$$\begin{aligned}0 &= \sum_{j=1}^N \alpha_j v(x_j) = \lambda(v) = \left\langle v, \sum_{j=1}^N \alpha_j\Phi(\cdot,x_j)\right\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \langle v, \lambda^x(\Phi(\cdot,x))\rangle_{\mathcal{N}_\Phi(\Omega)} = \langle v, \lambda^x(\Phi(\cdot,x))+q\rangle_{\mathcal{N}_\Phi(\Omega)} - \underbrace{\langle v,q\rangle_{\mathcal{N}_\Phi(\Omega)}}_{=0} = \\ &= \langle v, s_{f,X}\rangle_{\mathcal{N}_\Phi(\Omega)}.\end{aligned}$$

The chain of inequality holds because $q \in \mathcal{P}$ is in the null space of $\langle\cdot,\cdot\rangle_{\mathcal{N}_\Phi(\Omega)}$.

We can conclude with Pythagorean theorem, indeed

$$|s|^2_{\mathcal{N}_\Phi(\Omega)} = |s-s_{f,X}+s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} = |\underbrace{s-s_{f,X}}_{(s-s_{f,X})_{|X}=0}|^2_{\mathcal{N}_\Phi(\Omega)} + |s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)} \geq |s_{f,X}|^2_{\mathcal{N}_\Phi(\Omega)},$$

for every $s \in \mathcal{N}_\Phi(\Omega)$ such that $s(x_j)=f_j$ for $j=1,\dots,N$.

□

The last minimal property for radial basis function interpolant involves the cardianal function of Theorem 4.1. By equation (4.5) we can write the interpolat in cardinal form

$$s_{f,X}(x) = \sum_{j=1}^{N} f(x_j) u_j^*(x).$$

Theorem 6.4 *Suppose that $\Phi \in \mathcal{C}^{2k}(\Omega \times \Omega)$ is a conditionally positive definite kernel with respect to the finite-dimensional space $\mathcal{P} \subseteq \mathcal{C}^k(\Omega)$ and that $X = \{x_1, \ldots, x_N\}$ is $\mathcal{P}$-unisolvent. Fix $x \in \Omega$ and $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ then $D^\alpha(u^*)(x)$ is the solution of the minimization problem*

$$\inf_{\substack{u \in \mathbb{R}^N \\ \sum_{j=1}^N u_j p(x_j) = D^\alpha p(x)}} \sup_{\substack{f \in \mathcal{N}_\Phi(\Omega) \\ |f|_{\mathcal{N}_\Phi(\Omega)} = 1}} \left| D^\alpha(f)(x) - \sum_{j=1}^{N} u_j f(x_j) \right|.$$

Proof
Fix $f \in \mathcal{N}_\Phi(\Omega)$. If we apply to $D^\alpha(f)(x)$ Theorem 2.17 and to $\{f(x_j)\}_{j=1,\ldots,N}$ Theorem 2.3 then

$$\begin{aligned} D^\alpha(f)(x) - \sum_{j=1}^{N} u_j f(x_j) &= D^\alpha(\Pi_{\mathcal{P}}(f))(x) + \langle f, D_2^\alpha(G(\cdot,x))\rangle_{\mathcal{N}_\Phi(\Omega)} + \\ &\quad - \underbrace{\sum_{j=1}^{N} u_j \Pi_{\mathcal{P}}(f)(x_j)}_{D^\alpha(\Pi_{\mathcal{P}}(f))(x)} - \sum_{j=1}^{N} u_j \langle f, G(\cdot, x_j)\rangle_{\mathcal{N}_\Phi(\Omega)} = \\ &= \left\langle f, D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} u_j G(\cdot, x_j) \right\rangle_{\mathcal{N}_\Phi(\Omega)}, \end{aligned}$$

from which we obtain

$$\left| D^\alpha(f)(x) - \sum_{j=1}^{N} u_j f(x_j) \right| \leq |f|_{\mathcal{N}_\Phi(\Omega)} \left| D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} u_j G(\cdot, x_j) \right|_{\mathcal{N}_\Phi(\Omega)}.$$

We can conclude with Theorem 4.2 and Theorem 4.4

$$\sup_{\substack{f \in \mathcal{N}_\Phi(\Omega) \\ |f|_{\mathcal{N}_\Phi(\Omega)} = 1}} \left| D^\alpha(f)(x) - \sum_{j=1}^{N} u_j f(x_j) \right| = \left| D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} u_j G(\cdot, x_j) \right|_{\mathcal{N}_\Phi(\Omega)} = \sqrt{\mathcal{Q}(u)}.$$

The first equality holds if we choose

$$f(\cdot) = \frac{D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} u_j G(\cdot, x_j)}{\left| D_2^\alpha(G(\cdot,x)) - \sum_{j=1}^{N} u_j G(\cdot, x_j) \right|_{\mathcal{N}_\Phi(\Omega)}}.$$

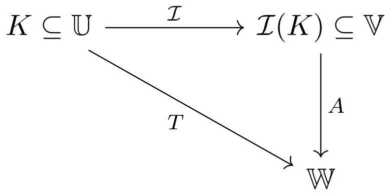
□

If we consider optimality from an abstract point of view [46, 47], we can show that interpolation with radial basis functions is an optimal algorithm.

Let $\mathbb{U}, \mathbb{V}$ and $\mathbb{W}$ be three normed linear spaces. Let $K \subseteq \mathbb{U}$. We assume that we have some information on the element of K through the linear mapping $\mathcal{I} : \mathbb{U} \to \mathbb{V}$, called information operator. We have another linear operator $T : \mathbb{U} \to \mathbb{W}$, called target operator. Our goal is to reconstruct for each $x \in K$ the target $T(x) \in \mathbb{W}$ from the information $\mathcal{I}(x) \in \mathbb{V}$. In a more mathematical way, we want to find a map, called algorithm, $A : \mathcal{I}(K) \subseteq \mathbb{V} \to \mathbb{W}$ that minimizes

$$\inf_{A} \sup_{x \in K} \|A(\mathcal{I}(x)) - T(x)\|_{\mathbb{W}}. \tag{6.1}$$

If A^* minimizes the quantity in equation (6.1) then A^* is called optimal algorithm (it can be not linear).



In this general setting we can give a sufficient condition on an algorithm $A : \mathcal{I}(K) \subseteq \mathbb{V} \to \mathbb{W}$ to be optimal.

Theorem 6.5 *Suppose that K is symmetric, i.e. for each $x \in K$ then $-x \in K$. If there exists a map $F : \mathcal{I}(K) \to \mathbb{U}$ such that for $x \in K$*

- $x - F(\mathcal{I}(x)) \in K$
- $\mathcal{I}(x - F(\mathcal{I}(x))) = 0$

then $T \circ F : \mathcal{I}(K) \to \mathbb{W}$ is an optimal algorithm.

Proof
Fix $x \in K$ such that $\mathcal{I}(x) = 0$, then using the linearity $\mathcal{I}(-x) = -\mathcal{I}(x) = 0$ with $-x \in K$, by symmetry of K. With such $x \in K$ we have, for an arbitrary algorithm $A : \mathcal{I}(K) \subseteq \mathbb{V} \to \mathbb{W}$,

$$\begin{aligned}
\|T(x)\|_{\mathbb{W}} &= \frac{1}{2}\|T(x) - A(0) + T(x) + A(0)\|_{\mathbb{W}} \leq \\
&\leq \frac{1}{2}(\|T(x) - A(0)\|_{\mathbb{W}} + \|T(x) + A(0)\|_{\mathbb{W}}) \leq \\
&\leq \max\{\|T(x) - A(0)\|_{\mathbb{W}}, \|T(x) + A(0)\|_{\mathbb{W}}\} = \\
&= \max\{\|T(x) - A(\mathcal{I}(x))\|_{\mathbb{W}}, \|T(x) + A(\mathcal{I}(-x))\|_{\mathbb{W}}\} \overset{T \text{ linear}}{=} \\
&= \max\{\|T(x) - A(\mathcal{I}(x))\|_{\mathbb{W}}, \| - T(-x) + A(\mathcal{I}(-x))\|_{\mathbb{W}}\} \leq \\
&\leq \sup_{y \in K} \|A(\mathcal{I}(y)) - T(y)\|_{\mathbb{W}}.
\end{aligned}$$

We proved that

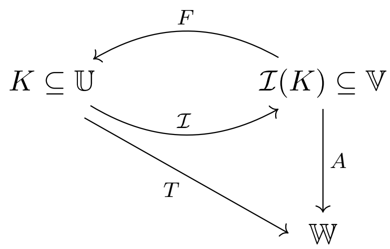
$$\sup\{\|T(x)\|_{\mathbb{W}} : x \in K, \mathcal{I}(x) = 0\} \le \inf_A \sup_{y\in K} \|A(\mathcal{I}(y)) - T(y)\|_{\mathbb{W}}.$$

We compute, with linearity of T,

$$\sup_{y\in K} \|T \circ F(\mathcal{I}(y)) - T(y)\|_{\mathbb{W}} = \sup_{y\in K} \|T(F(\mathcal{I}(y)) - y)\|_{\mathbb{W}} \le \sup\{\|T(x)\|_{\mathbb{W}} : x \in K, \mathcal{I}(x) = 0\}$$

by symmetry of K and hypothesis.

□



Theorem 6.5 seems difficult to apply because $T \circ F$ involves the target operator, which is unknown.

We give some examples of application of Theorem 6.5 which are useful for our purposes.

We fix $\mathbb{U}$ and $\mathbb{W}$ to be the native space $\mathcal{N}_\Phi(\Omega)$ of a conditionally positive definite kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ with respect to $\mathcal{P}$. The target map $T : \mathbb{U} \to \mathbb{W}$ is the identity and the information operator $\mathcal{I} : \mathbb{U} \to \mathbb{V} = \mathbb{R}^N$ is given by $f \longmapsto f_{|X}$ with $X = \{x_1, \dots, x_N\} \subseteq \Omega$ a $\mathcal{P}$-unisolvent set of distinct points. For us K is equals to $B_{\mathcal{N}_\Phi(\Omega)}(0, 1)$.

Theorem 6.6 *Among all maps* $A : \mathcal{N}_\Phi(\Omega)_{|X} \to \mathcal{N}_\Phi(\Omega)$, *interpolation on* X *is optimal, i.e. it minimizes*

$$\inf_A \sup_{f\in B_{\mathcal{N}_\Phi(\Omega)}(0,1)} \|A(f_{|X}) - f\|_{\mathcal{N}_\Phi(\Omega)}.$$

Proof
To apply Theorem 6.5 we define as $F : \mathbb{R}^N \to \mathcal{N}_\Phi(\Omega)$ by $f = (f_1, \dots, f_N) \longmapsto s_{f,X}$. With this definition we have that

$$\mathcal{I}(f) = (f(x_1), \dots, f(x_N))^\top = \mathcal{I}(s_{f,X}) = \mathcal{I}(F((f(x_1), \dots, f(x_N))^\top)) = \mathcal{I}(F(\mathcal{I}(f)))$$

and

$$|f - F(\mathcal{I}(f))|_{\mathcal{N}_\Phi(\Omega)} = |f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \overset{\text{Theorem 2.9}}{\le} |f|_{\mathcal{N}_\Phi(\Omega)} \le 1,$$

for $f \in B_{\mathcal{N}_\Phi(\Omega)}(0, 1)$. We can finish by showing that $T \circ F$ is the interpolation operator:

$$T(F(f)) = T(s_{f,X}) = s_{f,X}.$$

□

By keeping the same notation we defined above and changing $\mathbb{W}$ to be $\mathbb{R}$ and $T : \mathcal{N}_\Phi(\Omega) \to \mathbb{R}$ to be the point evaluation functional at $x \in \Omega$ we can state the following:

Theorem 6.7 *Among all maps $A : \mathcal{N}_\Phi(\Omega)_{|X} \to \mathbb{R}$, the point evaluation at $x \in \Omega$ of the interpolant on X is optimal, i.e. it minimizes*

$$\inf_A \sup_{f \in B_{\mathcal{N}_\Phi(\Omega)}(0,1)} |A(f_{|X}) - f(x)|.$$

Proof
The proof is equivalent to the proof of Theorem 6.6. We can conclude the proof by showing that $T \circ F$ is the evaluation at $x \in \Omega$ of the interpolant:

$$T(F(f)) = T(s_{f,X}) = s_{f,X}(x).$$

□

Chapter 7

Analyzing the convergence of RL-RBF method

The purpose of this section is to generalize the findings of [2] and analyze the proofs in this new context.

7.1 Quasi-uniform Shepard method

The key points of the proofs in [2] suggest this more general definition.

Definition 7.1 *Fix a decreasing function $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $\lim_{n\to+\infty} \frac{\varphi(n+1)}{\varphi(n)}$ exists and it is strictly smaller than 1. A process that defines, for every quasi-uniform set $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ with respect to c_{qu}, a family of functions $u_j = u_j^X : \Omega \to \mathbb{R}$ for $1 \leq j \leq N$ is a quasi-uniform Shepard method with respect to φ if there exists constants $h_0, C, c > 0$ such that*

- $\sum_{j=1}^N u_j(x) \geq c$ *for all* $x \in \Omega$,
- $|u_j(x)| \leq C\varphi\left(\frac{\|x-x_j\|_2}{q_X}\right)$ *for all* $x \in \Omega$ *and* $j = 1, \ldots, N$,

provided $h_{X,\Omega} \leq h_0$.

The process in the Definition 7.1 naturally defines a quasi-interpolation scheme for $f \in \mathcal{C}(\Omega)$ with constant reproduction as

$$z_{f,X}(x) = \frac{\sum_{j=1}^N f(x_j)u_j(x)}{\sum_{j=1}^N u_j(x)}. \tag{7.1}$$

The quasi-interpolant is well-defined because of the condition on the denominator stated in Definition 7.1 and if $k \in \mathbb{R}$ then $z_{k,X} = k$. The process is interpolatory if the functions $\{u_j\}_{j=1,\ldots,N}$ are cardinal, i.e. $u_j(x_i) = \delta_{i,j}$ for $i, j = 1, \ldots, N$.

We claim that this quasi-interpolation process admits linear convergence with respect to $h_{X,\Omega}$, under suitable assumption on the domain Ω, as Shepard method with compactly supported basis functions (Theorem 3.1,Theorem 3.7).

The following result proves the stability of the method and it will be useful to prove convergence.

Theorem 7.1 *Suppose that $X = \{x_1, \dots, x_N\} \subseteq \mathbb{R}^d$ is a quasi-uniform data set with respect to c_{qu} and $\{u_1, \dots, u_N\}$ is given by a quasi-uniform Shepard process with respect to φ (Definition 7.1) then for every $\ell \in \mathbb{N}$ there exists a constant $K = K(\ell, C, \varphi, d)$ such that*

$$\sum_{j=1}^{N} \|x - x_j\|_2^\ell |u_j(x)| \leq K h_{X,\Omega}^\ell \quad \text{for } x \in \Omega.$$

Proof
Fix $x \in \Omega$. We define a sequence of sets $\{E_n\}_{n\in\mathbb{N}}$ that covers $\mathbb{R}^d$ as

$$E_n = \{y \in \mathbb{R}^d : nq_X \leq \|y - x\|_2 \leq (n+1)q_X\}.$$

We note that if $x_j \in E_n$ then

$$B(x_j, q_X) \subseteq \{y \in \mathbb{R}^d : (n-1)q_X \leq \|y - x\|_2 \leq (n+2)q_X\},$$

because if $y \in B(x_j, q_X)$ then $nq_X - q_X \leq \|x_j - x\|_2 - \|x_j - y\|_2 \leq \|y - x\|_2 \leq \|y - x_j\|_2 + \|x_j - x\| \leq q_X + (n+1)q_X$. Since $\{B(x_j, q_X)\}_{j=1,\dots,N}$ are pairwise disjoint

$$\bigcup_{j:x_j\in E_n} B(x_j, q_X) \subseteq \overline{B(x, (n+2)q_X)} \setminus B(x, (n-1)q_X)$$

a volume comparison as in Theorem 5.7 gives for $n \geq 1$

$$\#\{x_j : x_j \in E_n\} \leq (n+2)^d - (n-1)^d \leq 3^d n^{d-1} \leq 3^d (n+1)^{d-1},$$

that holds also for $n = 0$ because $2^d \leq 3^d$.

By remarking that if $x_j \in E_n$ then $n \leq \frac{\|x_j - x\|_2}{q_X}$ and if we split the sum over the sets $\{E_n\}_{n\in\mathbb{N}}$ we obtain

$$\begin{aligned}
\sum_{j=1}^{N} \|x - x_j\|_2^\ell |u_j(x)| &\leq \sum_{n=0}^{+\infty} \sum_{x_j\in E_n} \|x - x_j\|_2^\ell |u_j(x)| \leq \sum_{n=0}^{+\infty} \sum_{x_j\in E_n} \|x - x_j\|_2^\ell C\varphi\left(\frac{\|x - x_j\|_2}{q_X}\right) \leq \\
&\leq \sum_{n=0}^{+\infty} \sum_{x_j\in E_n} \|x - x_j\|_2^\ell C\varphi(n) \leq \sum_{n=0}^{+\infty} \sum_{x_j\in E_n} (n+1)^\ell q_X^\ell C\varphi(n) \leq \\
&\leq \sum_{n=0}^{+\infty} 3^d C (n+1)^{d+\ell-1} q_X^\ell \varphi(n) \leq h_{X,\Omega}^\ell \underbrace{3^d C \sum_{n=0}^{+\infty} (n+1)^{d+\ell-1} \varphi(n)}_{K}.
\end{aligned}$$

The series is convergent because of the ratio test:

$$\frac{(n+2)^{d+\ell-1}\varphi(n+1)}{(n+1)^{d+\ell-1}\varphi(n)} = \left(\frac{n+2}{n+1}\right)^{d+\ell-1}\frac{\varphi(n+1)}{\varphi(n)} \xrightarrow{n\to+\infty} \lim_{n\to+\infty}\frac{\varphi(n+1)}{\varphi(n)} < 1.$$

□

We note that the hypothesis of X to be quasi-uniform is not necessary in this proof, but it will be fundamental for a convergence result. By Theorem 7.1 we remark that the Lebesgue function of the quasi-interpolation scheme in equation (7.1) is uniformly bounded by

$$\sum_{j=1}^{N}\frac{|u_j(x)|}{|\sum_{i=1}^{N}u_i(x)|} \leq \frac{3^d C}{c}\sum_{n=0}^{+\infty}(n+1)^{d-1}\varphi(n),$$

which show that the scheme is stable and by tuning the parameters of the method C, c and φ it is possible to reach a better stability.

Before proceeding to the proof of convergence we have to add some hypotheses on the domain Ω.

Definition 7.2 *Let Ω be a subset of $\mathbb{R}^d$ and $f : \Omega \to \mathbb{R}$ then we call modulus of continuity of f the following function*

$$\omega_f(\delta) = \sup\{|f(x) - f(y)| : x, y \in \Omega, \|x - y\|_2 \leq \delta\}.$$

Theorem 7.2 *[48] Let $\Omega \subseteq \mathbb{R}^d$ and suppose that there exists $\gamma \geq 1$ such that any two points $x, y \in \Omega$ can be joined with a rectifiable curve $\Gamma \subseteq \Omega$ with length $|\Gamma| \leq \gamma\|x - y\|_2$, then if $0 < \varepsilon < \delta$*

$$\omega_f(\delta) \leq 2\lceil\gamma\rceil\frac{\delta}{\varepsilon}\omega_f(\varepsilon).$$

Proof

Since $\{]m\varepsilon, (m+1)\varepsilon]\}_{m\in\mathbb{N}}$ covers $\mathbb{R}_{>0}$ then there exists $m \in \mathbb{N}\setminus\{0\}$ such that $m\varepsilon < \delta \leq (m+1)\varepsilon$ ($m = 0$ implies $\delta \leq \varepsilon$, that is a contradiction). If we prove that $\omega_f((m+1)\varepsilon) \leq \lceil\gamma\rceil(m+1)\omega_f(\varepsilon)$ then we can conclude because

$$\omega_f(\delta) \leq \omega_f((m+1)\varepsilon) \leq \lceil\gamma\rceil(m+1)\omega_f(\varepsilon) \leq 2\lceil\gamma\rceil\frac{\delta}{\varepsilon}\omega_f(\varepsilon),$$

the last inequality hold because

$$\varepsilon(m+1) \leq \varepsilon 2m \leq 2\delta.$$

Let us fix $x, y \in \Omega$ such that $\|x-y\|_2 \leq (m+1)\varepsilon$ and a rectifiable curve Γ in Ω joining x and y with $|\Gamma| \leq \gamma\|x-y\|_2 \leq \lceil\gamma\rceil\|x-y\|_2$. We can parameterize Γ with a continuous function

$x : [0, |\Gamma|] \to \Omega$ such that the length of the curve $|x_{|[s,s']}|$ is $|s' - s|$ for $|\Gamma| \geq s' > s \geq 0$ and $x(0) = y, x(|\Gamma|) = x$. We define the following set of points $\{s_0, \ldots, s_{\lceil\gamma\rceil(m+1)}\} \subseteq [0, |\Gamma|]$ as

$$s_i = i\frac{|\Gamma|}{\lceil\gamma\rceil(m+1)} \qquad \text{for } i = 0, \ldots, \lceil\gamma\rceil(m+1).$$

We claim that $s_i - s_{i-1} \leq \varepsilon$ because

$$\frac{|\Gamma|}{\lceil\gamma\rceil(m+1)} \leq \frac{\lceil\gamma\rceil\|x-y\|_2}{\lceil\gamma\rceil(m+1)} \leq \varepsilon.$$

Finally,

$$|f(x) - f(y)| = |f(x(|\Gamma|)) - f(x(0))| \leq \sum_{i=1}^{\lceil\gamma\rceil(m+1)} |f(x(s_i)) - f(x(s_{i-1}))| \leq \lceil\gamma\rceil(m+1)\omega_f(\varepsilon),$$

where the last inequality holds because $\|x(s_i) - x(s_{i-1})\|_2 \leq |x_{|[s_{i-1},s_i]}| \leq |s_i - s_{i-1}| \leq \varepsilon$ for $i = 1, \ldots, \lceil\gamma\rceil(m+1)$.

□

Theorem 7.3 *Suppose that $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ is a set of distinct points and fix $x \in \Omega$ then*

$$\#\{x_j : \|x - x_j\|_2 \leq \delta\} \leq \left(1 + \frac{\delta}{q_X}\right)^d.$$

Proof
We claim that if $\|x_j - x\|_2 \leq \delta$ for a $j \in \{1, \ldots, N\}$ then $B(x_j, q_X) \subseteq B(x, q_X + \delta)$, because if $y \in B(x_j, q_X)$ then $\|y - x\|_2 \leq \|y - x_j\|_2 + \|x_j - x\| < q_X + \delta$. Since $\{B(x_j, q_X)\}_{j=1,\ldots,N}$ are pairwise disjoint a volume comparison gives us:

$$\bigcup_{j:\|x_j - x\|_2 \leq \delta} B(x_j, q_X) \subseteq B(x, q_X + \delta),$$

that implies

$$\#\{x_j : \|x - x_j\|_2 \leq \delta\} q_X^d \leq (q_X + \delta)^d.$$

□

Theorem 7.4 *Suppose that $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ is a quasi-uniform data set with respect to c_{qu} and $\{u_1, \ldots, u_N\}$ is given by a quasi-uniform Shepard process with respect to φ. Fix $r_0 > 0$ and define Ω^* to be the closure $\bigcup_{x\in\Omega} B(x, r_0)$. If $f \in \mathcal{C}^1(\Omega^*)$ and Ω satisfy the hypothesis of Theorem 7.2 then there exists a constant $K = K(C, c, \varphi, c_{qu}, d, \Omega)$ such that*

$$\|f - z_{f,X}\|_{L^\infty(\Omega)} \leq K h_{X,\Omega} \|f\|_{\mathcal{C}^1(\Omega^*)},$$

for $h_{X,\Omega} < \min\{r_0, h_0\}$.

Proof
By the constant reproduction property of the scheme in equation (7.1) we have for $x \in \Omega$

$$
\begin{aligned}
|f(x) - z_{f,X}(x)| = &\left| f(x) - \frac{\sum_{j=1}^N f(x_j)u_j(x)}{\sum_{j=1}^N u_j(x)} \right| = \left| \frac{\sum_{j=1}^N (f(x) - f(x_j))u_j(x)}{\sum_{j=1}^N u_j(x)} \right| \leq \\
\leq &\frac{1}{c}\sum_{j=1}^N |f(x) - f(x_j)||u_j(x)| \leq \\
\leq &\frac{1}{c}\sum_{j:\|x-x_j\|_2 \leq h_{X,\Omega}} |f(x) - f(x_j)||u_j(x)| + \frac{1}{c}\sum_{j:\|x-x_j\|_2 > h_{X,\Omega}} |f(x) - f(x_j)||u_j(x)| \leq \\
\leq &\frac{1}{c}\sum_{j:\|x-x_j\|_2 \leq h_{X,\Omega}} \omega_f(h_{X,\Omega}) \underbrace{|u_j(x)|}_{\leq C\varphi(0)} + \frac{1}{c}\sum_{j:\|x-x_j\|_2 > h_{X,\Omega}} \omega_f(\|x - x_j\|_2)|u_j(x)|.
\end{aligned}
$$

If we apply Theorem 7.3 to the first sum and Theorem 7.2 to the second sum we have

$$
\begin{aligned}
|f(x) - z_{f,X}(x)| \leq &\frac{1}{c}\sum_{j:\|x-x_j\|_2 \leq h_{X,\Omega}} \omega_f(h_{X,\Omega}) \underbrace{|u_j(x)|}_{\leq C\varphi(0)} + \frac{1}{c}\sum_{j:\|x-x_j\|_2 > h_{X,\Omega}} \omega_f(\|x - x_j\|_2)|u_j(x)| \leq \\
\leq &\frac{1}{c}C\varphi(0)\omega_f(h_{X,\Omega})\left(1 + \frac{h_{X,\Omega}}{q_X}\right)^d + \frac{1}{c}\sum_{j:\|x-x_j\|_2 > h_{X,\Omega}} 2\lceil\gamma\rceil \frac{\|x - x_j\|_2}{h_{X,\Omega}} \omega_f(h_{X,\Omega})|u_j(x)| \leq \\
\leq &\omega_f(h_{X,\Omega})\left(\frac{1}{c}C\varphi(0)(1 + c_{qu})^d + \frac{2\lceil\gamma\rceil}{h_{X,\Omega}c}\sum_{j=1}^N \|x - x_j\|_2|u_j(x)|\right) \leq \\
\leq &\omega_f(h_{X,\Omega})\frac{1}{c}\max\{C\varphi(0)(1 + c_{qu})^d, 2\lceil\gamma\rceil\}\left(1 + \frac{1}{h_{X,\Omega}}\sum_{j=1}^N \|x - x_j\|_2|u_j(x)|\right) \leq \\
\leq &\omega_f(h_{X,\Omega})\underbrace{\frac{1}{c}\max\{C\varphi(0)(1 + c_{qu})^d, 2\lceil\gamma\rceil\}(1 + K(1, C, \varphi, d))}_{K},
\end{aligned}
$$

where the last inequality holds for Theorem 7.1.

We can conclude by remarking $\omega_f(h_{X,\Omega}) \leq h_{X,\Omega}\|f\|_{\mathcal{C}^1(\Omega^*)}$, because if $x, y \in \Omega$ with $\|x - y\|_2 \leq h_{X,\Omega}$ then by defining $q : [0,1] \to \mathbb{R}$ as $q(t) = f(x + t(y - x))$ for $t \in [0,1]$ and applying Lagrange Theorem we have

$$
f(y) - f(x) = \frac{q(1) - q(0)}{1 - 0} = \frac{\mathrm{d}\,q}{\mathrm{d}\,t}(\xi) = \nabla f(x + \xi(y - x)) \cdot (y - x),
$$

with $x + \xi(y - x) \in \Omega^*$ since $\|x + \xi(y - x) - x\|_2 = |\xi|\|y - x\|_2 \leq h_{X,\Omega} < r_0$.

□

7.1.1 An example for quasi-uniform Shepard method

The goal of this section is to build a method that satisfies the hypothesis of Definition 7.1.

Let us start by fixing a positive definite function $\Phi : \mathbb{R}^d \to \mathbb{R}$ that satisfy the hypothesis of Theorem 1.26 with support in the closure of the unit ball $B(0,1)$. It is important in this construction to consider the scaled version of Φ with a parameter $\delta \in]0,1[$.

Since we work in the quasi-uniform setting of Definition 3.6 then we have

$$q_X \leq h_{X,\Omega} \leq c_{qu} q_X$$

and we assume that there exists constants $\gamma \in]0,1[$ and $c_\gamma > 1$ such that

$$\gamma c_\gamma h_{X,\Omega} \leq \delta \leq c_\gamma h_{X,\Omega}, \tag{7.2}$$

which becomes for quasi-uniformity of X

$$\gamma c_\gamma q_X \leq \delta \leq c_\gamma c_{qu} q_X. \tag{7.3}$$

If we suppose that $\delta \geq h_{X,\Omega}$ this implies that every $x \in \Omega$ is in the support of at least one of the functions $\{\Phi_\delta(\cdot - x_j)\}_{j=1,\ldots,N}$, because there exists $j \in \{1,\ldots,N\}$ such that $\|x - x_j\|_2$ is minimal that implies $\|x - x_j\|_2 \leq h_{X,\Omega} \leq \delta$.

For each quasi-uniform set $X = \{x_1,\ldots,x_N\}$ with respect to c_{qu} we choose $\{u_1,\ldots,u_N\}$ of the Definition 7.1 to be the cardinal functions of Φ_δ with respect to X (Theorem 4.1).

To continue we analyse the following theorem.

Theorem 7.5 *Suppose that $\Phi \in L^1(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d)$ has support in $\overline{B(0,1)}$ and that satisfies*

$$c_1(1 + \|\omega\|_2^2)^{-s} \leq \widehat{\Phi}(\omega) \leq c_2(1 + \|\omega\|_2^2)^{-s}$$

with $s > \frac{d}{2}$ and two constant $0 < c_1 \leq c_2$. If $X = \{x_1,\ldots,x_N\}$ is a set of distinct points then, under the condition $\delta \geq q_X$, we have

$$\|A^{-1}_{\Phi_\delta,X}\|_2 \leq C(\Phi,d)\left(\frac{\delta}{q_X}\right)^{2s-d} \qquad \text{and} \qquad \|A_{\Phi_\delta,X}\|_2 \leq \left(1 + \frac{\delta}{q_X}\right)^d \Phi(0),$$

that implies

$$cond_2(A_{\Phi_\delta,X}) \leq C(\Phi,d)\Phi(0)\left(\frac{\delta}{q_X}\right)^{2s-d}\left(1 + \frac{\delta}{q_X}\right)^d.$$

Proof
The first inequality is a consequence of equation (5.6), indeed

$$\|A^{-1}_{\Phi_\delta,X}\|_2 = \frac{1}{\lambda_{\min}(A_{\Phi_\delta,X})} \leq C(\Phi,d)\left(\frac{\delta}{q_X}\right)^{2s-d}.$$

From equation (5.3) we have that there exists $i \in \{1, \dots, N\}$ such that

$$\lambda_{\max}(A_{\Phi_\delta,X}) - \Phi(0) \leq |\lambda_{\max}(A_{\Phi_\delta,X}) - \Phi_\delta(0)| \leq \sum_{\substack{j=1 \\ j\neq i}}^{N} |\Phi_\delta(x_i - x_j)| =$$

$$= \sum_{\substack{j:\|x_i - x_j\|\leq\delta \\ j\neq i}}^{N} |\Phi_\delta(x_i - x_j)| \leq \left(\left(1 + \frac{\delta}{q_X}\right)^d - 1\right)\Phi(0),$$

that conclude the proof since $\|A_{\Phi_\delta,X}\|_2 = \lambda_{\max}(A_{\Phi_\delta,X})$.

□

First of all, since we are working in a quasi-uniform setting, we state that the construction of the cardinal functions $\{u_1, \dots, u_N\}$ has a reasonable computational cost and it is stable. The matrix $A_{\Phi,X}$ in Theorem 4.1 is sparse and each row has an upper bound on the maximum number of non-zero element: if we apply Theorem 7.3 for $j = 1, \dots, N$ we have

$$\#\{x_k : \|x_k - x_j\|_2 \leq \delta\} \leq \left(1 + \frac{\delta}{q_X}\right)^d \leq (1 + c_\gamma c_{qu})^d$$

and with Theorem 7.5 we achieve

$$\begin{aligned} \mathrm{cond}_2(A_{\Phi_\delta,X}) &\leq C(\Phi,d)\Phi(0)\left(\frac{\delta}{q_X}\right)^{2s-d}\left(1 + \frac{\delta}{q_X}\right)^d \leq \\ &\leq C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1 + c_\gamma c_{qu})^d, \end{aligned} \tag{7.4}$$

furthermore, the stability of the interpolation is guaranteed by equation (5.2) since

$$\frac{1}{\lambda_{\min}(A_{\Phi_\delta,X})} = \|A^{-1}_{\Phi_\delta,X}\|_2 \leq C(\Phi,d)\left(\frac{\delta}{q_X}\right)^{2s-d} \leq C(\Phi,d)(c_\gamma c_{qu})^{2s-d}. \tag{7.5}$$

All these constants are independent of the cardinality of X, that lead the interpolation scheme to be computationally efficient in the stationary case (the radius of the support δ is proportional to the separation distance q_X).

Another sign of efficiency is given by the point evaluation at $x \in \Omega$ of the interpolant of a function $f : \Omega \to \mathbb{R}^d$ because (equation (2.1))

$$s_{f,X}(x) = \sum_{j=1}^{N} \alpha_j \Phi_\delta(x - x_j) = \sum_{j:\|x - x_j\|_2\leq\delta} \alpha_j \Phi_\delta(x - x_j)$$

and the number of non-zero summands are bounded by $(1 + c_\gamma c_{qu})^d$ (Theorem 7.3).

Our goal is to prove that for $i = 1, \dots, N$

$$|u_i(x)| \leq K(\Phi, d, c_\gamma c_{qu}) e^{-\nu(\Phi,d,c_\gamma c_{qu})\frac{\|x - x_i\|_2}{q_X}}.$$

To prove this we need a more general definition of matrix, we need a different way to index its components.

Definition 7.3 *We say that a multivariate matrix A is a finitely supported function $A : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}$. If we denote with A_{ij} the scalar $A(i, j)$ we have a natural extension of symmetry and positive definiteness.*

- *A is symmetric if $A_{ij} = A_{ji}$ for all $i, j \in \mathbb{Z}^d$,*
- *A is positive definite if*
$$\sum_{i,j\in\mathbb{Z}^d} y_i y_j A_{ij} > 0$$
for every finitely supported non-zero real sequence $\{y_j\}_{j\in\mathbb{Z}^d}$ such that the support of $(y_i y_j)_{i,j\in\mathbb{Z}^d}$ is contained in the support of A
- *A is R-banded with $R > 0$ if $A_{ij} = 0$ whenever $\|i - k\|_2 > R$ with $i, j \in \mathbb{Z}^d$.*

It has a fundamental importance for our work the following result of [49].

Theorem 7.6 *Let $A = (A_{ij})_{i,j\in\mathbb{Z}^d}$ be a symmetric, positive definite matrix which is R-banded with $R \in \mathbb{N}_{\geq 1}$. Then for $i, j \in \mathbb{Z}^d$*
$$|(A^{-1})_{ij}| \leq 2\|A^{-1}\|_2 \mu^{\|i-j\|_2},$$
where
$$\mu = \left(\frac{\sqrt{cond_2(A)} - 1}{\sqrt{cond_2(A)} + 1} \right)^{\frac{1}{R}}.$$

We have to translate $A_{\Phi_\delta,X}$ is this new context.

Theorem 7.7 *(D. Hensley) Let $\{z_1, \ldots, z_N\} \subseteq \mathbb{R}^d$ such that $\|z_j - z_k\|_2 \geq \sqrt{d}$ if $j \neq k$. Define*
$$\nu_j = (\lfloor z_j^1 \rfloor, \ldots, \lfloor z_j^d \rfloor) \quad \text{for } j = 1, \ldots, N.$$
Then the points $\{\nu_j\}_{j=1,\ldots,N}$ are pairwise distinct and if $\|\nu_j - \nu_k\|_2 \geq R$ with $R \geq 4\sqrt{d}$ then $\|z_j - z_k\|_2 \geq \frac{R}{2}$.

Proof
Suppose by contradiction that $\nu_j = \nu_k$ with $j \neq k$. This implies that $\lfloor z_j^\ell \rfloor = \lfloor z_k^\ell \rfloor$ for $\ell = 1, \ldots, d$. Since for $x \in \mathbb{R}$ holds $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$ we have that
$$\lfloor z_j^\ell \rfloor - 1 - \lfloor z_k^\ell \rfloor < z_j^\ell - z_k^\ell < \lfloor z_j^\ell \rfloor + 1 - \lfloor z_k^\ell \rfloor,$$

which gives us $|z_j^\ell - z_k^\ell| < 1$ for $\ell = 1, \dots, d$. In this situation we obtain

$$d \le \|z_j - z_k\|_2^2 = \sum_{\ell=1}^{d} |z_j^\ell - z_k^\ell|^2 < d,$$

that is a contradiction. To finish the proof we note that

$$\|\nu_j - \nu_k\|_2 \le \|\nu_j - z_j\|_2 + \|z_j - z_k\|_2 + \|z_k - \nu_k\|_2 < 2\sqrt{d} + \|z_j - z_k\|_2,$$

because $\sum_{\ell=1}^{d} |\lfloor z_i^\ell \rfloor - z_i^\ell|^2 < d$ for $i = 1, \dots, N$. If $\|\nu_j - \nu_k\|_2 \ge R \ge 4\sqrt{d}$ then

$$\|z_j - z_k\|_2 \ge \|\nu_j - \nu_k\|_2 - 2\sqrt{d} \ge R - 2\sqrt{d} = \frac{R}{2} + \frac{R}{2} - 2\sqrt{d} \ge \frac{R}{2}.$$

□

Let us fix some notation. We define

$$z_j = \frac{\sqrt{d}}{2q_X} x_j \quad \text{for } j = 1, \dots, N$$

and we move the points $\{z_j\}_{j=1,\dots,N}$ in a unique way to the lattice $\mathbb{Z}^d$ as

$$y_j = (\lfloor z_j^1 \rfloor, \dots, \lfloor z_j^d \rfloor) \in \mathbb{Z}^d \quad \text{for } j = 1, \dots, N. \tag{7.6}$$

The points $\{y_j\}_{j=1,\dots,N}$ are pairwise distinct because we can apply Theorem 7.7 since for $j \ne k$

$$\|z_j - z_k\|_2 = \frac{\sqrt{d}}{2q_X} \|x_j - x_k\|_2 \ge \sqrt{d}.$$

We define $A_{\Phi_\delta, X}$ by

$$A_{\Phi_\delta, X}(y_i, y_j) = \Phi_\delta(x_i - x_j) \quad \text{for } i, j = 1, \dots, N. \tag{7.7}$$

If we define $R(c_\gamma c_{qu}) = \max\{c_\gamma c_{qu}, 4\}\sqrt{d}$ then we claim that $A_{\Phi_\delta, X}$ is $R(c_\gamma c_{qu})$-banded. By Theorem 7.7 we can obtain from $\|y_i - y_j\|_2 \ge R(c_\gamma c_{qu})$ that

$$\frac{\|x_i - x_j\|_2}{2q_X} \sqrt{d} \ge \frac{R}{2} \ge \frac{c_\gamma c_{qu} \sqrt{d}}{2},$$

that implies with equation (7.3)

$$\|x_i - x_j\|_2 \ge c_\gamma c_{qu} q_X \ge \delta \Rightarrow A_{\Phi_\delta, X}(y_i, y_j) = 0.$$

Theorem 7.8 *Suppose that $X = \{x_1, \dots, x_N\}$ is a quasi-uniform data set with respect to c_{qu} and $\Phi : \mathbb{R}^d \to \mathbb{R}$ is a positive definite function with support in $\overline{B(0,1)}$ that satisfies the hypothesis of Theorem 1.26. If we choose $\delta \ge h_{X,\Omega}$ as in equation (7.2) then*

$$|(A_{\Phi_\delta, X}^{-1})_{ij}| \le C(\Phi, d, c_\gamma c_{qu}) \mu(\Phi, d, c_\gamma c_{qu})^{\|y_i - y_k\|_2},$$

where $\{y_i\}_{i=1,\dots,N}$ are defined in equation (7.6). Moreover $\mu(\Phi, d, c_\gamma c_{qu}) \in]0, 1[$.

Proof
We apply Theorem 7.6 to the matrix $A_{\Phi_\delta,X}$ defined in equation (7.7). By the definition of $A_{\Phi_\delta,X}$ we proved that it is $R(c_\gamma c_{qu})$-banded with $R(c_\gamma c_{qu}) = \max\{c_\gamma c_{qu}, 4\}\sqrt{d}$ and it is possible to conclude that

$$|(A^{-1}_{\Phi_\delta,X})_{ij}| = |(A^{-1}_{\Phi_\delta,X})_{y_i y_j}| \leq 2\|A^{-1}_{\Phi_\delta,X}\|_2 \widetilde{\mu}^{\|y_i - y_j\|_2}$$

with

$$\widetilde{\mu} = \left(\frac{\sqrt{\mathrm{cond}_2(A_{\Phi_\delta,X})} - 1}{\sqrt{\mathrm{cond}_2(A_{\Phi_\delta,X})} + 1} \right)^{\frac{1}{R(c_\gamma c_{qu})}}.$$

Since the function $g(x) = \frac{x-1}{x+1}$ is monotonically increasing due to the fact that $\frac{\mathrm{d}g}{\mathrm{d}x}(x) = \frac{2}{(x+1)^2} > 0$ we can bound $\widetilde{\mu}$ as

$$\widetilde{\mu} \leq \left(\frac{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1 + c_\gamma c_{qu})^d} - 1}{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1 + c_\gamma c_{qu})^d} + 1} \right)^{\frac{1}{R(c_\gamma c_{qu})}} = \mu(\Phi, d, c_\gamma c_{qu}),$$

where we use equation (7.4). $\mu(\Phi, d, c_\gamma c_{qu}) < 1$ because g is increasing and $g(x) = \frac{x-1}{x+1} \nearrow 1^-$ as $x \to +\infty$. With equation (7.5) we can conclude

$$C(\Phi, d, c_\gamma c_{qu}) = 2C(\Phi, d)(c_\gamma c_{qu})^{2s-d}.$$

□

We can finally prove the exponential decay of $\{u_1, \ldots, u_N\}$.

Theorem 7.9 *Suppose that $X = \{x_1, \ldots, x_N\}$ is a quasi-uniform data set with respect to c_{qu} and $\Phi : \mathbb{R}^d \to \mathbb{R}$ is a positive definite function with support in $\overline{B(0,1)}$ that satisfies the hypothesis of Theorem 1.26. If we choose $\delta \geq h_{X,\Omega}$ as in equation (7.2) then there exists $K(\Phi, d, c_\gamma c_{qu}) > 0$ and $\nu(\Phi, d, c_\gamma c_{qu}) > 0$ such that*

$$|u_i(x)| \leq K(\Phi, d, c_\gamma c_{qu}) e^{-\nu(\Phi,d,c_\gamma c_{qu}) \frac{\|x - x_i\|_2}{q_X}} \qquad \text{for } x \in \mathbb{R}^d \text{ and } i = 1, \ldots, N.$$

Proof
First of all we prove the relation $||x| - \lfloor y \rfloor| \geq |x - y| - 1$ for $x, y \in \mathbb{R}$. Without loss of generality we can suppose $x \geq y$. Since $\lfloor x \rfloor + 1 > x \geq y \geq \lfloor y \rfloor$ then $\lfloor x \rfloor \geq \lfloor y \rfloor$, that implies $|x - y| = x - y \leq \lfloor x \rfloor + 1 - \lfloor y \rfloor = |\lfloor x \rfloor - \lfloor y \rfloor| + 1$.

To continue we need to prove $\sqrt{d}\|x\|_2 \geq \|x\|_1 \geq \|x\|_2$ for $x \in \mathbb{R}^d$.

$$\sum_{j=1}^{d} |x_j| = \sum_{j=1}^{d} |x_j| \cdot 1 = (|x_1|, \ldots, |x_d|) \cdot (1, \ldots, 1) \leq$$
$$\leq \|(|x_1|, \ldots, |x_d|)\|_2 \|(1, \ldots, 1)\|_2 = \|x\|_2 \sqrt{d}.$$

By recalling that $\sqrt{|a|+|b|} \leq \sqrt{|a|} + \sqrt{|b|}$ we have

$$\|x\|_2 = \sqrt{\sum_{j=1}^d |x_j|^2} \leq \sum_{j=1}^d \sqrt{|x_j|^2} = \|x\|_1.$$

Now we can obtain the following chain of inequalities

$$\begin{aligned}\|y_i - y_j\|_2 \geq \frac{1}{\sqrt{d}}\|y_i - y_j\|_1 = \frac{1}{\sqrt{d}}\sum_{k=1}^d |\lfloor z_i^k \rfloor - \lfloor z_j^k \rfloor| \geq \\ \geq \frac{1}{\sqrt{d}}\sum_{k=1}^d |z_i^k - z_j^k| - \frac{d}{\sqrt{d}} = \frac{1}{\sqrt{d}}\|z_i - z_j\|_1 - \sqrt{d} \geq \\ \geq \frac{1}{\sqrt{d}}\|z_i - z_j\|_2 - \sqrt{d} = \frac{1}{2q_X}\|x_i - x_j\|_2 - \sqrt{d}.\end{aligned}$$

With Theorem 7.8 we achieve

$$\begin{aligned}|(A^{-1}_{\Phi_\delta,X})_{ij}| \leq C(\Phi,d,c_\gamma c_{qu})\mu(\Phi,d,c_\gamma c_{qu})^{\|y_i - y_k\|_2} \leq \\ \leq C(\Phi,d,c_\gamma c_{qu})\mu(\Phi,d,c_\gamma c_{qu})^{\frac{\|x_i - x_j\|}{2q_X}}\mu(\Phi,d,c_\gamma c_{qu})^{-\sqrt{d}}\end{aligned}$$

because $\log(\mu(\Phi,d,c_\gamma c_{qu})) < 0$. If we fix $\nu(\Phi,d,c_\gamma c_{qu}) = -\frac{1}{2}\log(\mu(\Phi,d,c_\gamma c_{qu})) > 0$ we have

$$|(A^{-1}_{\Phi_\delta,X})_{ij}| \leq C(\Phi,d,c_\gamma c_{qu})\mu(\Phi,d,c_\gamma c_{qu})^{-\sqrt{d}} e^{-\nu(\Phi,d,c_\gamma c_{qu})\frac{\|x_i - x_j\|}{q_X}}.$$

If $x \in \mathbb{R}^d$ and $\|x - x_j\|_2 \leq \delta$ for $j \in \{1,\ldots,N\}$ then

$$\|x_i - x_j\|_2 \geq \|x_i - x\|_2 - \|x_j - x\|_2 \geq \|x_i - x\|_2 - \delta.$$

With Theorem 4.1 we can write

$$\begin{aligned}|u_i(x)| &= \left|\sum_{j=1}^N (A^{-1}_{\Phi_\delta,X})_{ij}\Phi_\delta(x - x_j)\right| \leq \sum_{j=1}^N |(A^{-1}_{\Phi_\delta,X})_{ij}||\Phi_\delta(x - x_j)| = \\ &= \sum_{j:\|x - x_j\|_2 \leq \delta} |(A^{-1}_{\Phi_\delta,X})_{ij}||\Phi_\delta(x - x_i)| \leq \|\Phi\|_{L^\infty(\mathbb{R}^d)} \sum_{j:\|x - x_j\|_2 \leq \delta} |(A^{-1}_{\Phi_\delta,X})_{ij}| \leq \\ &\leq \|\Phi\|_{L^\infty(\mathbb{R}^d)} \sum_{j:\|x - x_j\|_2 \leq \delta} C(\Phi,d,c_\gamma c_{qu})\mu(\Phi,d,c_\gamma c_{qu})^{-\sqrt{d}} e^{-\nu(\Phi,d,c_\gamma c_{qu})\frac{\|x_i - x_j\|}{q_X}} \leq \\ &\leq \|\Phi\|_{L^\infty(\mathbb{R}^d)} C(\Phi,d,c_\gamma c_{qu})\mu(\Phi,d,c_\gamma c_{qu})^{-\sqrt{d}} e^{\nu(\Phi,d,c_\gamma c_{qu})\frac{\delta}{q_X}} \sum_{j:\|x - x_j\|_2 \leq \delta} e^{-\nu(\Phi,d,c_\gamma c_{qu})\frac{\|x_i - x\|}{q_X}} \leq \\ &\leq \underbrace{\|\Phi\|_{L^\infty(\mathbb{R}^d)} C(\Phi,d,c_\gamma c_{qu})\mu(\Phi,d,c_\gamma c_{qu})^{-\sqrt{d}} e^{\nu(\Phi,d,c_\gamma c_{qu})c_\gamma c_{qu}}(1 + c_\gamma c_{qu})^d}_{K(\Phi,d,c_\gamma c_{qu})} e^{-\nu(\Phi,d,c_\gamma c_{qu})\frac{\|x_i - x\|}{q_X}}.\end{aligned} \tag{7.8}$$

The last inequality holds because of Theorem 7.3. The constants $C(\Phi,d,c_\gamma c_{qu})$ and $\mu(\Phi,d,c_\gamma c_{qu})$ are defined in Theorem 7.8.

$\square$

Theorem 7.10 *Suppose that $X = \{x_1, \ldots, x_N\}$ is a quasi-uniform data set with respect to c_{qu} and $\Phi : \mathbb{R}^d \to \mathbb{R}$ is a positive definite function with support in $\overline{B(0,1)}$ that satisfies the hypothesis of Theorem 1.26. If we choose $\delta \geq h_{X,\Omega}$ as in equation (7.2) then there exists $K_L(\Phi, d, c_\gamma c_{qu}, \gamma c_\gamma) > 0$ such that*

$$|u_i(x) - u_i(y)| \leq \frac{K_L(\Phi, d, c_\gamma c_{qu}, \gamma c_\gamma)}{q_X} \|x - y\|_2 \quad \textit{for } x, y \in \mathbb{R}^d \textit{ and } i = 1, \ldots, N.$$

Proof

By recalling that $\partial_k(\Phi_\delta(\cdot - x_j)) = \frac{1}{\delta}\partial_k \Phi\left(\frac{\cdot - x_j}{\delta}\right)$ for $j = 1, \ldots, N, k = 1, \ldots, d$ then by substituting $\partial_k u_i(x)$ with $u_i(x)$ in equation (7.8) we have

$$|\partial_k u_i(x)| \leq \frac{K(\Phi, d, c_\gamma c_{qu})}{\delta} e^{-\nu(\Phi, d, c_\gamma c_{qu}) \frac{\|x_i - x\|}{q_X}} \leq \frac{K(\Phi, d, c_\gamma c_{qu})}{\delta}$$

for $x \in \mathbb{R}^d, i = 1, \ldots, N, k = 1, \ldots, d$. The constants involved are defined by

$$\nu(\Phi, d, c_\gamma c_{qu}) = -\frac{1}{2} \log(\mu(\Phi, d, c_\gamma c_{qu})) > 0$$

and

$$K(\Phi, d, c_\gamma c_{qu}) = \|\Phi\|_{C^1(\mathbb{R}^d)} C(\Phi, d, c_\gamma c_{qu}) \mu(\Phi, d, c_\gamma c_{qu})^{-\sqrt{d}} e^{\nu(\Phi, d, c_\gamma c_{qu}) c_\gamma c_{qu}} (1 + c_\gamma c_{qu})^d$$

with $C(\Phi, d, c_\gamma c_{qu})$ and $\mu(\Phi, d, c_\gamma c_{qu})$ as in Theorem 7.8.

Fix $x, y \in \mathbb{R}^d$ and define $q : [0, 1] \to \mathbb{R}$ as $q(t) = u_i(y + t(x - y))$ for $t \in [0, 1]$ and $i \in \{1, \ldots, N\}$. From Lagrange Theorem we have

$$u_i(x) - u_i(y) = \frac{q(1) - q(0)}{1 - 0} = \frac{\mathrm{d}\, q}{\mathrm{d}\, t}(\xi) = \nabla u_i(y + \xi(x - y)) \cdot (x - y),$$

so

$$|u_i(x) - u_i(y)| \leq \|\nabla u_i(y + \xi(x - y))\|_2 \|x - y\|_2 \leq \sqrt{d} \frac{K(\Phi, d, c_\gamma c_{qu})}{\delta} \|x - y\|_2.$$

From equation (7.3) we obtain

$$K_L(\Phi, d, c_\gamma c_{qu}, \gamma c_\gamma) = \sqrt{d} \frac{K(\Phi, d, c_\gamma c_{qu})}{\gamma c_\gamma}.$$

$\square$

If we define $\varphi(x) = K(\Phi, d, c_\gamma c_{qu}) e^{-\nu(\Phi, d, c_\gamma c_{qu}) x}$ for $x \in \mathbb{R}_{\geq 0}$ then φ is decreasing and

$$\frac{\varphi(n+1)}{\varphi(n)} = \frac{K(\Phi, d, c_\gamma c_{qu}) e^{-\nu(\Phi, d, c_\gamma c_{qu})(n+1)}}{K(\Phi, d, c_\gamma c_{qu}) e^{-\nu(\Phi, d, c_\gamma c_{qu}) n}} = e^{-\nu(\Phi, d, c_\gamma c_{qu})} < 1.$$

From the facts proved in Theorem 7.9 we can obtain

$$|u_i(x)| \leq \varphi\left(\frac{\|x - x_i\|_2}{q_X}\right) \quad \text{for } i = 1, \ldots, N$$

and if Definition 7.1 holds then Theorem 7.4 guarantees an $\mathcal{O}(h_{X,\Omega})$ convergence rate.

7.1.2 A conjecture for rescaled localized RBFs

Now we can state a conjecture.

Theorem 7.11 *We can choose* c_γ *in equation* (7.2) *that if* $X = \{x_1, \dots, x_N\}$ *is a quasi-uniform data set with respect to* c_{qu}*,* $\Phi : \mathbb{R}^d \to \mathbb{R}$ *is a positive definite function with support in* $\overline{B(0,1)}$ *that satisfies the hypothesis of Theorem 1.26 and* $\delta \geq h_{X,\Omega}$ *as in equation* (7.2) *then there exists a constant* $c > 0$ *such that*

$$\sum_{j=1}^{N} u_j(x) > c \qquad \textit{for } x \in \Omega.$$

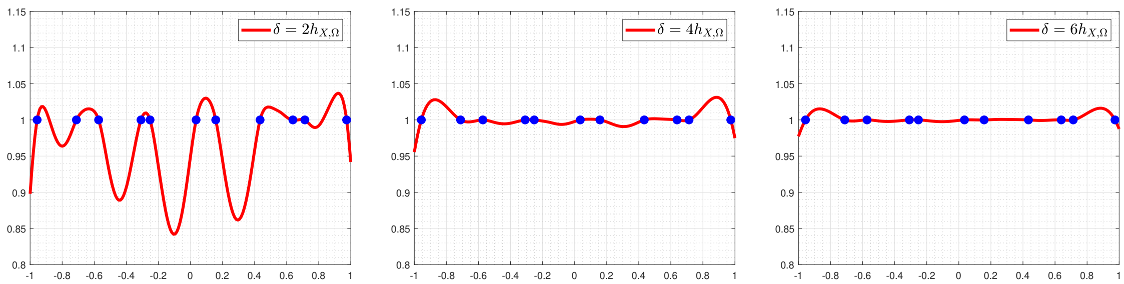


Figure 7.1: Interpolant $s_{1,X,\delta}$ on 11 uniformly perturbed equispaced points in $[-1,1]$ as the choice of delta varies. The compactly supported RBF is the Wendland's function $\phi_{1,1} \in \mathcal{C}^2([-1,1])$ [52, 53].

From Figure 7.1 we notice that when the ratio between δ and $h_{X,\Omega}$ gets bigger the approximation gets better. In this case c_γ in equation (7.2) is bigger and implies that more nodes are used locally to produce the approximant. The drawback is that the condition number of $A_{\Phi_\delta,X}$ increases (equation (7.4)) but the matrix remains sparse.

Now we will give some theoretical intuitions that convinced us of the veracity of Theorem 7.11.

We introduce a result of [49].

Theorem 7.12 *With notation of Theorem 7.6 if* $A = (A_{ij})_{i,j\in\mathbb{Z}^d}$ *is a symmetric, positive definite matrix which is* R*-banded with* $R \in \mathbb{N}_{\geq 1}$ *then*

$$\|A^{-1}\|_\infty \leq 2\|A^{-1}\|_2 \left(\frac{1+\mu^{1/\sqrt{d}}}{1-\mu^{1/\sqrt{d}}}\right)^d = 2\|A^{-1}\|_2 \left(\tanh\left(\frac{1}{2Rd}\log\left(\frac{\kappa+1}{\kappa-1}\right)\right)\right)^{-d},$$

where

$$\kappa = \sqrt{cond_2(A)} \qquad \textit{and} \qquad \mu = \left(\frac{\sqrt{cond_2(A)}-1}{\sqrt{cond_2(A)}+1}\right)^{\frac{1}{R}}.$$

Moreover, if $\|A^{-1}\|_2 \leq \alpha$ *and* $cond_2(A) \leq \beta$ *then*

$$\|A^{-1}\|_\infty \leq 2\alpha \left(\frac{1+\nu}{1-\nu}\right)^d$$

with

$$\nu = \left(\frac{\sqrt{\beta}-1}{\sqrt{\beta}+1}\right)^{\frac{1}{R\sqrt{d}}}.$$

Theorem 7.12 with equation (7.4), equation (7.5) and construction in equation 7.7 gives us

$$\|A^{-1}_{\Phi_\delta,X}\|_\infty \leq C_\infty(\Phi, d, c_\gamma c_{qu}). \tag{7.9}$$

Theorem 7.13 *Suppose that* $X = \{x_1, \dots, x_N\}$ *is a quasi-uniform data set with respect to* c_{qu} *and* $\Phi : \mathbb{R}^d \to \mathbb{R}$ *is a positive definite function with support in* $\overline{B(0,1)}$ *that satisfies the hypothesis of Theorem 1.26. If we choose* $\delta \geq h_{X,\Omega}$ *as in equation* (7.2) *then*

$$\|s_{f,X,\delta}\|_{L^\infty(\Omega)} \leq (1 + c_\gamma c_{qu})^d \|\Phi\|_{L^\infty(\Omega)} C_\infty(\Phi, d, c_\gamma c_{qu}) \|f\|_{L^\infty(\Omega)}.$$

Proof
By equation (2.1)

$$s_{f,X,\delta}(x) = \sum_{j=1}^{N} \alpha_j \Phi_\delta(x - x_j),$$

where $\alpha \in \mathbb{R}^d$ solves $A_{\Phi_\delta,X}\alpha = f_{|X}$. Thus,

$$|s_{f,X,\delta}(x)| = \left|\sum_{j=1}^{N} \alpha_j \Phi_\delta(x - x_j)\right| = \left|\sum_{j:\|x-x_j\|_2 \leq \delta} \alpha_j \Phi_\delta(x - x_j)\right| \overset{\text{Theorem 7.3}}{\leq}$$
$$\leq (1 + c_\gamma c_{qu})^d \|\Phi\|_{L^\infty(\Omega)} \|\alpha\|_\infty = (1 + c_\gamma c_{qu})^d \|\Phi\|_{L^\infty(\Omega)} \|A^{-1}_{\Phi_\delta,X} f_{|X}\|_\infty \leq$$
$$\leq (1 + c_\gamma c_{qu})^d \|\Phi\|_{L^\infty(\Omega)} \|A^{-1}_{\Phi_\delta,X}\|_\infty \|f_{|X}\|_\infty.$$

We can conclude with equation (7.9).

□

We proved that there is no Runge phenomenon if we perform interpolation with scaled compactly supported radial basis function. The scaling has to be proportional to the fill distance $h_{X,\Omega}$ and X is a quasi-uniform data set.

If we apply Theorem 7.13 to $f = 1$ we have

$$\left|\sum_{j=1}^{N} u_j(x)\right| \leq (1 + c_\gamma c_{qu})^d \|\Phi\|_{L^\infty(\Omega)} C_\infty(\Phi, d, c_\gamma c_{qu}),$$

where $\{u_1, \dots, u_N\}$ are the cardinal functions of Φ_δ with respect to X (Theorem 4.1). If there are oscillations in the interpolant $s_{1,X,\delta}$ these are bounded.

We can obtain an explicit formula for the difference $s_{f,X,\delta} - 1$. Since Φ_δ is positive definite then Φ_δ is also conditionally positive definite with respect to $\Pi_0(\mathbb{R}^d) = \langle 1 \rangle$ and we compare the cardinal functions provided by the two different interpolation method.

The cardinal functions $\{u_1, \dots, u_N\}$ of Φ_δ considered as a positive definite function satisfy

$$A_{\Phi_\delta,X} u(x) = R(x),$$

instead the cardinal functions $\{\widetilde{u}_1, \dots, \widetilde{u}_N\}$ of Φ_δ considered as a conditionally positive definite function with respect to $\Pi_0(\mathbb{R}^d) = \langle 1 \rangle$ satisfy

$$\begin{pmatrix} A_{\Phi_\delta,X} & P \\ P^\top & 0 \end{pmatrix} \begin{pmatrix} \widetilde{u}(x) \\ v(x) \end{pmatrix} = \begin{pmatrix} R(x) \\ S(x) \end{pmatrix}, \tag{7.10}$$

where $R(x) = (\Phi(x, x_1), \dots, \Phi(x, x_N))^\top \in \mathbb{R}^N$, $S(x) = 1 \in \mathbb{R}$ and $P = (1, \dots, 1)^\top \in \mathbb{R}^N$.

With this relation we obtain

$$A_{\Phi_\delta,X}(u(x) - \widetilde{u}(x)) = R(x) - R(x) + Pv(x) = Pv(x), \tag{7.11}$$

where $\lambda(x) = 2v(x)$ is the Lagrange multiplier associated to the optimal solution of the minimization problem in Theorem 4.4 for $\alpha = 0$. Moreover, $v(x_j) = 0$ for $j = 1, \dots, N$.

We can write explicitly the difference $s_{1,X,\delta} - 1$ as

$$\sum_{j=1}^N u_j(x) - \sum_{j=1}^N \widetilde{u}_j(x) = P^\top (u(x) - \widetilde{u}(x)) = P^\top A_{\Phi_\delta,X}^{-1} Pv(x),$$

where $v(x)$ can also be characterized with equation (7.10) as

$$Nv(x) = P^\top Pv(x) = P^\top (R(x) - A_{\Phi_\delta,X} \widetilde{u}(x)).$$

Since $Pv(x) = R(x) - A_{\Phi_\delta,X} \widetilde{u}(x)$, we note that for $j = 1, \dots, N$

$$\begin{aligned} v(x) = (R(x) - A_{\Phi_\delta,X} \widetilde{u}(x))_j &= \Phi_\delta(x - x_j) - \sum_{i=1}^N \Phi_\delta(x_j - x_i) \widetilde{u}_i(x) = \\ &= \Phi_\delta(x - x_j) - s^c_{\Phi_\delta(\cdot - x_j),X,\delta}(x), \end{aligned} \tag{7.12}$$

where $s^c_{f,X,\delta}$ is the interpolant on X built from Φ_δ considered as a conditionally positive definite function with respect to $\Pi_0(\mathbb{R}^d) = \langle 1 \rangle$. The expression in equation (7.12) as a function of x_j is a function in $F^c_{\Phi_\delta}(\Omega)$ if Φ_δ is a conditionally positive definite function. We have also for $j = 1, \dots, N$

$$v(x) = \lambda(\Phi_\delta(\cdot - x_j)) \quad \text{with } \lambda = \delta_x - \sum_{i=1}^N \widetilde{u}_i(x) \delta_{x_i} \in L_{\langle 1 \rangle}(\Omega) \subseteq L(\Omega),$$

that with Theorem 2.7 it becomes

$$|v(x)| \leq \|\lambda\|_{\Phi_\delta} \|\Phi_\delta(\cdot - x_j)\|_{\mathcal{N}_{\Phi_\delta}(\Omega)} = \Phi(0)\|\lambda\|_{\Phi_\delta}.$$

We study $\|\lambda\|_{\Phi_\delta}$ with $\lambda = \delta_x - \sum_{i=1}^N \widetilde{u}_i(x)\delta_{x_i}$.

$$\|\lambda\|^2_{\Phi_\delta} = \Phi_\delta(x,x) - 2\sum_{i=1}^N \widetilde{u}_i(x)\Phi_\delta(x,x_i) + \sum_{i,j=1}^N \widetilde{u}_i(x)\widetilde{u}_j(x)\Phi_\delta(x_i,x_j) = (P^c_{\Phi_\delta,X}(x))^2,$$

where $P^c_{\Phi_\delta,X}$ is the power function of the conditionally positive definite function Φ_δ with respect to X.

Finally equation (7.9) and equation (7.11) gives us

$$\begin{aligned}\|u(x) - \widetilde{u}(x)\|_\infty = \|A^{-1}_{\Phi_\delta,X} P v(x)\|_\infty \leq \|A^{-1}_{\Phi_\delta,X}\|_\infty |v(x)| \leq \\ \leq C_\infty(\Phi, d, c_\gamma c_{qu})\Phi(0) P^c_{\Phi_\delta,X}(x).\end{aligned}$$

If we can prove that $\|P^c_{\Phi,X}\|_{L^\infty(\Omega)} \leq K r(h_{X,\Omega})$ with $r :]0,+\infty[\to \mathbb{R}$ increasing such that $\lim_{x\to 0^+} r(x) = 0$ then equation (4.9) gives us

$$\begin{aligned}\|u(x) - \widetilde{u}(x)\|_\infty &\leq C_\infty(\Phi, d, c_\gamma c_{qu})\Phi(0) P^c_{\Phi,X/\delta}(x/\delta) \leq \\ &\leq C_\infty(\Phi, d, c_\gamma c_{qu})\Phi(0) K r\left(\frac{h_{X,\Omega}}{\delta}\right) \leq \\ &\leq C_\infty(\Phi, d, c_\gamma c_{qu})\Phi(0) K r\left(\frac{1}{\gamma c_\gamma}\right) < +\infty.\end{aligned}$$

Under some conditions, we prove that $|u_i(x) - \widetilde{u}_i(x)|$ are uniformly bounded for $i = 1, \ldots, N$ and $\sum_{j=1}^N \widetilde{u}_j(x) = 1$.

A different approach allows us to obtain other similar results.

$$\begin{pmatrix} A_{\Phi_\delta,X} & P \\ P^\top & 0 \end{pmatrix} \begin{pmatrix} \mathbb{I}_N & -A^{-1}_{\Phi_\delta,X}P \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_{\Phi_\delta,X} & -A_{\Phi_\delta,X}A^{-1}_{\Phi_\delta,X}P + P \\ P^\top \mathbb{I}_N & -P^\top A^{-1}_{\Phi_\delta,X}P \end{pmatrix} = \begin{pmatrix} A_{\Phi_\delta,X} & 0 \\ P^\top & -P^\top A^{-1}_{\Phi_\delta,X}P \end{pmatrix}$$

We note that $-P^\top A^{-1}_{\Phi_\delta,X}P < 0$ because $A_{\Phi_\delta,X}$ is positive definite. Since

$$\left\langle \begin{pmatrix} A_{\Phi_\delta,X} & 0 \\ P^\top & -P^\top A^{-1}_{\Phi_\delta,X}P \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} A_{\Phi_\delta,X} & 0 \\ 0 & -P^\top A^{-1}_{\Phi_\delta,X}P \end{pmatrix} \right\rangle$$

then all the matrix involved are invertible so

$$\begin{pmatrix} A_{\Phi_\delta,X} & P \\ P^\top & 0 \end{pmatrix} = \begin{pmatrix} A_{\Phi_\delta,X} & 0 \\ P^\top & -P^\top A^{-1}_{\Phi_\delta,X}P \end{pmatrix} \begin{pmatrix} \mathbb{I}_N & -A^{-1}_{\Phi_\delta,X}P \\ 0 & 1 \end{pmatrix}^{-1}.$$

We can compute

$$\begin{pmatrix} A_{\Phi_\delta,X} & P \\ P^\top & 0 \end{pmatrix}^{-1} = \begin{pmatrix} \mathbb{I}_N & -A^{-1}_{\Phi_\delta,X}P \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_{\Phi_\delta,X} & 0 \\ P^\top & -P^\top A^{-1}_{\Phi_\delta,X}P \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} \mathbb{I}_N & -A^{-1}_{\Phi_\delta,X}P \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A^{-1}_{\Phi_\delta,X} & 0 \\ \frac{P^\top A^{-1}_{\Phi_\delta,X}}{P^\top A^{-1}_{\Phi_\delta,X}P} & -\frac{1}{P^\top A^{-1}_{\Phi_\delta,X}P} \end{pmatrix} =$$

$$= \begin{pmatrix} A^{-1}_{\Phi_\delta,X} - \frac{A^{-1}_{\Phi_\delta,X}PP^\top A^{-1}_{\Phi_\delta,X}}{P^\top A^{-1}_{\Phi_\delta,X}P} & \frac{A^{-1}_{\Phi_\delta,X}P}{P^\top A^{-1}_{\Phi_\delta,X}P} \\ \frac{P^\top A^{-1}_{\Phi_\delta,X}}{P^\top A^{-1}_{\Phi_\delta,X}P} & -\frac{1}{P^\top A^{-1}_{\Phi_\delta,X}P} \end{pmatrix}.$$

We can exploit

$$v(x) = \frac{P^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P} - \frac{1}{P^\top A^{-1}_{\Phi_\delta,X}P}$$

and

$$\widetilde{u}(x) = A^{-1}_{\Phi_\delta,X}R(x) - \frac{A^{-1}_{\Phi_\delta,X}PP^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P} + \frac{A^{-1}_{\Phi,X}P}{P^\top A^{-1}_{\Phi_\delta,X}P} =$$

$$= u(x) - \frac{A^{-1}_{\Phi_\delta,X}PP^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P} + \frac{A^{-1}_{\Phi,X}P}{P^\top A^{-1}_{\Phi_\delta,X}P}.$$

We can estimate

$$|v(x)| = \left|\frac{P^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P} - \frac{1}{P^\top A^{-1}_{\Phi_\delta,X}P}\right| \leq \left|\frac{P^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P}\right| + \left|\frac{1}{P^\top A^{-1}_{\Phi_\delta,X}P}\right| \leq$$

$$\leq \frac{\|P\|_2\|A^{-1}_{\Phi_\delta,X}R(x)\|_2}{\|P\|_2^2\frac{1}{\lambda_{\max}(A_{\Phi_\delta,X})}} + \frac{1}{\|P\|_2^2\frac{1}{\lambda_{\max}(A_{\Phi_\delta,X})}} \leq \frac{\|P\|_2\frac{1}{\lambda_{\min}(A_{\Phi_\delta,X})}\|R(x)\|_2}{\|P\|_2^2\frac{1}{\lambda_{\max}(A_{\Phi_\delta,X})}} + \frac{1}{\|P\|_2^2\frac{1}{\lambda_{\max}(A_{\Phi_\delta,X})}} \leq$$

$$\leq \frac{\lambda_{\max}(A_{\Phi_\delta,X})}{\lambda_{\min}(A_{\Phi_\delta,X})}\frac{\Phi(0)(1+c_\gamma c_{qu})^d}{\sqrt{N}} + \frac{\lambda_{\max}(A_{\Phi_\delta,X})}{N} = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) = \mathcal{O}\left(h^{\frac{d}{2}}_{X,\Omega}\right),$$

where we used Theorem 7.3 and Theorem 3.6. We proved that v converges to 0 for $h_{X,\Omega} \to 0^+$ uniformly in $x \in \Omega$.

In the same way

$$\|\widetilde{u}(x) - u(x)\|_\infty \leq \|\widetilde{u}(x) - u(x)\|_2 = \left\|-\frac{A^{-1}_{\Phi_\delta,X}PP^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P} + \frac{A^{-1}_{\Phi_\delta,X}P}{P^\top A^{-1}_{\Phi_\delta,X}P}\right\|_2 \leq$$

$$\leq \left\|\frac{A^{-1}_{\Phi_\delta,X}PP^\top A^{-1}_{\Phi_\delta,X}R(x)}{P^\top A^{-1}_{\Phi_\delta,X}P}\right\|_2 + \left\|\frac{A^{-1}_{\Phi_\delta,X}P}{P^\top A^{-1}_{\Phi_\delta,X}P}\right\|_2 \leq$$

$$\leq \frac{\|A^{-1}_{\Phi_\delta,X}\|_2\|PP^\top\|_2\|A^{-1}_{\Phi_\delta,X}\|_2\|R(x)\|_2}{\|P\|_2^2\frac{1}{\lambda_{\max}(A_{\Phi_\delta,X})}} + \frac{\frac{1}{\lambda_{\min}(A_{\Phi_\delta,X})}\|P\|_2}{\|P\|_2^2\frac{1}{\lambda_{\max}(A_{\Phi_\delta,X})}}.$$

Since $PP^\top$ is symmetric and

$$PP^\top = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in M_N(\mathbb{R})$$

then

$$\|PP^\top\|_2^2 = \lambda_{\max}((PP^\top)^\top PP^\top) = \lambda_{\max}((PP^\top)^2) = \lambda_{\max}(PP^\top)^2 = N^2,$$

where the last equality holds because $\mathrm{rank}(PP^\top) = N-1$ and $(1,\dots,1)^\top \in \mathbb{R}^N$ is an eigenvector with eigenvalue N.

We can conclude

$$\|\widetilde{u}(x) - u(x)\|_2 \le \frac{\lambda_{\max}(A_{\Phi_\delta,X})\Phi(0)(1+c_\gamma c_{qu})^d}{\lambda_{\min}(A_{\Phi_\delta,X})^2} + \frac{\lambda_{\max}(A_{\Phi_\delta,X})}{\lambda_{\min}(A_{\Phi_\delta,X})}\frac{1}{\sqrt{N}},$$

that is bounded. We note that the constants (polynomials of degree 0) that appears in the definition of the cardinal functions $\{\widetilde{u}_1,\dots,\widetilde{u}_N\}$ converge to 0 as $h_{X,\Omega} \to 0^+$. We can also state that the burden of reproducing the constants is distributed over all the basis functions.

Moreover

$$N\Phi(0) = \mathrm{trace}\begin{pmatrix} A_{\Phi_\delta,X} & P \\ P^\top & 0 \end{pmatrix} = \mathrm{trace}(A_{\Phi_\delta,X}) = \sum_{j=1}^{N} \lambda_j(A_{\Phi_\delta,X}) > 0$$

and the two matrices can not have the same eigenvalues because they are both invertible.

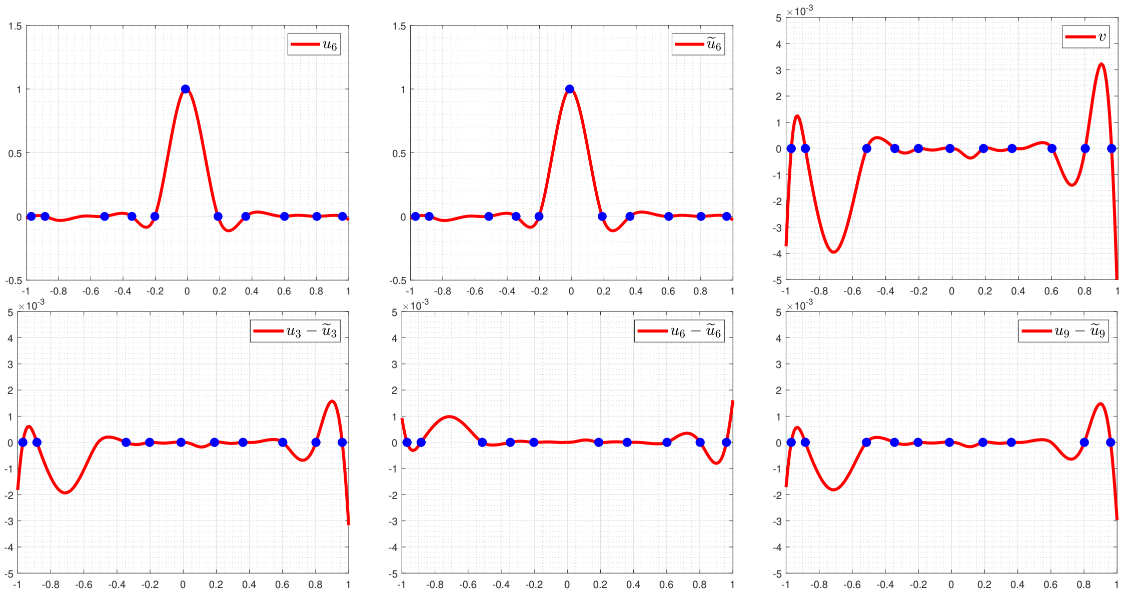


Figure 7.2: Comparison of cardinal functions of the interpolation process with the compactly supported Wendland's function $\phi_{1,1}$ on 11 uniformly perturbed equispaced nodes in $[-1,1]$ with $\delta = 5h_{X,\Omega}$.

Figure 7.2 confirms the theoretical results, in fact the function v tends to zero. We can also point out that the difference between $u(x)$ and $\widetilde{u}(x)$ it is not only bounded but also seems to approach zero.

Now we consider well-known methods to establish convergence. To prove the conjecture in Theorem 7.11 we can not use standard error estimates not even to prove a bound on the difference between 1 and $s_{1,X,\delta}$. Theorem 4.3 gives us for $f \in \mathcal{N}_{\Phi_\delta}(\Omega)$ and $x \in \Omega$

$$|f(x) - s_{f,X,\delta}(x)| \leq P_{\Phi_\delta,X}(x)|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}.$$

Under some conditions we can prove that there exists an increasing function $r :]0, +\infty[\to \mathbb{R}$ such that $P_{\Phi,X}(x) \leq Kr(h_{X,\Omega})$ for $x \in \Omega$ and K depends on Ω only via its cone condition angle θ (Definition 3.2). We note that the angle in the cone condition does not change if we scale the domain Ω. If Ω satisfies an interior cone condition with angle ϑ and radius r then Ω/δ satisfies an interior cone condition with angle ϑ and radius r/δ, indeed by imposing $\xi(x/\delta) = \xi(x)$ then

$$\frac{x}{\delta} + \lambda y \in \frac{\Omega}{\delta} \Leftrightarrow x + \delta\lambda y \in \Omega,$$

that holds if $\lambda \in [0, r/\delta]$, $\|y\|_2 = 1$ and $\left\langle y, \xi\left(\frac{x}{\delta}\right)\right\rangle = \langle y, \xi(x)\rangle \geq \cos(\vartheta)$.

Before continuing we check that $\frac{h_{X,\Omega}}{\delta} = h_{\frac{X}{\delta},\frac{\Omega}{\delta}}$,indeed

$$h_{\frac{X}{\delta},\frac{\Omega}{\delta}} = \sup_{x \in \frac{\Omega}{\delta}} \min_{1 \leq j \leq N} \left\| x - \frac{x_j}{\delta} \right\|_2 = \sup_{x \in \Omega} \min_{1 \leq j \leq N} \left\| \frac{x}{\delta} - \frac{x_j}{\delta} \right\|_2 = \frac{h_{X,\Omega}}{\delta}.$$

From equation (4.9) we obtain for $x \in \Omega$

$$\begin{aligned} |f(x) - s_{f,X,\delta}(x)| &\leq P_{\Phi,X/\delta}(x/\delta)|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq Kr(h_{X/\delta,\Omega/\delta})|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)} = \\ &= Kr\left(\frac{h_{X,\Omega}}{\delta}\right)|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq Kr\left(\frac{1}{\gamma c_\gamma}\right)|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}. \end{aligned}$$

This proves that in general if the scaling factor δ is proportional to the fill-distance $h_{X,\Omega}$ we do not have convergence. For our purposes it would be enough to have a uniform bound on δ of $|\cdot|_{\mathcal{N}_{\Phi_\delta}(\Omega)}$, unfortunately this is not the case because from equation (2.23) we have

$$\|f\|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq (2\pi)^{-\frac{d}{4}} C_s c_1^{-\frac{1}{2}} \delta^{-\frac{d}{2}} \|f\|_{H^s(\Omega)} \tag{7.13}$$

where C_s is the norm of the extension operator in Theorem 2.18 and c_1 comes from the initial hypothesis stated in Theorem 1.26. The right-hand side of the inequality in equation (7.13) diverges as $\delta \to 0^+$.

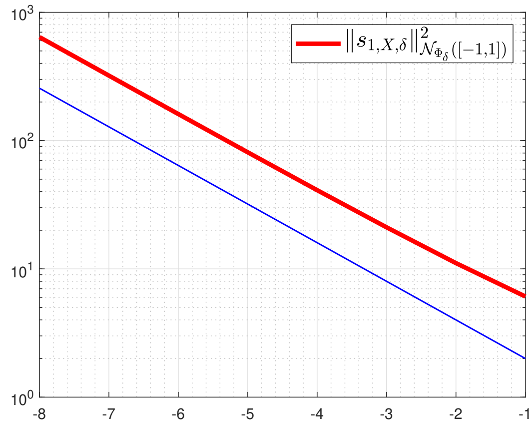


Figure 7.3: Approximation of $\|1\|^2_{\mathcal{N}_{\Phi_\delta}([-1,1])}$ as $\delta \in \left[\frac{1}{2^8}, \frac{1}{2}\right]$ with $\|s_{1,X,\delta}\|^2_{\mathcal{N}_{\Phi_\delta}([-1,1])}$. X consists of 2500 equispaced nodes in $[-1,1]$ and Φ is the Wendland's function $\phi_{1,1}$. The blue line allows us to identify the correct slope (equation (7.13)).

An interesting observation which can be drawn from Figure 7.3 is that the inequality in equation (7.13) reaches a sharp inequality, so we can not bound $\|1\|_{\mathcal{N}_{\Phi_\delta}([-1,1])}$ as δ decreases. This observation makes sense because convergence in the uniform norm implies convergence in the native space. By recalling that $\overline{F_{\Phi_\delta}(\Omega)} = \mathcal{N}_{\Phi_\delta}(\Omega)$ (equation (1.11)) then if $g \in \mathcal{N}_{\Phi_\delta}(\Omega)$ we can choose $\widetilde{g} \in F_{\Phi_\delta}(\Omega)$ such that $|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}|g-\widetilde{g}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq \frac{\varepsilon}{3}$. We can compute

$$\begin{aligned}
&|\langle f, g\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} - \langle s_{f,X,\delta}, g\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)}| = \\
=&|\langle f, g-\widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} + \langle f, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} - \langle s_{f,X,\delta}, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} - \langle s_{f,X,\delta}, g-\widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)}| \leq \\
\leq&|f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}|g-\widetilde{g}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} + |\langle f, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} - \langle s_{f,X,\delta}, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)}| + |f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}|g-\widetilde{g}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq \\
\leq&\frac{2}{3}\varepsilon + |\langle f, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} - \langle s_{f,X,\delta}, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)}|,
\end{aligned}$$

where we use $|s_{f,X,\delta}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq |f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}$ (Theorem 2.9). If $\widetilde{g} = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j)$ then by the reproduction property of $\mathcal{N}_{\Phi_\delta}(\Omega)$ (Definition 1.7)

$$\langle f, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} - \langle s_{f,X,\delta}, \widetilde{g}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} = \sum_{j=1}^N \alpha_j (f(x_j) - s_{f,X,\delta}(x_j)) \xrightarrow[h_{X,\Omega}\to 0]{} 0,$$

that holds for Theorem 4.3 (δ is fixed). Since Riesz Theorem works on $\mathcal{N}_{\Phi_\delta}(\Omega)$ (Theorem 1.6) then $s_{f,X,\delta} \rightharpoonup f$ (convergence in the weak topology) that implies [54]

$$\limsup_{h_{X,\Omega}\to 0} |s_{f,X,\delta}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq |f|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq \liminf_{h_{X,\Omega}\to 0} |s_{f,X,\delta}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \implies \lim_{h_{X,\Omega}\to 0} |s_{f,X,\delta}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} = |f|_{\mathcal{N}_{\Phi_\delta}(\Omega)}.$$

We have also $s_{f,X,\delta} \xrightarrow[h_{X,\Omega}\to 0]{|\cdot|_{\mathcal{N}_{\Phi_\delta}(\Omega)}} f$ because

$$|f - s_{f,X,\delta}|^2_{\mathcal{N}_{\Phi_\delta}(\Omega)} = |f|^2_{\mathcal{N}_{\Phi_\delta}(\Omega)} - 2\langle f, s_{f,X,\delta}\rangle_{\mathcal{N}_{\Phi_\delta}(\Omega)} + |s_{f,X,\delta}|^2_{\mathcal{N}_{\Phi_\delta}(\Omega)} \xrightarrow[h_{X,\Omega}\to 0]{} 0$$

by weak convergence.

We note that also the restriction of the functions we want to approximate does not lead to our goal, in fact equation (4.11) gives us under the same conditions on the growth of the power function

$$\begin{aligned}\|f - s_{f,X,\delta}\|_{L^\infty(\Omega)} &\leq C^2 r (h_{X,\Omega})^2 \|T_\delta^{-1}(f)\|_{L^2(\Omega)} \sqrt{\mathcal{L}(\Omega)} \leq \\ &\leq C^2 r \left(\frac{h_{X,\Omega}}{\delta}\right)^2 \|T_\delta^{-1}(f)\|_{L^2(\Omega)} \sqrt{\mathcal{L}(\Omega)} \leq \\ &\leq C^2 r \left(\frac{1}{\gamma c_\gamma}\right)^2 \|T_\delta^{-1}(f)\|_{L^2(\Omega)} \sqrt{\mathcal{L}(\Omega)},\end{aligned}$$

where $\mathcal{L}$ is the Lebesgue measure on $\mathbb{R}^d$ and the integral operator T_δ (equation (2.18)) depends on δ. Also in this case we would need a uniform bound on δ and a necessary condition to use this inequality is a restriction on the functions to approximate:

$$f \in \bigcap_{0<\delta\leq 1} T_\delta(L^2(\Omega)).$$

In Theorem 6.3 and Theorem 6.6 we proved some optimality properties of the interpolation process with (conditionally) positive definite kernel, in particular it holds that

$$|s_{1,X,\delta}|_{\mathcal{N}_{\Phi_\delta}(\Omega)} = \min\{|s|_{\mathcal{N}_{\Phi_\delta}(\Omega)} : s \in \mathcal{N}_{\Phi_\delta}(\Omega) \text{ such that } s(x_j) = 1 \text{ for } j = 1, \dots, N\}.$$

Since from Theorem 1.26 and Theorem 2.21 we have $\mathcal{N}_{\Phi_\delta}(\Omega) = H^s(\Omega)$ and the norms $\|\cdot\|_{\mathcal{N}_{\Phi_\delta}(\Omega)}$ and $\|\cdot\|_{H^s(\Omega)}$ are equivalent if $\|s_{1,X,\delta}\|_{\mathcal{N}_{\Phi_\delta}(\Omega)}$ is optimal then we expect that also $\|s_{1,X,\delta}\|_{H^s(\Omega)}$ is almost optimal, which it might mean that the derivatives involved are almost flat and the interpolant $s_{1,X,\delta}$ does not have too much oscillations.

When the native space of a positive definite kernel is also a Sobolev space with equivalent norms the error estimates of the interpolation process can also be obtained with the so-called sampling inequalities [50].

Suppose that $f \in H^s(\Omega)$ and $\Omega \subseteq \mathbb{R}^d$ is a bounded Lipschitz domain then a typical sampling inequality let us to achieve

$$\|f\|_{L^\infty(\Omega)} \leq C h_{X,\Omega}^{s-d/2} |u|_{H^s(\Omega)} + \|f_{|X}\|_\infty. \tag{7.14}$$

We can apply the inequality of equation (7.14) in our context, since $f - s_{f,X,\delta} \in H^s(\Omega)$, obtaining

$$\|f - s_{f,X,\delta}\|_{L^\infty(\Omega)} \leq C h_{X,\Omega}^{s-d/2} |f - s_{f,X,\delta}|_{H^s(\Omega)} \leq C h_{X,\Omega}^{s-d/2} \|f - s_{f,X,\delta}\|_{H^s(\Omega)},$$

because $(f - s_{f,X,\delta})_{|X} = 0$. With the optimality properties of the interpolant $s_{f,X,\delta}$ we can get rid of $s_{f,X,\delta}$ in $\|f - s_{f,X,\delta}\|_{H^s(\Omega)}$, indeed

$$
\begin{aligned}
\|f - s_{f,X,\delta}\|_{H^s(\Omega)} &\leq (2\pi)^{\frac{d}{4}} c_2^{\frac{1}{2}} \delta^{\frac{d}{2}-s} \|f - s_{f,X,\delta}\|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq \\
&\leq (2\pi)^{\frac{d}{4}} c_2^{\frac{1}{2}} \delta^{\frac{d}{2}-s} (\|f\|_{\mathcal{N}_{\Phi_\delta}(\Omega)} + \|s_{f,X,\delta}\|_{\mathcal{N}_{\Phi_\delta}(\Omega)}) \overset{\text{Theorem 6.3}}{\leq} \\
&\leq 2(2\pi)^{\frac{d}{4}} c_2^{\frac{1}{2}} \delta^{\frac{d}{2}-s} \|f\|_{\mathcal{N}_{\Phi_\delta}(\Omega)} \leq 2C_s c_1^{-\frac{1}{2}} \delta^{-\frac{d}{2}} c_2^{\frac{1}{2}} \delta^{\frac{d}{2}-s} \|f\|_{H^s(\Omega)} = \\
&= 2C_s c_1^{-\frac{1}{2}} c_2^{\frac{1}{2}} \delta^{-s} \|f\|_{H^s(\Omega)},
\end{aligned}
$$

where we used equation (2.23) and equation (2.24). The constants involved comes from Theorem 1.26 (c_1, c_2) and Theorem 2.18 (C_s). Finally we obtain

$$
\begin{aligned}
\|f - s_{f,X,\delta}\|_{L^\infty(\Omega)} &\leq C h_{X,\Omega}^{s-d/2} 2C_s c_1^{-\frac{1}{2}} c_2^{\frac{1}{2}} \delta^{-s} \|f\|_{H^s(\Omega)} = \\
&= 2CC_s c_1^{-\frac{1}{2}} c_2^{\frac{1}{2}} \left(\frac{h_{X,\Omega}}{\delta}\right)^s h_{X,\Omega}^{-d/2} \|f\|_{H^s(\Omega)} \leq \\
&\leq 2CC_s c_1^{-\frac{1}{2}} c_2^{\frac{1}{2}} \left(\frac{1}{\gamma c_\gamma}\right)^s h_{X,\Omega}^{-d/2} \|f\|_{H^s(\Omega)}.
\end{aligned}
\tag{7.15}
$$

As we have seen in equation (7.13) also the right-hand side of the inequality in equation (7.15) diverges as $h_{X,\Omega} \to 0^+$. Also in this case when the scaling factor δ is proportional to the fill-distance $h_{X,\Omega}$ we can not prove convergence or a bound for the approximation error.

In any case it is worth noting that from equation (7.15) there exists K independent of $f \in H^s(\Omega)$ such that

$$
\|f - s_{f,X,\delta}\|_{L^\infty(\Omega)} \leq K h_{X,\Omega}^{s-d/2} \|f\|_{H^s(\Omega)}, \tag{7.16}
$$

that proves convergence (when δ is fixed) but it suffers the so-called curse of dimension. We have seen in Theorem 3.6 that when X is quasi-uniform then $h_{X,\Omega}$ is proportional to $(\frac{1}{N})^{\frac{1}{d}}$ where $N = \#X$ that with equation (7.16) gives us

$$
\|f - s_{f,X,\delta}\|_{L^\infty(\Omega)} \leq K \left(\frac{1}{N}\right)^{\frac{s}{d}-\frac{1}{2}}.
$$

When the space dimension d becomes larger, with the same number of sampling points, we have a slower convergence. Also since the smoothness of the Sobolev space $H^s(\Omega)$ depends on s then the inequality $2s > d$ is restrictive for large d.

7.1.3 Constants analysis for exponential decaying cardinal functions

Numerical evidences (Figure 7.1) show that when c_γ grows then $s_{1,X,\delta}$ has less oscillations but the constants that characterize the exponential decay of the cardinal functions $\{u_1, \ldots, u_N\}$ (Theorem 7.9) do not have a good behaviour. Since

$$
\nu(\Phi, d, c_\gamma c_{qu}) = -\frac{1}{2} \log(\mu(\Phi, d, c_\gamma c_{qu}))
$$

with

$$\mu(\Phi, d, c_\gamma c_{qu}) = \left(\frac{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} - 1}{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} + 1} \right)^{\frac{1}{R(c_\gamma c_{qu})}}$$

and $R(c_\gamma c_{qu}) = \max\{c_\gamma c_{qu}, 4\}\sqrt{d}$, the constant that rules how flat is the exponential can be written as

$$\nu(\Phi, d, c_\gamma c_{qu}) = -\frac{1}{2} \frac{1}{\max\{c_\gamma c_{qu}, 4\}\sqrt{d}} \log \left(\frac{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} - 1}{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} + 1} \right),$$

which permits us to say that $\nu(\Phi, d, c_\gamma c_{qu})$ decreases when c_γ grows, because $\frac{1}{\max\{c_\gamma c_{qu}, 4\}\sqrt{d}}$ is decreasing as function of c_γ and also

$$-\log \left(\frac{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} - 1}{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} + 1} \right) > 0$$

is positive and decreasing ($g(x) = \frac{x-1}{x+1}$ is monotonically increasing).

Because of the exponential decay of the cardinal functions it seems natural to study

$$\sum_{j=1}^{N} u_j(x) \quad \text{for } x \in \Omega$$

as

$$\sum_{j:\|x-x_j\| \le q_X} u_j(x) + \sum_{j:\|x-x_j\| > q_X} u_j(x)$$

With similar computations of the proof of Theorem 7.1 we obtain

$$\sum_{j:\|x-x_j\| > q_X} u_j(x) \ge -3^d K(\Phi, d, c_\gamma c_{qu}) \sum_{n=1}^{+\infty} (n+1)^{d-1} e^{-\nu(\Phi,d,c_\gamma c_{qu})n},$$

which does not give an exploitable lower bound because when c_γ becomes larger the exponential becomes flatter ($\nu(\Phi, d, c_\gamma c_{qu})$ decreases), so we can not get a better lower-bound for the sum of the cardinal functions that are centered far from the point $x \in \Omega$.

It is worth noting that from Theorem 7.9 the constant $K(\Phi, d, c_\gamma c_{qu})$ seems to grow at least exponentially with c_γ. This is not true because

$$\nu(\Phi, d, c_\gamma c_{qu}) c_\gamma c_{qu} = -\frac{1}{2} \frac{1}{\max\{c_\gamma c_{qu}, 4\}\sqrt{d}} \log \left(\frac{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} - 1}{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} + 1} \right) c_\gamma c_{qu}$$

that becomes with $c_\gamma c_{qu} \ge 4$

$$-\frac{1}{2} \frac{1}{\sqrt{d}} \log \left(\frac{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} - 1}{\sqrt{C(\Phi,d)\Phi(0)(c_\gamma c_{qu})^{2s-d}(1+c_\gamma c_{qu})^d} + 1} \right),$$

which converges to 0^+ as $c_\gamma \to +\infty$.

Another insight that can be derived from Theorem 4.1 is that

$$R(x)^\top u(x) = R(x)^\top A_{\Phi_\delta,X}^{-1} R(x) \geq \frac{\|R(x)\|_2^2}{\lambda_{\max}(A_{\Phi_\delta,X})}.$$

Since $R(x) = (\Phi_\delta(x-x_1), \ldots, \Phi_\delta(x-x_N))$ then

$$R(x)^\top u(x) = \sum_{j=1}^N \Phi_\delta(x-x_j)u_j(x) = \sum_{j:\|x-x_j\|_2\leq\delta} \Phi_\delta(x-x_j)u_j(x),$$

so

$$\|R(x)\|_2 \left(\sum_{j:\|x-x_j\|_2\leq\delta} u_j^2(x) \right)^{\frac{1}{2}} \geq \frac{\|R(x)\|_2^2}{\lambda_{\max}(A_{\Phi_\delta,X})},$$

which becomes

$$\left(\sum_{j:\|x-x_j\|_2\leq\delta} u_j^2(x) \right)^{\frac{1}{2}} \geq \frac{\|R(x)\|_2}{\lambda_{\max}(A_{\Phi_\delta,X})}.$$

If $x \in \Omega$ then there exists $j \in \{1, \ldots, N\}$ such that $\|x - x_j\|_2 \leq h_{X,\Omega} \leq \delta$, thus we can obtain with Theorem 7.3

$$(1 + c_\gamma c_{qu})^d \max_{j:\|x-x_j\|_2\leq\delta} \{|u_j(x)|\} \geq \frac{\min_{y\in B(0,1/(\gamma c_\gamma))} |\Phi(y)|}{\lambda_{\max}(A_{\Phi_\delta,X})}$$

and with Theorem 7.5 we can conclude

$$\max_{j:\|x-x_j\|_2\leq\delta} \{|u_j(x)|\} \geq \frac{\min_{y\in B(0,1/(\gamma c_\gamma))} |\Phi(y)|}{(1 + c_\gamma c_{qu})^{2d}\Phi(0)}.$$

7.2 Fast decaying polynomial reproduction

Imitating the Definition 3.1 we want to relax the hypothesis of compact support but to hope for the same kind of convergence we must guarantee some sort of locality.

Definition 7.4 *Fix a decreasing function $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $\lim_{n\to+\infty} \frac{\varphi(n+1)}{\varphi(n)}$ exists and it is strictly smaller than 1. A process that defines for every set $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ a family of functions $u_j = u_j^X : \Omega \to \mathbb{R}$ for $1 \leq j \leq N$ provides fast decaying polynomial reproduction of degree ℓ on Ω with respect to φ if there exists a constant C, h_0 such that*

- $\sum_{j=1}^N p(x_j)u_j = p$ *for each* $p \in \pi_\ell(\mathbb{R}^d)$,
- $|u_j(x)| \leq C\varphi\left(\frac{\|x-x_j\|_2}{q_X}\right)$ *for all* $x \in \Omega$ *and* $j = 1, \ldots, N$

are satisfied for all X with $h_{X,\Omega} \leq h_0$.

By carefully reading Theorem 7.1 we can obtain a similar result.

Theorem 7.14 *Suppose that* $X = \{x_1, \ldots, x_N\} \subseteq \mathbb{R}^d$ *is an arbitrary data set and* $\{u_1, \ldots, u_N\}$ *is given by a fast decaying polynomial reproduction method with respect to* φ *(Definition 7.4) then for every* $\ell \in \mathbb{N}$ *there exists a constant* $K = K(\ell, C, \varphi, d)$ *such that*

$$\sum_{j=1}^{N} \|x - x_j\|_2^\ell |u_j(x)| \leq K h_{X,\Omega}^\ell \quad \textit{for } x \in \Omega.$$

with

$$K = 3^d C \sum_{n=0}^{+\infty} (n+1)^{d+\ell-1} \varphi(n).$$

If $f : \Omega \to \mathbb{R}$ is a function we can build with Definition 7.4 a stable quasi-interpolation process by

$$z_{f,X} = \sum_{j=1}^{N} f(x_j) u_j(x) \tag{7.17}$$

with Lebesgue function bounded by $3^d C \sum_{n=0}^{+\infty} (n+1)^{d-1} \varphi(n)$. The quasi-interpolation method of Definition 7.1 respect the Definition 7.4 for quasi-uniform data sets. Since we dropped the hypothesis of quasi-uniformity on the distribution of X to get a convergence result we need stronger assumptions on the function f to reconstruct from X.

Theorem 7.15 *Suppose that* $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ *is an arbitrary data set and* $\{u_1, \ldots, u_N\}$ *is given by a fast decaying polynomial reproduction process with respect to* φ *of order* $m \in \mathbb{N}$. *Define* Ω^* *to be the closure of the convex hull of* Ω. *If* $f \in \mathcal{C}^{m+1}(\Omega^*)$ *then there exists a constant* $K = K(C, \varphi, d, m)$ *such that*

$$\|f - z_{f,X}\|_{L^\infty(\Omega)} \leq K h_{X,\Omega}^{m+1} \|f\|_{\mathcal{C}^{m+1}(\Omega^*)},$$

for $h_{X,\Omega} \leq h_0$.

Proof

Let p be an arbitrary polynomial from $\pi_m(\mathbb{R}^d)$. Using Definition 7.4 we obtain

$$\begin{aligned} |f(x) - z_{f,X}(x)| &\leq |f(x) - p(x)| + \left| p(x) - \sum_{j=1}^{N} f(x_j) u_j(x) \right| \leq \\ &\leq |f(x) - p(x)| + \sum_{j=1}^{N} |p(x_j) - f(x_j)| |u_j(x)| \end{aligned}$$

To end the proof we choose p to be the Taylor polynomial of f around x of order m. For $y \in \Omega$ there exists $\xi \in \Omega^*$ such that

$$f(y) - \sum_{|\alpha| \leq m} \frac{D^\alpha f(x)}{\alpha!} (y - x)^\alpha = \sum_{|\alpha| = m+1} \frac{D^\alpha f(\xi)}{\alpha!} (y - x)^\alpha.$$

Remarking that $|(y-x)^\alpha| = \prod_{i=1}^d |y_i - x_i|^{\alpha_i} \le \prod_{i=1}^d \|y-x\|^{\alpha_i} \le \|y-x\|^{|\alpha|}$ we can conclude for $j = 1, \dots, N$

$$|p(x_j) - f(x_j)| \le \sum_{|\alpha|=m+1} \frac{\|D^\alpha(f)\|_{L^\infty(\Omega^*)}}{\alpha!} \|x_j - x\|_2^{m+1} \le \left(\sum_{|\alpha|=m+1} \frac{1}{\alpha!} \right) |f|_{\mathcal{C}^{m+1}(\Omega^*)} \|x_j - x\|_2^{m+1},$$

that gives us

$$\begin{aligned} |f(x) - z_{f,X}(x)| &\le \left(\sum_{|\alpha|=m+1} \frac{1}{\alpha!} \right) |f|_{\mathcal{C}^{m+1}(\Omega^*)} \sum_{j=1}^N \|x_j - x\|_2^{m+1} |u_j(x)| \\ &\le \left(\sum_{|\alpha|=m+1} \frac{1}{\alpha!} \right) K(m+1, C, \varphi, d) |f|_{\mathcal{C}^{m+1}(\Omega^*)} h_{X,\Omega}^{m+1}, \end{aligned}$$

where $K(m+1, C, \varphi, d)$ comes from the application of Theorem 7.14.

□

7.2.1 An example for polynomial reproduction with fast decay

We want to build an example for Definition 7.4 with a $\mathcal{C}^\infty(\Omega)$ interpolant (equation (7.17)). As local polynomial reproduction with consider a minimization problem, i.e. moving least squares. By recalling Definition 3.3 we have for $x \in \Omega$ that the value $z_{f,X}(x)$ of the moving least squares approximant is given by $z_{f,X}(x) = p^*(x)$ where p^* is the solution of

$$\min \left\{ \sum_{i=1}^N (f(x_i) - p(x_i))^2 e^{-\nu \left(\frac{\|x - x_i\|_2}{\delta} \right)^2} : p \in \pi_m(\mathbb{R}^d) \right\}, \tag{7.18}$$

where $\nu \in \mathbb{R}_{>0}$ is a fixed parameter and instead δ will depend on the data set X. Since the weight function is strictly positive

$$w(x, y) = e^{-\nu \left(\frac{\|x-y\|_2}{\delta} \right)^2} > 0$$

we can prove a similar result of Theorem 3.4.

Theorem 7.16 *Suppose that for every $x \in \Omega$ the set $X = \{x_1, \dots, x_N\}$ is $\pi_m(\mathbb{R}^d)$-unisolvent. In this situation, the problem stated in equation (7.18) is uniquely solvable and the solution $z_{f,X}(x) = p^*(x)$ can be represented as*

$$z_{f,X}(x) = \sum_{i=1}^N f(x_i) a_i^*(x),$$

where the coefficients $a_i^*(x)$ are determined by minimizing the quadratic form

$$\sum_{i=1}^N \frac{a_i(x)^2}{e^{-\nu \left(\frac{\|x - x_i\|_2}{\delta} \right)^2}} \tag{7.19}$$

under the constraints

$$\sum_{i=1}^{N} p(x_i)a_i(x) = p(x), \qquad p \in \pi_m(\mathbb{R}^d). \tag{7.20}$$

Theorem 7.16 let us to characterize the shape of the quasi-interpolant $z_{f,X}$ as

Theorem 7.17 *The functions $\{a_j^*\}_{j=1,\ldots,N}$ of Theorem 7.16 have the form*

$$a_j^*(x) = e^{-\nu\left(\frac{\|x-x_j\|_2}{\delta}\right)^2} \sum_{k=1}^{Q} \lambda_k(x) p_k(x_j)$$

where $\{\lambda_k(x)\}_{k=1,\ldots,Q}$ are the unique solutions of

$$\sum_{k=1}^{Q} \lambda_k(x) \sum_{j=1}^{N} e^{-\nu\left(\frac{\|x-x_j\|_2}{\delta}\right)^2} p_k(x_j)p_\ell(x_j) = p_\ell(x), \qquad 0 \le \ell \le Q. \tag{7.21}$$

Moreover the approximant $z_{f,X} \in \mathcal{C}^\infty(\Omega)$. In this context $\{p_1, \ldots, p_Q\}$ is a basis for $\pi_m(\mathbb{R}^d)$.

For m=0 Theorem 7.17 let us to conclude

$$a_j^*(x) = \frac{e^{-\nu\left(\frac{\|x-x_j\|_2}{\delta}\right)^2}}{\sum_{i=1}^{N} e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \qquad \text{for each } j \in \{1, \ldots, N\}.$$

For $m = 0$ the minimization problem described in equation (7.18) has the solution

$$z_{f,X}(x) = \sum_{j=1}^{N} f(x_j) \underbrace{\frac{e^{-\nu\left(\frac{\|x-x_j\|_2}{\delta}\right)^2}}{\sum_{i=1}^{N} e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}}}_{a_j^*(x)},$$

which is a particular instance of the Shepard approximation method [22] and it reproduces constants exactly.

In Theorem 3.7 we have a similar result for compactly supported basis functions.

Theorem 7.18 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is compact a satisfies an interior cone condition with angle $\vartheta \in]0, \pi/2[$ and radius $r > 0$. Fix $m \in \mathbb{N}$. Let h_0, C_1 and C_2 denote the constants of Theorem 3.2. Suppose that $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ is a quasi-uniform data sets with respect to $c_{qu} > 0$ and $h_{X,\Omega} \le h_0$. Let δ be as in equation (7.2). Then the basis functions $\{a_j^*(x)\}_{j=1,\ldots,N}$ of Theorem 7.17 provide local polynomial reproduction with fast decay (Definition 7.4), with certain constants C, h_0 and function φ that can be derived explicitly.*

Proof
The first property of Definition 7.4 is a consequence of equation (7.18) and Theorem 7.16 that define the moving least squares method.

To prove the second property we bound the following quantity

$$\sum_{i=1}^{N} \frac{|a_i^*(x)|^2}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}}.$$

There exists $\{\widetilde{u}_j(x)\}_{j=1,\dots,N}$ providing local polynomial reproduction (Theorem 3.2) such that $\widetilde{u}_j$ is supported in $\overline{B(x_j, C_2 h_{X,\Omega})}$ for $j = 1, \dots, N$. The minimal property stated in Theorem 7.16 gives

$$\begin{aligned}\sum_{i=1}^{N} \frac{|a_i^*(x)|^2}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} &\leq \sum_{i=1}^{N} \frac{|\widetilde{u}_i(x)|^2}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} = \sum_{i\in\widetilde{I}(x)} \frac{|\widetilde{u}_i(x)|^2}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \leq \\ &\leq \frac{1}{e^{-\nu\left(\frac{C_2 h_{X,\Omega}}{\delta}\right)^2}} \sum_{i\in\widetilde{I}(x)} |\widetilde{u}_i(x)|^2 \leq \\ &\leq \frac{1}{e^{-\nu\left(\frac{C_2 h_{X,\Omega}}{\delta}\right)^2}} \left(\sum_{i\in\widetilde{I}(x)} |\widetilde{u}_i(x)|\right)^2 \leq \\ &\leq \frac{C_1^2}{e^{-\nu\left(\frac{C_2 h_{X,\Omega}}{\delta}\right)^2}} \leq \frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}},\end{aligned}$$

where $\widetilde{I}(x) = \left\{j \in \{1, \dots, N\} : x_j \in \overline{B(x, C_2 h_{X,\Omega})}\right\}$. With the last computation we obtained for $i = 1, \dots, N$

$$\frac{|a_i^*(x)|^2}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \leq \frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}} \Rightarrow |a_i^*(x)|^2 \leq \frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}} e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}$$

that gives us

$$|a_i^*(x)| \leq \sqrt{\frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}}} e^{-\frac{\nu}{2}\left(\frac{\|x-x_i\|_2}{\delta}\right)^2} \leq \sqrt{\frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}}} e^{-\frac{\nu}{2}\left(\frac{\|x-x_i\|_2}{c_\gamma c_{qu} q_X}\right)^2}.$$

Since $e^{-x^2} \leq e e^{-x}$ for $x \in \mathbb{R}$ because

$$-x^2 \leq -x + 1 \Leftrightarrow x \leq x^2 + 1$$

that holds for $x \in [-1, 1]$ and $x \in \mathbb{R} \setminus [-1, 1]$, we obtain

$$|a_i^*(x)| \leq \underbrace{e\sqrt{\frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}}}}_{C} e^{-\sqrt{\frac{\nu}{2}}\frac{\|x-x_i\|_2}{c_\gamma c_{qu} q_X}} = C\varphi\left(\frac{\|x-x_i\|_2}{q_X}\right)$$

for $i = 1, \dots, N$ with

$$\varphi(x) = e^{-\sqrt{\frac{\nu}{2}} \frac{1}{c_\gamma c_{qu}} x}.$$

□

In Theorem 3.7 we ask that X is unisolvent locally, which leads to oversampling, instead in theorem 7.18 X is unisolvent because there are not local arguments.

The same construction that leads to Theorem 7.18 can also be made with

$$w(x, y) = e^{-\nu \frac{\|x-y\|_2}{\delta}} > 0$$

and we obtain

$$C = \sqrt{\frac{C_1^2}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)}}} \quad \text{and} \quad \varphi(x) = e^{-\frac{\nu}{2} \frac{1}{c_\gamma c_{qu}} x}. \tag{7.22}$$

The construction that provides Theorem 7.18 can be made more general if

$$w(x, y) = \varphi\left(\frac{\|x - y\|_2}{q_X}\right) > 0$$

and satisfies the hypothesis in Definition 7.4, obtaining

$$C = \sqrt{\frac{C_1^2}{\varphi(C_2 c_{qu})}} \quad \text{and} \quad \widetilde{\varphi}(x) = \sqrt{\varphi(x)}.$$

For $m = 0$ the minimization problem that gives us the approximant $z_{f,X}(x)$ has the solution

$$z_{f,X}(x) = \sum_{j=1}^{N} f(x_j) \underbrace{\frac{\varphi\left(\frac{\|x-x_j\|_2}{q_X}\right)}{\sum_{i=1}^{N} \varphi\left(\frac{\|x-x_i\|_2}{q_X}\right)}}_{a_j^*(x)}.$$

For all the previous constructions Theorem 7.17 let us to analyse the computational complexity of the method for the evaluation of the approximant at a point $x \in \Omega$. To compute $\{\lambda_1(x), \dots, \lambda_Q(x)\}$ we need to solve a $Q \times Q$ linear system, so the computational cost is $\mathcal{O}(Q^3)$. The cost to build the matrix of the linear system is $\mathcal{O}(NQ^2)$, where N is the number of data sites. To compute $\{a_1^*(x), \dots, a_N^*(x)\}$ we need $\mathcal{O}(NQ)$ because for each basis function we perform a number of multiplications and sums proportional to Q. After all to build the basis functions at a point $x \in \Omega$ the computational cost is

$$\mathcal{O}(Q^3 + NQ^2 + NQ) = \mathcal{O}(N)$$

and we have to add $\mathcal{O}(N)$ to compute the value of the approximant. If we have to compute the value of the approximant at M points the algorithm leads to a computational cost of $\mathcal{O}(NM)$.

In Theorem 7.17 appears a basis for the polynomial space and, even if it is not relevant from a theoretical point of view, a careful choice of the basis can lead to a more stable and efficient implementation.

We point out that also the construction of Theorem 3.7 satisfy the requests of Definition 7.4 because for $i = 1, \dots, N$ we have

$$|a_i^*(x)| \leq \sum_{j=1}^{N} |a_j^*(x)| \leq \widetilde{C}_1,$$

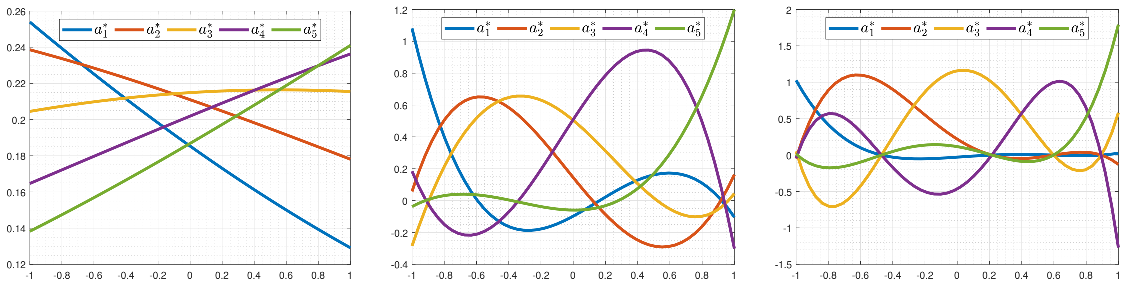
so if we choose $\widetilde{K}$ such that $\widetilde{C}_1 \leq \widetilde{K} e^{-\widetilde{C}_2}$ then for $x \in \Omega \cap B(x_i, \widetilde{C}_2 h_{X,\Omega})$ we obtain

$$|a_i^*(x)| \leq \widetilde{C}_1 \leq \widetilde{K} e^{-\widetilde{C}_2} \leq \widetilde{K} e^{-\frac{\|x-x_i\|_2}{h_{X,\Omega}}} \leq \widetilde{K} e^{-\frac{1}{c_{qu}} \frac{\|x-x_i\|_2}{q_X}}.$$

7.2.2 Numerical test

The purpose of this section is to show some numerical experiments that confirm the theoretical results regarding fast decaying polynomial reproduction methods (Definition (7.4)). As regards basis functions with compact support numerical evidence can be found in [55]. We focus on basis functions which also allow us to construct smooth approximants.

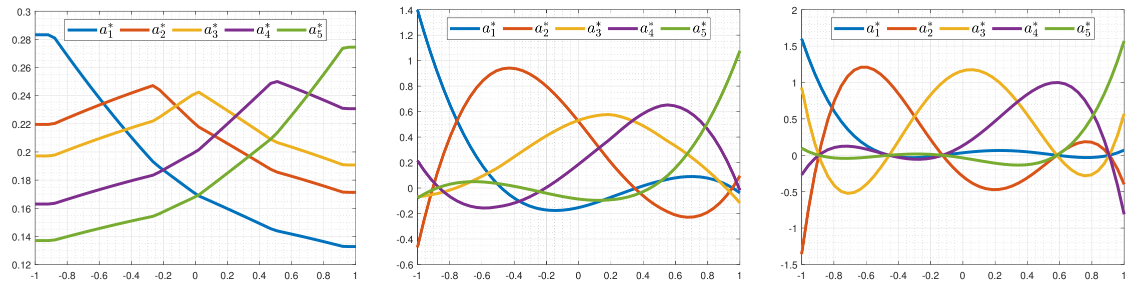
We will start by analyzing the basis functions of Theorem 7.16.



(a) m=0 (b) m=3 (c) m=4

Figure 7.4: Basis functions of Theorem 7.16. The weight function coincides with $e^{-x^2} \in \mathcal{C}^\infty(\mathbb{R})$ and the approxiomation nodes are 5 uniformly perturbed equispaced nodes in $[-1, 1]$. From left to right the basis functions reproduce the polynomials of degree $0, 3$ and 4 respectively. In this numerical test $\delta = 5h_{X,\Omega}$.

We can confirm the results on the differentiability of the basis functions of Theorem 7.17 with Figure 7.4: the basis functions are smooth for each polynomial space they reproduce.



(a) m=0 (b) m=3 (c) m=4

Figure 7.5: Basis functions of Theorem 7.16. The weight function coincides with $e^{-x} \in \mathcal{C}(\mathbb{R})$ and the approxiomation nodes are 5 uniformly perturbed equispaced nodes in $[-1, 1]$. From left to right the basis functions reproduce the polynomials of degree $0, 3$ and 4 respectively. In this numerical test $\delta = 5h_{X,\Omega}$.

Figure 7.5 points out an interesting consideration. Theorem 7.17 guarantees a priori that $\{a_1^*, \ldots, a_5^*\} \subseteq \mathcal{C}([-1, 1])$ and this is confirmed, but when the degree of the polynomials to be reproduced increases then also the smoothness of the basis functions increases from a practical point of view. We have seen in Theorem 7.15 that the smoothness of the weight functions does not affect the convergence rate but a smooth approximant can be useful for applications.

To produce Figure 7.4 and Figure 7.5 we used the system in Theorem 7.17 and as polynomial basis we choose Chebyshev polynomials of the first kind [56].

Now we discuss some experiments to confirm Theorem 7.15 numerically. We approximate different functions on equispaced nodes (quasi-uniform data set) in $[-1, 1]$. We use Theorem 7.17 to get the approximant and the polynomial basis is the Chebyshev polynomial basis of the first kind. The scaling of δ is constant because it does not influence the convergence rate even if appropriate choices of the parameter can improve the stability of the method (Theorem 7.14, equation (7.22)). We fix $\delta = 5h_{X,\Omega}$.

As weight functions in equation (7.18) we used

$$w_1(x, y) = e^{-\left(\frac{|x-y|}{\delta}\right)^2} \quad \text{and} \quad w_2(x, y) = e^{-\frac{|x-y|}{\delta}}.$$

The following numerical experiments approximate the functions

$$\begin{aligned} f_1(x) &= \sin(\pi x), \\ f_2(x) &= 6x^6 + 5x^3 + x^2, \\ f_3(x) &= \frac{1}{1 + 30x^2}. \end{aligned}$$

In the following graphs the blue line allows us to check the correct slope of the approximation error (Theorem 7.15).

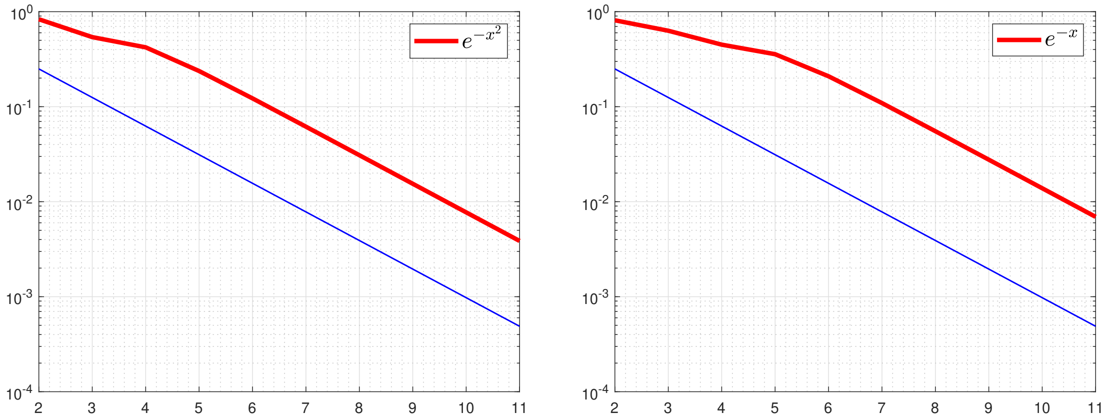


Figure 7.6: Convergence rate of the approximation error $\|f_1 - z_{f_1,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 0.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	5.40e-01	2.37e-01	6.16e-02	1.54e-02	3.86e-03	0
e^{-x}	6.30e-01	3.57e-01	1.09e-01	2.77e-02	6.93e-03	0

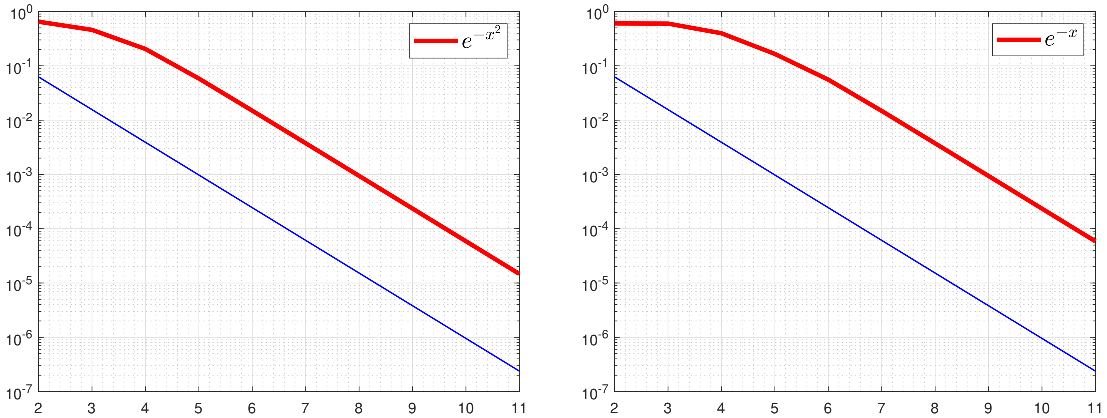


Figure 7.7: Convergence rate of the approximation error $\|f_1 - z_{f_1,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 1.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	4.60e-01	5.85e-02	3.76e-03	2.35e-04	1.47e-05	1
e^{-x}	6.00e-01	1.66e-01	1.48e-02	9.37e-04	5.86e-05	1

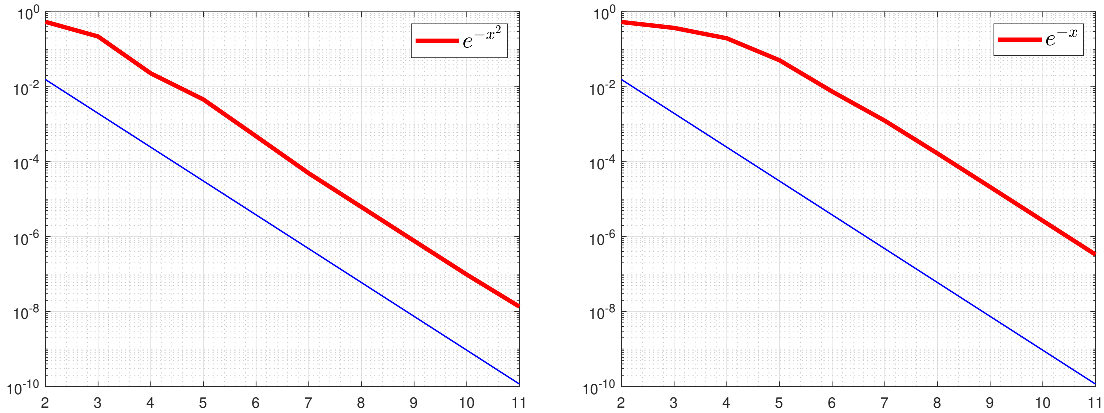


Figure 7.8: Convergence rate of the approximation error $\|f_1 - z_{f_1,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 2.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	2.20e-01	4.58e-03	4.91e-05	7.73e-07	1.35e-08	2
e^{-x}	3.72e-01	5.10e-02	1.24e-03	2.12e-05	3.33e-07	2

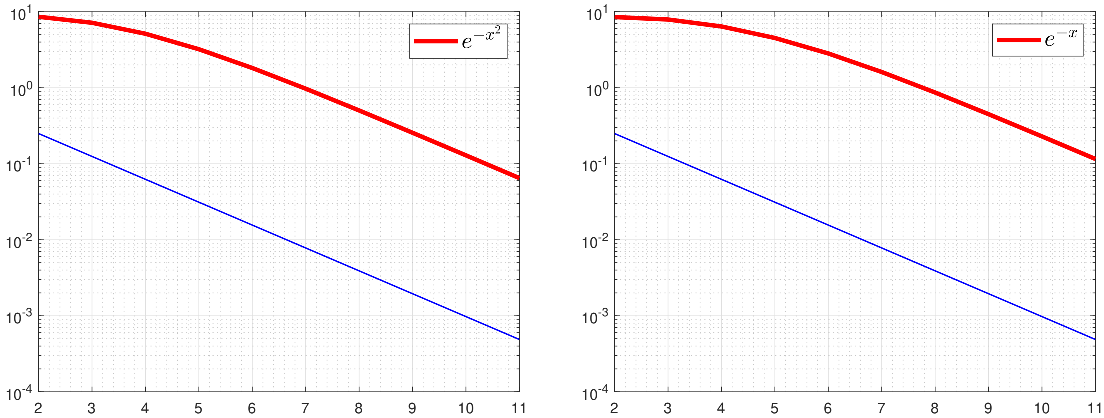


Figure 7.9: Convergence rate of the approximation error $\|f_2 - z_{f_2,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 0.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	7.17e+00	3.22e+00	9.74e-01	2.56e-01	6.49e-02	0
e^{-x}	7.91e+00	4.52e+00	1.61e+00	4.50e-01	1.16e-01	0

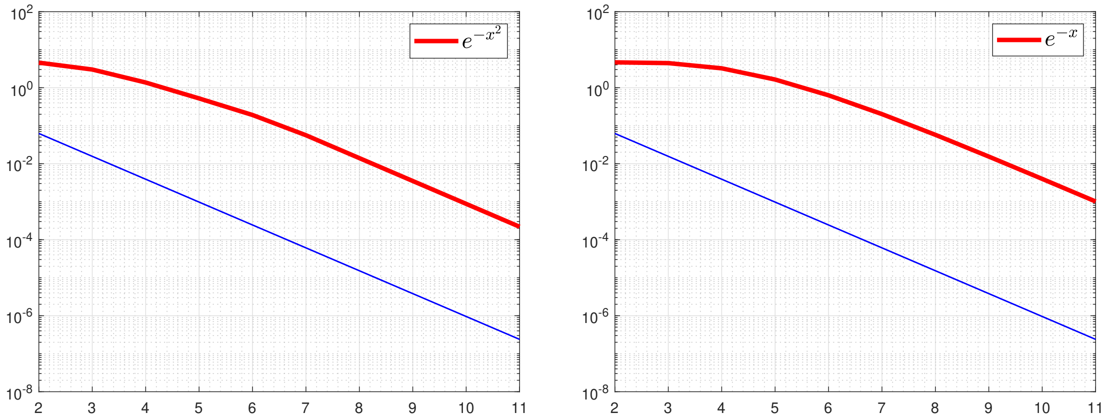


Figure 7.10: Convergence rate of the approximation error $\|f_2 - z_{f_2,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 1.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	3.01e+00	5.22e-01	5.62e-02	3.51e-03	2.19e-04	1
e^{-x}	4.44e+00	1.64e+00	2.02e-01	1.54e-02	1.01e-03	1

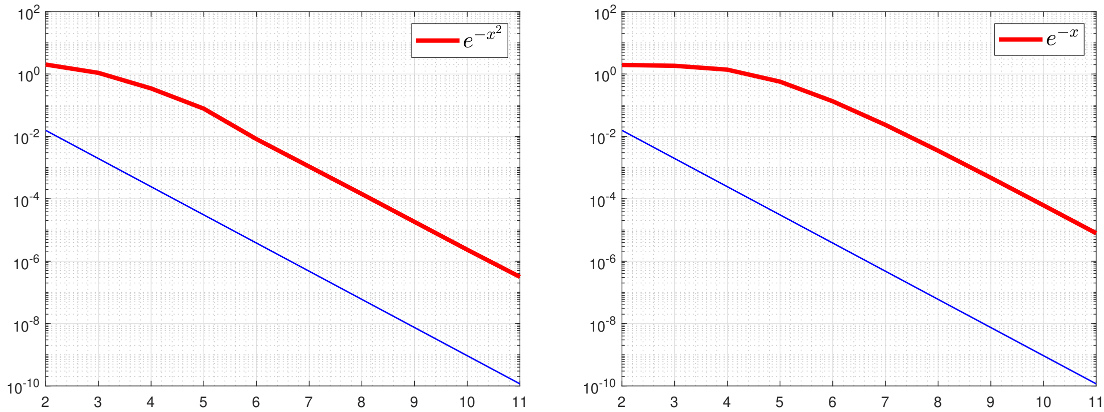


Figure 7.11: Convergence rate of the approximation error $\|f_2 - z_{f_2,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 2.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	1.09e+00	7.76e-02	1.08e-03	1.82e-05	3.14e-07	2
e^{-x}	1.84e+00	5.70e-01	2.34e-02	4.73e-04	7.87e-06	2

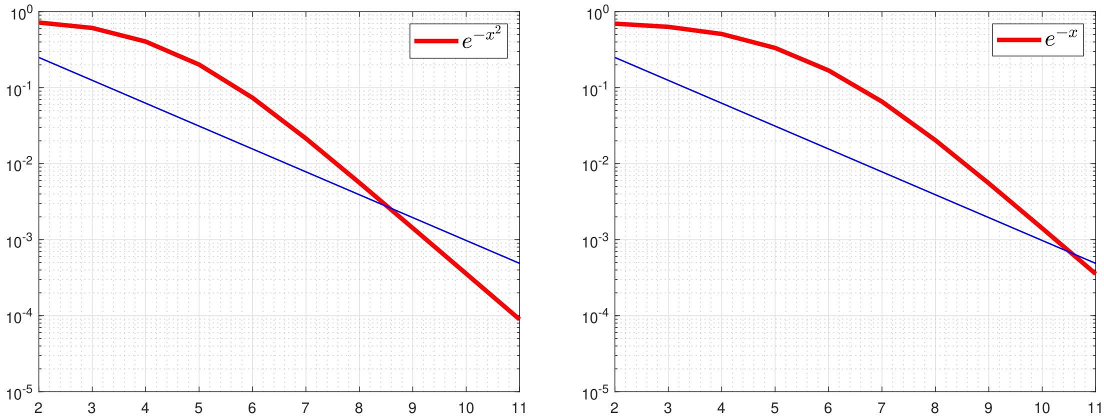


Figure 7.12: Convergence rate of the approximation error $\|f_3 - z_{f_3,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 0.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	6.11e-01	2.01e-01	2.15e-02	1.42e-03	8.94e-05	0
e^{-x}	6.32e-01	3.34e-01	6.56e-02	5.52e-03	3.56e-04	0

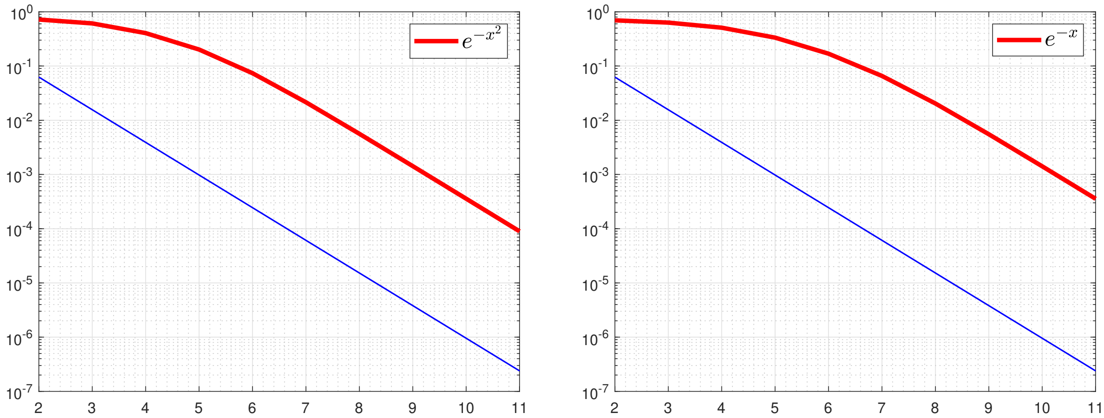


Figure 7.13: Convergence rate of the approximation error $\|f_3 - z_{f_3,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 1.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	6.11e-01	2.01e-01	2.15e-02	1.42e-03	8.94e-05	1
e^{-x}	6.32e-01	3.34e-01	6.56e-02	5.52e-03	3.56e-04	1

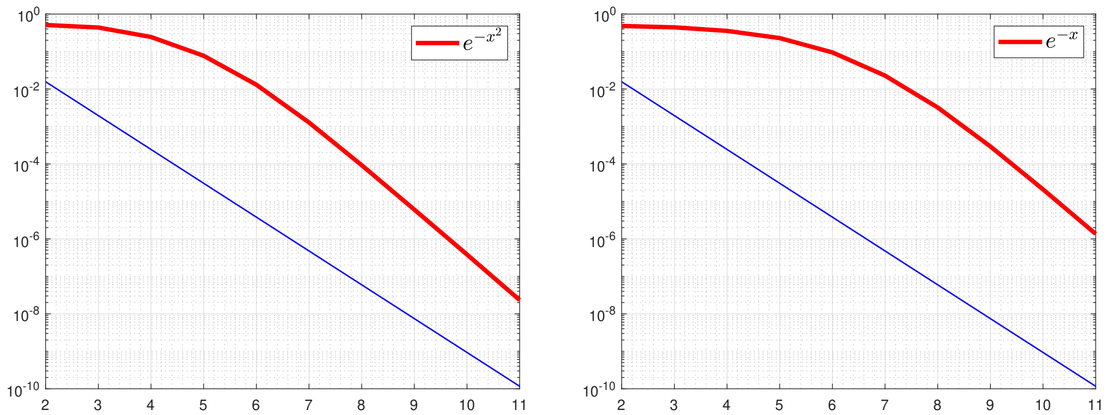


Figure 7.14: Convergence rate of the approximation error $\|f_3 - z_{f_3,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^{12}+1]$. The approximation method reproduces exactly polynomials of degree 2.

Nodes	17	65	257	1025	4097	Degree
e^{-x^2}	4.38e-01	7.72e-02	1.28e-03	6.05e-06	2.33e-08	2
e^{-x}	4.43e-01	2.29e-01	2.28e-02	2.94e-04	1.36e-06	2

Numerical experiments confirm the statement of Theorem 7.15 because when the number of approximation nodes increases, if the method reproduces exactly $\pi_m(\mathbb{R}^d)$, then the convergence rate stabilizes at the value $\mathcal{O}(h_{X,\Omega}^{m+1})$.

With Figure 7.12 and Figure 7.14 it is worth noting that the convergence rate can be even better, as happens when we approximate the Runge function f_3.

7.3 Approximation with the 1-norm

The approximation scheme in section 7.2.1 is convergent but in general the basis functions do not have compact support. This feature has negative repercussions when we compute the value of the approximant (we have to do N multiplications). Moreover, the system in the Theorem 7.17 is dense and could lead to numerical instability.

We want to construct a new approximant by imitating Theorem 7.16 and substituting a weighted 2-norm with a weighted 1-norm.

We want an approximant $z_{f,X}$ with the following form

$$z_{f,X}(x) = \sum_{i=1}^{N} f(x_i)a_i^*(x),$$

where the coefficients $a_i^*(x)$ are determined by minimizing

$$\sum_{i=1}^{N} \frac{|a_i(x)|}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \tag{7.23}$$

under the constraints

$$\sum_{i=1}^{N} p(x_i)a_i(x) = p(x), \qquad p \in \pi_m(\mathbb{R}^d).$$

Since the optimization problem in equation (7.23) is a feasible bounded linear program because X is $\pi_m(\mathbb{R}^d)$-unisolvent, the approximation method is well-defined and it reproduces $\pi_m(\mathbb{R}^d)$ exactly. Our next step is to prove the same converge rate of Theorem 7.15.

Theorem 7.19 *Suppose that $\Omega \subseteq \mathbb{R}^d$ is compact a satisfies an interior cone condition with angle $\vartheta \in]0, \pi/2[$ and radius $r > 0$. Fix $m \in \mathbb{N}$. Let h_0, C_1 and C_2 denote the constants of Theorem 3.2. Suppose that $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ is a quasi-uniform data sets with respect to $c_{qu} > 0$ and $h_{X,\Omega} \le h_0$. Let δ be as in equation (7.2). Then the basis functions $\{a_j^*(x)\}_{j=1,\ldots,N}$ of equation 7.23 provide local polynomial reproduction with fast decay (Definition 7.4), with certain constants C, h_0 and function φ that can be derived explicitly.*

Proof

The first property of Definition 7.4 is a consequence of equation (7.23) that defines the optimization problem and its constraints.

To prove the second property we bound the following quantity

$$\sum_{i=1}^{N} \frac{|a_i^*(x)|}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}}.$$

There exists $\{\widetilde{u}_j(x)\}_{j=1,\ldots,N}$ providing local polynomial reproduction (Theorem 3.2) such that $\widetilde{u}_j$ is supported in $\overline{B(x_j, C_2 h_{X,\Omega})}$ for $j = 1, \ldots, N$. The minimal property stated in equation (7.23) gives

$$\begin{aligned}\sum_{i=1}^{N} \frac{|a_i^*(x)|}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \le \sum_{i=1}^{N} \frac{|\widetilde{u}_i(x)|}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} = \sum_{i\in\widetilde{I}(x)} \frac{|\widetilde{u}_i(x)|}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \le \\ \le \frac{1}{e^{-\nu\left(\frac{C_2 h_{X,\Omega}}{\delta}\right)^2}} \sum_{i\in\widetilde{I}(x)} |\widetilde{u}_i(x)| \le \\ \le \frac{C_1}{e^{-\nu\left(\frac{C_2 h_{X,\Omega}}{\delta}\right)^2}} \le \frac{C_1}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}},\end{aligned}$$

where $\widetilde{I}(x) = \left\{ j \in \{1, \ldots, N\} : x_j \in \overline{B(x, C_2 h_{X,\Omega})} \right\}$. With the last computation we obtained for $i = 1, \ldots, N$

$$\frac{|a_i^*(x)|}{e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}} \le \frac{C_1}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}} \Rightarrow |a_i^*(x)| \le \frac{C_1}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}} e^{-\nu\left(\frac{\|x-x_i\|_2}{\delta}\right)^2}$$

that gives us

$$|a_i^*(x)| \le \frac{C_1}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}} e^{-\nu\left(\frac{\|x-x_i\|_2}{c_\gamma c_{qu} q_X}\right)^2}.$$

Since $e^{-x^2} \le e e^{-x}$ for $x \in \mathbb{R}$ because

$$-x^2 \le -x + 1 \Leftrightarrow x \le x^2 + 1$$

that holds for $x \in [-1, 1]$ and $x \in \mathbb{R} \setminus [-1, 1]$, we obtain

$$|a_i^*(x)| \le \underbrace{e\frac{C_1}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)^2}}}_{C} e^{-\sqrt{\nu}\frac{\|x-x_i\|_2}{c_\gamma c_{qu} q_X}} = C\varphi\left(\frac{\|x - x_i\|_2}{q_X}\right)$$

for $i = 1, \ldots, N$ with

$$\varphi(x) = e^{-\sqrt{\nu}\frac{1}{c_\gamma c_{qu}}x}.$$

□

The same construction that leads to Theorem 7.19 can also be made with

$$w(x, y) = e^{-\nu\frac{\|x-y\|_2}{\delta}} > 0$$

and we obtain

$$C = \frac{C_1}{e^{-\nu\left(\frac{C_2}{\gamma c_\gamma}\right)}} \quad \text{and} \quad \varphi(x) = e^{-\nu\frac{1}{c_\gamma c_{qu}}x}. \tag{7.24}$$

The construction that provides Theorem 7.19 can be made more general if

$$w(x, y) = \varphi\left(\frac{\|x - y\|_2}{q_X}\right) > 0$$

and satisfies the hypothesis in Definition 7.4, obtaining

$$C = \frac{C_1}{\varphi(C_2 c_{qu})} \quad \text{and} \quad \widetilde{\varphi}(x) = \varphi(x).$$

In this context the computational cost depends on the algorithm we use to solve the optimization problem in equation (7.23). We analyse the simplex method [51] so we rewrite equation (7.23) in standard form.

We will use the following notation:

$$w(x) = \left(\frac{1}{e^{-\nu\left(\frac{\|x-x_1\|_2}{\delta}\right)^2}}, \ldots, \frac{1}{e^{-\nu\left(\frac{\|x-x_N\|_2}{\delta}\right)^2}}\right)^\top \in \mathbb{R}^N$$

is the weight vector, $P = (p_j(x_i))_{i=1,\ldots,N, j=1,\ldots,Q} \in M_{N,Q}(\mathbb{R}^d)$ and $S(x) = (p_1(x), \ldots, p_Q(x))^\top \in \mathbb{R}^Q$, where $\{p_1, \ldots, p_Q\}$ is a basis for $\pi_m(\mathbb{R}^d)$.

The optimization problem in equation (7.23) becomes

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} w_i(x)|a_i(x)| \\ \text{s.t.} \quad & P^\top a(x) = S(x), \end{aligned}$$

that in standard form is

$$\begin{aligned} \min \quad & w(x)^\top (a^+(x) + a^-(x)) \\ \text{s.t.} \quad & P^\top (a^+(x) - a^-(x)) = S(x) \\ & a^+(x) \geq 0, a^-(x) \geq 0. \end{aligned} \tag{7.25}$$

If the solution of the linear problem in equation (7.25) is $(\overline{a}^+(x), \overline{a}^-(x)) \in \mathbb{R}^{2N}$ then the solution of (7.23) will be $a^*(x) = \overline{a}^+(x) - \overline{a}^-(x) \in \mathbb{R}^N$.

If we use the simplex method to solve (7.25) then a solution can be a vertex of the polyhedron $\{a \in \mathbb{R}^N : P^\top a = S(x)\}$, so the number of non-zero components of a vertex solution is at most $\text{rank}(P^\top) = Q$. To compute the value of approximant in a point $x \in \Omega$ we need to solve (7.25) and then we perform at most Q multiplications and additions.

We state that if the weight function is continuous then under some conditions we can implement warm-start techniques.

It is useful to rewrite the linear program in (7.25) as

$$\begin{aligned} \max \quad & c(x)^\top \alpha(x) \\ \text{s.t.} \quad & A\alpha(x) = S(x) \\ & \alpha(x) \geq 0, \end{aligned} \tag{7.26}$$

where $c(x) = (-w(x), -w(x)) \in \mathbb{R}^{2N}, \alpha(x) = (a^+(x), a^-(x))$ and $A = (P^\top, -P^\top) \in M_{Q,2N}(\mathbb{R})$.

The simplex method return us not only a solution of the problem but also a base $B \subseteq \{1, \ldots, 2N\}$ such that $|B| = Q$, which is admissible for the primal and for the dual, i.e.

$$\begin{aligned} \overline{S}(x) &= A_B^{-1} S(x) \geq 0 \\ \overline{c}_N(x)^\top &= c_N(x)^\top - c_B(x)^\top A_B^{-1} A_N \leq 0, \end{aligned} \tag{7.27}$$

where $N = \{1, \ldots, 2N\} \setminus B$. If at least one of the inequalities in (7.27) does not reach zero in every component then in a neighborhood of $U \subseteq \Omega$ of x the base B is admissible respectively in the primal or in the dual, so we can restart the algorithm for $y \in U$ without computing a new admissible base.

In general we want that if $\|x - x_i\|_2 \gg 1$ then the weight $w_i(x) \gg 1$ so reasonably we expect $|a_i(x)| \ll 1$ or $|a_i(x)| = 0$. This consideration pushes us to analyze a reduction of the dimensionality of the problem in equation (7.25). We study the column generation approach.

Consider an equivalent formulation of (7.25)

$$\begin{aligned} \min \quad & c(x)^\top \alpha(x) \\ \text{s.t.} \quad & A\alpha(x) = S(x) \\ & \alpha(x) \geq 0, \end{aligned} \tag{7.28}$$

where $c(x) = (w(x), w(x)) \in \mathbb{R}^{2N}$, $\alpha(x) = (a^+(x), a^-(x))$ and $A = (P^\top, -P^\top) \in M_{Q,2N}(\mathbb{R})$.

The dual problem of (7.28) is

$$\begin{aligned} \min \quad & S(x)^\top u(x) \\ \text{s.t.} \quad & A^\top u(x) \leq c(x). \end{aligned} \tag{7.29}$$

To understand the dimensionality reduction is better to write (7.28) and (7.29) more explicitly.

$$\begin{aligned} \min \quad & \sum_{i=1}^{2N} c_i(x)\alpha_i(x) \\ \text{s.t.} \quad & \sum_{i=1}^{2N} A_{ij}\alpha_i(x) = S_j(x) \qquad && j = 1, \dots, Q \\ & \alpha_i(x) \geq 0 && i = 1, \dots, 2N. \end{aligned} \tag{7.30}$$

$$\begin{aligned} \max \quad & \sum_{i=1}^{Q} S_i(x)u_i(x) \\ \text{s.t.} \quad & \sum_{i=1}^{Q} A_{ji}u_i(x) \leq c_j(x) \qquad && j = 1, \dots, 2N. \end{aligned} \tag{7.31}$$

We fix $\widehat{\mathcal{S}} \subseteq \{1, \dots, 2N\}$ such that the reduced problem of (7.30) is admissible.

$$\begin{aligned} \min \quad & \sum_{i \in \widehat{\mathcal{S}}} c_i(x)\alpha_i(x) \\ \text{s.t.} \quad & \sum_{i \in \widehat{\mathcal{S}}} A_{ij}\alpha_i(x) = S_j(x) \qquad && j = 1, \dots, Q \\ & \alpha_i(x) \geq 0 && i \in \widehat{\mathcal{S}}. \end{aligned} \tag{7.32}$$

The dual of (7.32) (that is the reduced problem of (7.31)) becomes

$$\begin{aligned} \max \quad & \sum_{i=1}^{Q} S_i(x) u_i(x) \\ \text{s.t.} \quad & \sum_{i=1}^{Q} A_{ji} u_i(x) \leq c_j(x) \qquad j \in \widehat{\mathcal{S}}. \end{aligned}$$

If $(\overline{\alpha}_i(x))_{i\in\widehat{\mathcal{S}}}$ is a solution of (7.32) and $\overline{u}(x) \in \mathbb{R}^Q$ is its dual solution we can extend $(\overline{\alpha}_i(x))_{i\in\widehat{\mathcal{S}}}$ to an admissible solution of (7.30) imposing $\overline{\alpha}_i(x) = 0$ if $i \notin \widehat{\mathcal{S}}$. We remark that by duality theory

$$\sum_{i=1}^{2N} c_i(x)\overline{\alpha}_i(x) = \sum_{i\in\widehat{\mathcal{S}}} c_i(x)\overline{\alpha}_i(x) = \sum_{i=1}^{Q} S_i(x) u_i(x),$$

so $\overline{\alpha}(x)$ is a solution of (7.30) if $\overline{u}(x)$ is admissible in (7.31).

If $\overline{u}(x)$ is not admissible then there exits $j \in \{1, \ldots, 2N\}\backslash\widehat{\mathcal{S}}$ such that $\sum_{i=1}^{Q} A_{ji}u_i(x) > c_j(x)$, so we iterate the procedure adding to $\widehat{\mathcal{S}}$ the element j.

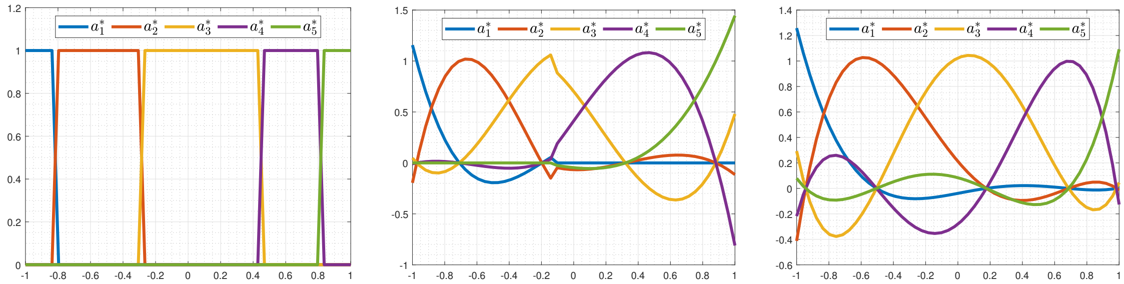
At each step we solve a problem that has less variable than (7.30) and then we perform an admissibility check that cost $\mathcal{O}(N)$. This procedure can be useful because we can guess where the non-zero components of the solution of (7.30) are and find the solution with less computational effort. We remark that a vertex solution has at most Q non-zero components.

We conclude remembering that we can improve the performance of the numerical scheme described in equation (7.23) using warm-start algorithms and the column generation approach.

7.3.1 Numerical test

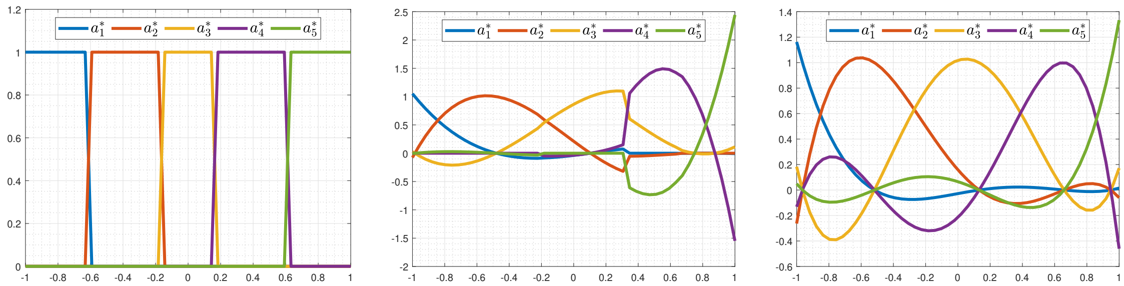
The goal of this section is to show some numerical experiments that confirm the theoretical results regarding approximation methods with $\|\cdot\|_1$ (equation (7.23)). We will study also compactly supported weight functions.

We start by analyzing the basis functions of equation (7.23).



(a) m=0 (b) m=3 (c) m=4

Figure 7.15: Basis functions of equation (7.23). The weight function coincides with $e^{-x^2} \in \mathcal{C}^\infty(\mathbb{R})$ and the approxiomation nodes are 5 uniformly perturbed equispaced nodes in $[-1, 1]$. From left to right the basis functions reproduce the polynomials of degree $0, 3$ and 4 respectively. In this numerical test $\delta = 5h_{X,\Omega}$.



(a) m=0 (b) m=3 (c) m=4

Figure 7.16: Basis functions of equation (7.23). The weight function coincides with $e^{-x} \in \mathcal{C}(\mathbb{R})$ and the approxiomation nodes are 5 uniformly perturbed equispaced nodes in $[-1, 1]$. From left to right the basis functions reproduce the polynomials of degree $0, 3$ and 4 respectively. In this numerical test $\delta = 5h_{X,\Omega}$.

As fast decaying polynomial reproduction methods Figure 7.15 and Figure 7.16 show an interesting result. We do not know the regularity of the basis functions $\{a_1^*, \dots, a_5^*\}$ but when the degree of the polynomials to be reproduced increases then also the smoothness of the basis functions increases from a practical point of view. With our numerical experiments this consideration does not depend on the weight functions used. We have seen in Theorem 7.19 that the smoothness of the weight functions does not affect the convergence rate but a smooth approximant can be useful for applications. From Figure 7.15 and Figure 7.16 we can underline a characteristic that derives from the method used to solve the linear optimization problem in equation (7.23). Since we use the simplex method then in each point of $[-1, 1]$ only $m + 1$ basis functions are different from zero.

To produce Figure 7.15 and Figure 7.16 we used as linear optimization solver Gurobi 10 with a tolerance on optimality conditions and constraints of 10^{-10}. As polynomial basis in equation (7.25) we choose Chebyshev polynomials of the first kind.

Now we discuss some experiments to confirm Theorem 7.15 numerically. We approximate different functions on equispaced nodes (quasi-uniform data set) in $[-1,1]$. We use Gurobi 10 with a tollerance of 10^{-10} to get the approximant and the polynomial basis is the Chebyshev polynomial basis of the first kind. To interface with the linear solver we use AMPL as modeling language. The scaling of δ is constant because it does not influence the convergence rate even if appropriate choices of the parameter can improve the stability of the method (Theorem 7.14, equation (7.24)). We fix $\delta = 5h_{X,\Omega}$.

As weight functions in equation (7.23) we used

$$w_1(x,y) = e^{-\frac{|x-y|}{\delta}} \quad \text{and} \quad w_2(x,y) = \phi_{1,1}\left(\frac{|x-y|}{\delta}\right),$$

where $\phi_{1,1}(r) = (1-r)_+^3(3r+1)$ is a $\mathcal{C}^2(\mathbb{R})$-Wendland's function [52].

The following numerical experiments approximate the functions

$$f_1(x) = \sin(\pi x),$$
$$f_2(x) = 6x^6 + 5x^3 + x^2,$$
$$f_3(x) = \frac{1}{1+30x^2}.$$

In the following graphs the blue line allows us to check the correct slope of the approximation error (Theorem 7.15).

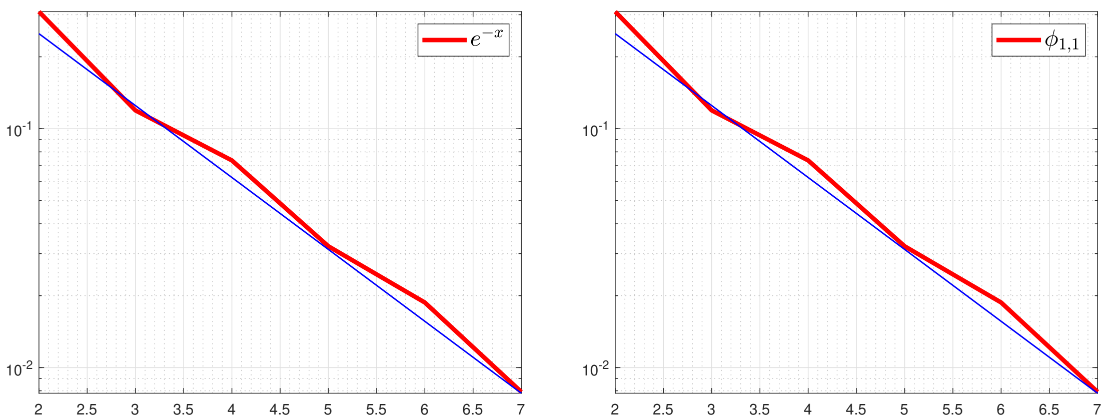


Figure 7.17: Convergence rate of the approximation error $\|f_1 - z_{f_1,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^8+1]$. The approximation method reproduces exactly polynomials of degree 0.

Nodes	9	17	33	65	129	257	Degree
e^{-x}	3.09e-01	1.19e-01	7.37e-02	3.22e-02	1.87e-02	7.91e-03	0
$\phi_{1,1}$	3.09e-01	1.19e-01	7.37e-02	3.22e-02	1.87e-02	7.91e-03	0

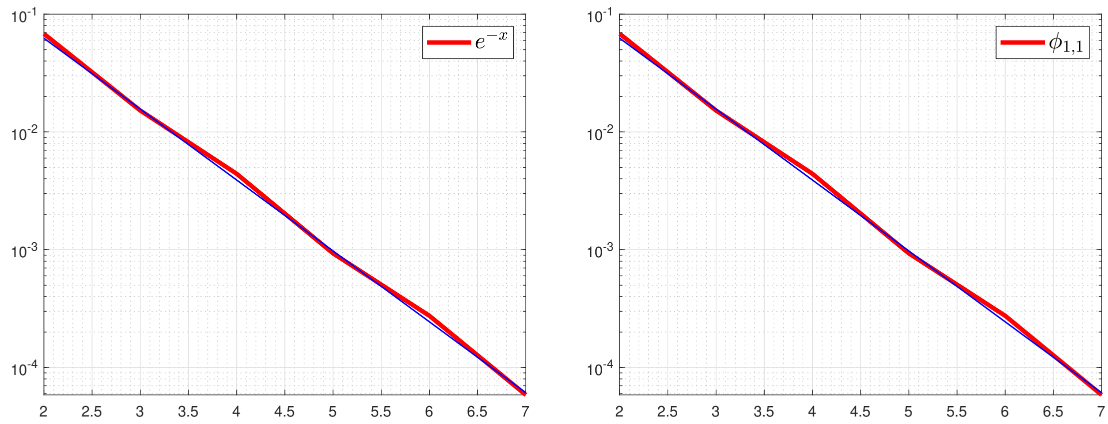


Figure 7.18: Convergence rate of the approximation error $\|f_1 - z_{f_1,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^8+1]$. The approximation method reproduces exactly polynomials of degree 1.

Nodes	9	17	33	65	129	257	Degree
e^{-x}	6.82e-02	1.52e-02	4.41e-03	9.31e-04	2.75e-04	5.86e-05	1
$\phi_{1,1}$	6.82e-02	1.52e-02	4.41e-03	9.31e-04	2.75e-04	5.86e-05	1

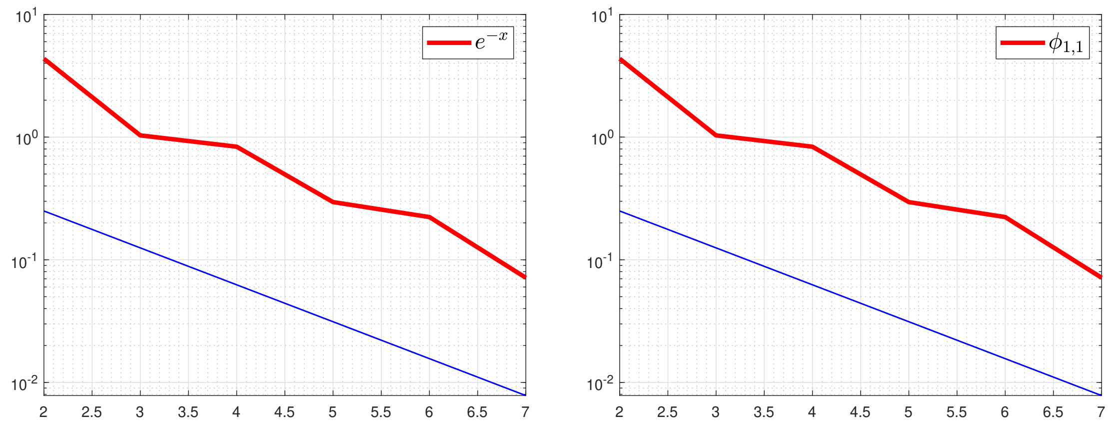


Figure 7.19: Convergence rate of the approximation error $\|f_2 - z_{f_2,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^8+1]$. The approximation method reproduces exactly polynomials of degree 0.

Nodes	9	17	33	65	129	257	Degree
e^{-x}	4.36e+00	1.03e+00	8.36e-01	2.95e-01	2.23e-01	7.14e-02	0
$\phi_{1,1}$	4.36e+00	1.03e+00	8.36e-01	2.95e-01	2.23e-01	7.14e-02	0

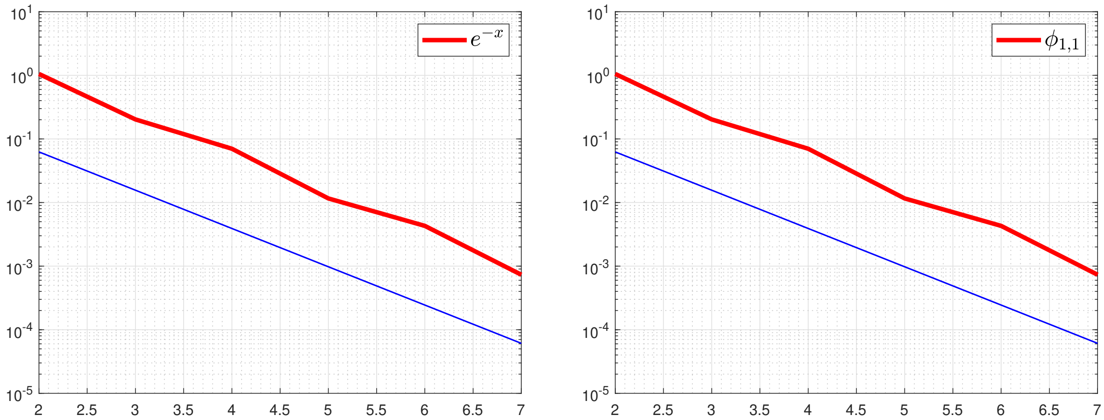


Figure 7.20: Convergence rate of the approximation error $\|f_2 - z_{f_2,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^8+1]$. The approximation method reproduces exactly polynomials of degree 1.

Nodes	9	17	33	65	129	257	Degree
e^{-x}	1.05e+00	2.03e-01	7.01e-02	1.16e-02	4.29e-03	7.32e-04	1
$\phi_{1,1}$	1.05e+00	2.03e-01	7.01e-02	1.16e-02	4.29e-03	7.32e-04	1

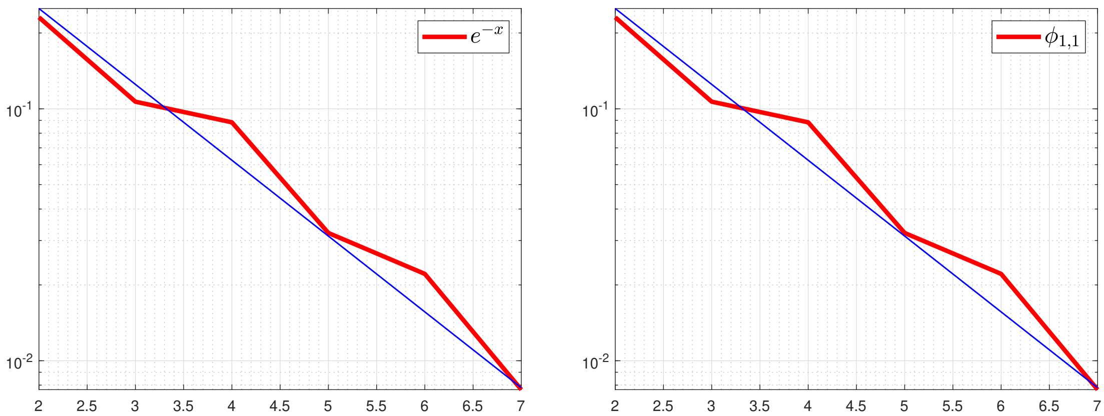


Figure 7.21: Convergence rate of the approximation error $\|f_3 - z_{f_3,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^8+1]$. The approximation method reproduces exactly polynomials of degree 0.

Nodes	9	17	33	65	129	257	Degree
e^{-x}	2.31e-01	1.07e-01	8.84e-02	3.21e-02	2.21e-02	7.68e-03	0
$\phi_{1,1}$	2.31e-01	1.07e-01	8.84e-02	3.21e-02	2.21e-02	7.68e-03	0

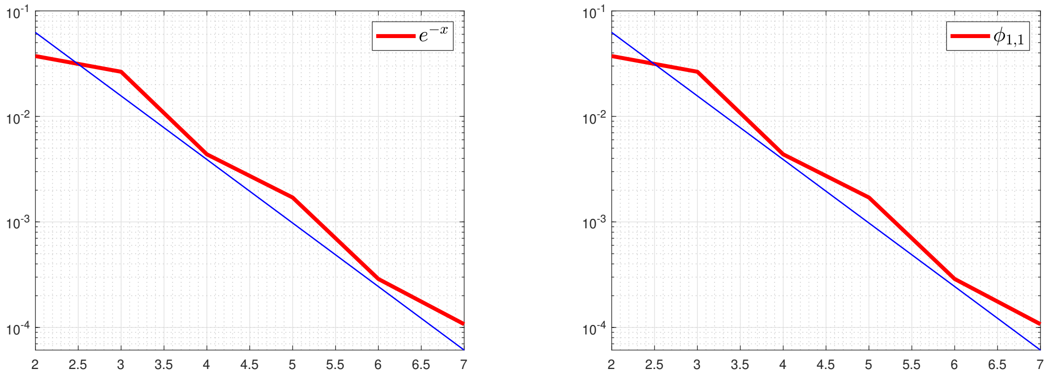
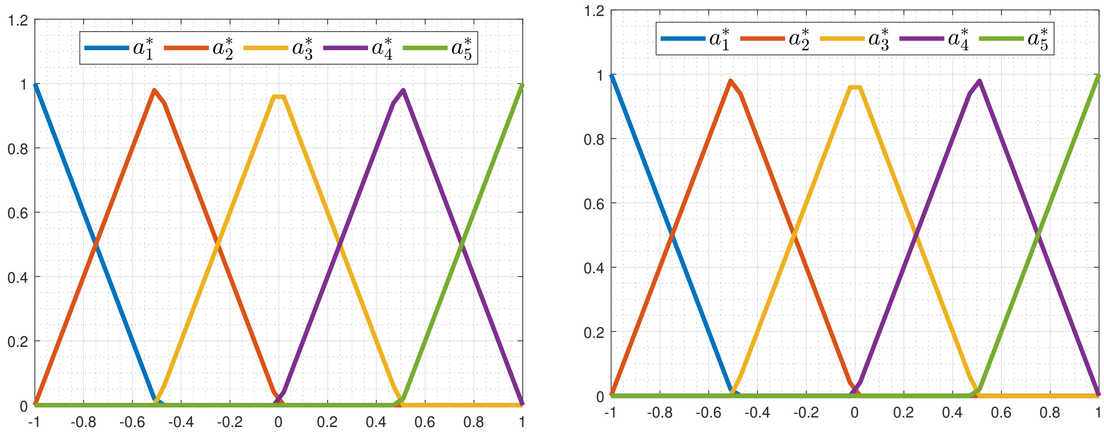


Figure 7.22: Convergence rate of the approximation error $\|f_3 - z_{f_3,X}\|_{L^\infty([-1,1])}$. The x-axis describes the number of equispaced nodes used to produce the approximant. $|X| \in [2^3+1, 2^8+1]$. The approximation method reproduces exactly polynomials of degree 1.

Nodes	9	17	33	65	129	257	Degree
e^{-x}	3.73e-02	2.65e-02	4.37e-03	1.70e-03	2.88e-04	1.07e-04	1
$\phi_{1,1}$	3.73e-02	2.65e-02	4.37e-03	1.70e-03	2.88e-04	1.07e-04	1

Numerical experiments confirm the statement of Theorem 7.15 because if the method reproduces exactly $\pi_m(\mathbb{R}^d)$ then the convergence rate is $\mathcal{O}(h_{X,\Omega}^{m+1})$.

We can note that the approximation errors for the different weight functions w_1 and w_2 coincide. This can be explained by showing that the basis functions $\{a_i^*\}_{i=1,\dots,N}$ (equation (7.23)) coincide when reproducing polynomials of degree 0 and 1.



(a) $w_1(x,y) = e^{-\frac{|x-y|}{\delta}}$

(b) $w_2(x,y) = \phi_{1,1}\left(\frac{|x-y|}{\delta}\right)$

Figure 7.23: Basis functions of equation (7.23). The approxiomation nodes are 5 equispaced nodes in $[-1,1]$. The basis functions reproduce the polynomials of degree 1. In this numerical test $\delta = 3h_{X,\Omega}$.

7.4 Conclusion

The rescaled localized radial basis function method (RL-RBF) proposed in [1] represents an efficient numerical scheme to interpolate functions on a scattered set of nodes. If radial basis functions with compact support are used the main idea is to maintain a limit on the number of nodes in the support of each function even if the number of nodes increases (we want to preserve the locality of the problem). We can achieve this by scaling the basis functions with a parameter δ proportional to the fill distance $h_{X,\Omega}$ to control the size of the support. The algorithm, that is an instance of the Shepard's method, reproduces exactly constants functions and for this reason linear convergence with respect to the fill distance is expected. Numerical evidences show this feature also on large-scale interpolation problems (the method is suitable for parallel computation). We have to wait for the results of [2] to read a proof of the convergence for RL-RBF. This proof works up to a conjecture in the quasi-uniform setting (this assumption is useful to scale all the basis functions with the same parameter, in a more general context the parameter δ also depends on the density of the nodes in different areas of the interpolation domain). If we restrict ourselves to a quasi-uniform framework and the native space of the radial basis function we are studying coincides with a Sobolev space then the conjecture can be reported as follows: we can determine a lower bound for the interpolant of the constant function 1 uniformly with respect to the choice of the parameter δ. The solution of this conjecture is important because we can obtain a sufficient condition on the number of nodes in the support of each basis function to guarantee the linear convergence of the numerical scheme. Other relevant properties concern the stability of the method (the condition number of the interpolation matrix, the maximum eigenvalue and the inverse of the minimum eigenvalue are bounded by a constant in a uniform way with respect to the choice of the scaling δ) and the effectiveness of the implementation (the interpolation matrix is sparse and therefore treatable also for large dimensions).

In this work we analyzed some practical and theoretical results to support the conjecture. Thanks to the estimates on the uniform norm of the inverse of the interpolation matrix we can state that the interpolant does not suffer from Runge's phenomenon. Moreover, deepening the relationship between the cardinal functions of a positive definite kernel with the cardinal functions of the same kernel thought as conditionally positive definite with respect to the constants we reached inequalities for the uniform norm of the difference of cardinal functions (we travel two different approaches that led to the same result, the first one uses the power function and the second exploits matrices properties). The approach with linear algebra gives us the possibility to estimate the polynomials of degree 0 which allow us to reproduce exactly the constant functions (the terms just mentioned tend to 0 when the number of nodes increases).

Our work continues generalizing the RL-RBF method by increasing the dimension of the polynomial space to be reproduced exactly. The goal is to determine a convergent method whose convergence rate is $\mathcal{O}(h_{X,\Omega}^{m+1})$ if all polynomials up to degree m can be approximated correctly. We modified the definition of local polynomial reproduction by replacing the compactness of the support with a fast decay of the basis functions. The RL-RBF method adapts to this new definition, while it does not reproduce polynomials locally because the cardinal functions have not compact support. In the quasi-uniform setting this new approximation scheme is convergent and stable (the Lebesgue constant depends on the dimension of the space, the dimension of the polynomial space that is reproduced and on the decaying of the basis functions). At this point in the analysis, smoothness plays no fundamental role. In addition to RL-RBF method we proposed a further approximation scheme which provides smooth quasi-interpolants ($\mathcal{C}^\infty(\mathbb{R}^d)$). The method approximates the value of an unknown function using the moving least squares technique with a Gaussian as weight function. The smoothness of the interpolant is inherited from the smoothness of the weight functions. In this case the decrease of the basis functions is controlled by the inverse of an exponential function. Since the solution of a moving least squares problem is the solution of a quadratic optimization problem we are able to analyze the computational cost of the method, which turns out to be linear in the number of nodes. Numerical tests confirm the theoretical results of convergence and also the smoothness of the basis functions. With the numerical tests we obtain an unexpected result: even if the weight function is only continuous, if we increase the dimension of the polynomial space to be reproduced then the basis functions turn out to be numerically smooth.

In these analysis and numerical tests we considered functions with global support and the matrices involved can be dense although small in size when the space to be reproduced is not too large. To address this difficulty we replaced the quadratic optimization problem deriving from moving least squares with a linear program on a polyhedron. Also in this area, with techniques similar to the previous ones, a stable and convergent method can be achieved with the same convergence rate. A vertex solution of the linear optimization problem allows us to control the number of non-zero basis functions at each point in the domain (the basis functions that are different from 0 are at most $m + 1$ if we reproduce all the polynomials up to degree m in a space of dimension 1). Since the weight functions try to locate the optimization problem, we expect that the value of the weight function corresponding to a node in the domain is large when the considered point is far from the node. This type of experience leads us to use column generation techniques to reduce the dimensionality of the problem (we can try to predict the non-zero basis functions because the number of them is bounded uniformly with respect to the fill distance). If the weight functions are continuous then warm start techniques in the primal and in the dual problem are easily applied. The numerical results confirm the theoretical evidences on convergence and even if we do not have any results on the smoothness of the approximant we get similar outcomes to the moving least squares method (numerically we can observe that the basis functions become smooth when the polynomial space to be reproduced gets bigger). The method has been tested with weight functions with global (exponential function) and compact (Wendland's functions) support. The convergence results obtained are equivalent.

This work could also be continued in the future by looking for the solution to the conjecture and analyzing the optimization problem of the least squares method with different norms and other weight functions, which guarantee a fast decay of the basis functions.

Bibliography

[1] S. Deparis, D. Forti, A. Quarteroni, *A rescaled localized radial Basis Function interpolation on non-Cartesian and nonconforming grids*, SIAM Journal on Scientific Computing, 36, 2014.

[2] S. De Marchi, H. Wendland, *On the convergence of the rescaled localized radial basis function method*, Applied Mathematics Letters, 99, 2020.

[3] S. Axler, *Measure, Integration and Real Analysis*, Graduate Texts in Mathematics, Springer, 282, 2020.

[4] W. Rudin, *Functional Analysis*, International series in pure and applied mathematics, McGraw-Hill, 1991.

[5] E.M. Stein, R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.

[6] N. Aronszajn, *Theory of Reproducing Kernels*, Transactions of the American Mathematical Society 68, 1950.

[7] H. Meschkowski, *Hilbertsche Räume mit Kernfunktionen*, Berlin-Göttingen-Heidelberg, Springer, 1962.

[8] R. Schaback, *Native Hilbert Spaces for Radial Basis Functions I*, New Developments in Approximation Theory, ISNM International Series of Numerical Mathematics, 132, 1999.

[9] R. Schaback, *A unified theory of radial basis functions: Native Hilbert spaces for radial basis functions II*, Journal of Computational and Applied Mathematics, 121, 2000.

[10] W.R. Madych, S.A. Nelson, *Multivariate Interpolation and Conditionally Positive Definite Functions*, Approx. Theory and its Applications 4, 1988.

[11] W.R. Madych, S.A. Nelson, *Multivariate Interpolation and Conditionally Positive Definite Functions II*, Math. Comp. 54, 1990.

[12] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2004.

[13] R. Schaback, H. Wendland, *Kernel Techniques: From Machine Learning to Meshless Methods*, Cambridge University Press, Acta Numerica, 2006.

[14] M. Reed, B. Simon, *Functional Analysis*, Methods of Modern Mathematical Physics, Elsevier Science, 1981.

[15] G.B. Folland, *Fourier analysis and its applications*, American Mathematical Soc., 2009.

[16] H. Wendland, *Multiscale analysis in Sobolev spaces on bounded domains*, Numer. Math., 116, 2010.

[17] A. Townsend, H. Wendland, *Multiscale analysis in Sobolev spaces on bounded domains with zero boundary values*, Journal of Numerical Analysis, 33, 2013.

[18] H. Mhaskar, F. Narcowich, J.D. Ward, *Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature* Mathematics of Computation, 70, 2001.

[19] P. Lancaster, K. Salkauskas, *Surfaces Generated by Moving Least Squares Methods*, Mathematics of Computation 37, 155, 1981.

[20] D.H. McLain, *Drawing Contours from Arbitrary Data Points*, The Computer Journal, 17, 4, 1974.

[21] D.H. McLain, *Two Dimensional Interpolation from Random Data*, Comput. J., 19, 1976.

[22] D. Shepard, *A Two-Dimensional Interpolation Function for Irregularly-Spaced Data*, Proceedings of the 1968 ACM National Conference, 1968.

[23] R.L. Hardy, *Multiquadric equations of topography and other irregular surfaces*, J. Geophys. Res., 76, 1971.

[24] R.L. Hardy, *Theory and applications of the multiquadric–biharmonic method*,Comput. Math. Appl., 1990.

[25] J. Duchon, *Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces*, Rev. Franc¸aise Automat. Informat. Rech. Oper. Anal. ´ Numer., 10, 1976.

[26] J. Duchon *Splines minimizing rotation-invariant semi-norms in Sobolev spaces*, Constructive Theory of Functions of Several Variables, Springer, Berlin, 1977.

[27] J. Duchon *Sur l'erreur d'interpolation des fonctions de plusieurs variables par les D^m-splines*, Rev. Franc¸aise Automat. Informat. Rech. Oper. Anal. Numer., 12, 1978.

[28] W. Pogorzelski, *Integral Equations and their Applications*, Volume 1, Oxford, Pergamon Press, 1966.

[29] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems. Kluwer Academic Publishers, 2000.

[30] R.A. DeVore, R.C. Sharpley, *Besov spaces on domains in* $\mathbb{R}^d$, Transactions of the American Mathematical Society, 335(2), 1993.

[31] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1971.

[32] W.R. Madych, S.A. Nelson, *Multivariate interpolation: a variational theory* Unpublished manuscript, 1983.

[33] Z. Wu, R. Schaback, *Local error estimates for radial basis function interpolation of scattered data*, IMA J. Numer. Anal., 13, 1993.

[34] K. Ball, *Eigenvalues of Euclidean distance matrices*, J. Approx. Theory, 8, 1992.

[35] F.J. Narcowich, J.D. Ward, *Norms of inverses for matrices associated with scattered data*, Curves and Surfaces, Boston, Academic Press, 1991.

[36] F.J. Narcowich, J.D. Ward, *Scattered-data interpolation on* $\mathbb{R}^d$*: error estimates for radial basis and band-limited functions*, SIAM J. Math. Anal., 36, 2004.

[37] F.J. Narcowich, J.D. Ward, *Norm estimates for the inverse of a general class of scattered-data radial-function interpolation matrices*, J. Approx. Theory, 1992.

[38] F.J. Narcowich, J.D. Ward, *On condition numbers associated with radial-function interpolation*, J. Math. Anal. Appl., 186, 1994.

[39] R. Schaback, *Error estimates and condition number for radial basis function interpolation*, Adv. Comput. Math., 3, 1995.

[40] E.M. Stein, R. Shakarchi, *Fourier Analysis: An Introduction (Princeton Lectures in Analysis I)*, Princeton University Press, 2003.

[41] G.N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge, Cambridge University Press, 1966.

[42] N.N. Lebedev, *Special Functions and their Applications* Englewood Cliffs, Prentice-Hall, 1965.

[43] R.A. Horn, C.R. Johnson, *Matrix Analysis*, New York, Cambridge University Press, 2013.

[44] J.C. Mairhuber, *On Haar's theorem concerning Chebychev approximation problems having unique solutions*, Proc. Amer. Math. Soc., 7, 1956.

[45] P.C. Curtis, *N-parameter families and best approximation*, Pacific J. Math., 9, 1959.

[46] M. Golomb, H.F. Weinberger, *Optimal approximation and error bounds*, On Numerical Approximation, Madison, University of Wisconsin Press, 1959.

[47] C.A. Micchelli, T.J. Rivlin, *A survey of optimal recovery*, Optimal Estimation in Approximation Theory,New York, Plenum Press, 1977.

[48] R. Farwig, *Rate of convergence of Shepard's global interpolation formula*, Mathematics of Computation, 46, 1986.

[49] B.J.C. Baxter, N. Sivakumar, and J.D. Ward, *Regarding the p-norms of radial basis interpolation matrices*, Constr. Approx., 10, 1994.

[50] C. Rieger, H. Wendland, *Sampling inequalities for sparse grids*, Numer. Math., 136, 2017.

[51] H. Karloff, *The Simplex Algorithm*, Linear Programming, Modern Birkhäuser Classics, Birkhäuser Boston, 2009.

[52] H. Wendland, *Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree*, Adv. Comput. Math., 1995.

[53] H. Wendland, *Error estimates for interpolation by compactly supported radial basis functions of minimal degree*, J. Approx. Theory, 93, 1998.

[54] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.

[55] G.E. Fasshauer, *Meshfree Approximation Methods with Matlab*, Interdisciplinary Mathematical Sciences, 2007.

[56] T.J. Rivlin, *Chebyshev Polynomials*, 2nd Edition, John Wiley and Sons, New York, 1990.