Better bases for kernel spaces

E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, G. B Wright

November 7, 2011

Abstract

In this article we investigate the feasibility of constructing stable, local bases for computing with kernels. In particular, we are interested in constructing families \( b = (b_\xi)_{\xi \in \Xi} \) that function as bases for kernel spaces \( S(k, \Xi) = \left\{ \sum_{\xi \in \Xi} a_\xi k(\cdot, \xi) \mid (a_\xi)_{\xi \in \Xi} \in \mathbb{R}^\Xi \right\} \) so that each basis function can be obtained by very few kernels

\[
 b_\xi = \sum_{\xi \in \Xi} A_{\zeta, \xi} k(\cdot, \xi) \quad A_{\zeta, \xi} = 0 \text{ for all but a few } \xi.
\]

This is reminiscent of the construction of the B-spline basis from the family of truncated power functions.

We demonstrate that for a large class of kernels (the Sobolev kernels as well as many kernels of polyharmonic and related type) such bases exist. In fact, the basis elements can be constructed using a combination of roughly \( O(\log N)^d \) kernels, where \( d \) is the local dimension of the manifold and \( N \) is the dimension of the kernel space (i.e. \( N = \#\Xi \)). Viewing this as a preprocessing step – the construction of the basis has computational cost \( O(N(\log N)^d) \). Furthermore, we prove that the new basis is \( L_p \) stable and satisfies polynomial decay estimates that are stationary with respect to the density of \( \Xi \).

1 Introduction

The purpose of this article is to investigate robust bases for spaces associated with a positive definite or conditionally positive definite kernel \( k : \mathbb{M} \times \mathbb{M} \to \mathbb{R} \), where \( \mathbb{M} \) is a \( C^\infty \) (closed) compact Riemannian manifold. The dimension of \( \mathbb{M} \) is \( d \). The kernels that we discuss below belong to a wide class that includes the thin-plate splines and similar kernels when \( \mathbb{M} = \mathbb{S}^d \) or \( SO(3) \).

\*Department of Mathematics, High Point University, High Point, NC 27262, USA.
\†Department of Mathematics, University of Hawaii, Honolulu, HI 96822, USA. Research supported by grant DMS-104794 from the National Science Foundation.
\‡Department of Mathematics, Texas A&M University, College Station, TX 77843, USA. Research supported by grant DMS-0807033 from the National Science Foundation.
\§Department of Mathematics, Texas A&M University, College Station, TX 77843, USA. Research supported by grant DMS-0807033 from the National Science Foundation.
\¶Department of Mathematics, Boise State University, Boise, ID 83725, USA. Research supported by grants DMS-0934581 and DMS-0540779 from the National Science Foundation.
The spaces associated with a kernel from this class are defined as follows. Let \( \Xi \subset M \) be a finite set of points, called centers, having cardinality \( N = \# \Xi \). The centers are scattered in the sense that they do not need to belong to a regular grid. In the positive definite case, the space \( S(k, \Xi) \) associated with the kernel \( k \) and the set \( \Xi \) is just \( S(k, \Xi) = \text{span}\{k(\cdot, \xi), \xi \in \Xi\} \); that is,

\[
S(k, \Xi) := \left\{ \sum_{\xi \in \Xi} a_\xi k(\cdot, \xi), \ a_\xi \in \mathbb{R} \right\}.
\]

The conditionally positive definite case is similar; we will discuss it in Section 4.1 – specifically in (4.3). For these spaces, if \( h := \max_{\xi \in \Xi} \text{dist}(x, \xi) \) is the mesh norm (fill distance), \( q := \frac{1}{2} \min_{\xi \neq \eta} \text{dist}(\eta, \xi) \) is the separation radius and \( \rho_\Xi := h/q \) is the mesh ratio, then as long as \( \rho_\Xi \leq \rho_0 \), where \( \rho_0 \) is fixed, Lebesgue constants are uniformly bounded and approximation rates for functions in Sobolev spaces \( W^m_p(M) \) are \( O(h^m) \), with the constants independent of other properties of \( \Xi \) [20, 19].

Two other remarkable properties of \( S(k, \Xi) \) concern its Lagrange basis, \( \{\chi_\xi(\cdot)\}_{\xi \in \Xi} \). Recall that in a Lagrange basis each basis function satisfies \( \chi_\xi(\eta) = \delta_{\xi, \eta} \) when \( \eta \in \Xi \). What was shown in [20, 19] is that \( \chi_\xi \) decays exponentially fast away from \( \xi \) for special kernels, and algebraically fast for many others.

Equally as important, as we shall prove below in Theorem 4.3 if we express the \( \chi_\xi \)'s in the standard basis,

\[
\chi_\xi = \sum_{\eta \in \Xi} A_{\xi, \eta} k(\cdot, \eta),
\]

where the coefficients \( A_{\xi, \eta} \) are well-known to be the entries of the inverse of the interpolation matrix, then \( |A_{\xi, \eta}| \) decays as a function of \( \text{dist}(\eta, \xi) \) at the same rate as \( |\chi_\xi(x)| \) decays in \( \text{dist}(x, \xi) \) – i.e., exponentially or algebraically, as the case may be. Prior to our work, the only provable results concerning decay of these coefficients were done by Fornberg [14] in the case of \( \mathbb{R} \) and \( \mathbb{R}^2 \) for gridded data, using Fourier techniques that do not carry over to the scattered case.

The difficulty with the Lagrange basis is that each \( \chi_\xi \) is computationally costly both to construct (as a linear combination of \( k(\cdot, \xi), \xi \in \Xi \)) and to compute with. Are there better bases? Here is what we would desire in a basis \( \{b_\xi\}_{\xi \in \Xi} \) for \( S(k, \Xi) \).

Each basis function should be highly localized and nearly scalable with respect to the mesh norm \( h \) of \( \Xi \). By this we mean that each basis element is of the form

\[
b_\xi = \sum_{\eta \in \Xi} A_{\xi, \eta} k(\cdot, \eta)
\]

where the \( \eta \)'s come from small subset of the centers \( \Xi \) and satisfy the following requirements:

\begin{itemize}
  \item[i)] \( \#\{A_{\rho, \xi} \neq 0\} = c(\# \Xi) \)
  \item[ii)] \( |b_\rho(x)| \leq \sigma \left( \frac{\text{dist}(x, \xi)}{h} \right) \)
\end{itemize}

where the cost \( c(N) \) is constant or slowly growing with \( N = \# \Xi \) and the function \( \sigma(x) \) decays rapidly: at an exponential rate \( \sigma(x) \leq Ce^{-\nu|x|} \) or at least at a fast polynomial rate.
\[ \sigma(x) \leq C(1 + |x|)^{-J}. \] The B-spline basis, constructed from the family of truncated power functions, is a model kernel method and in the \( d = 1 \) case provides an ideal solution to the problem we consider.

The main results of this paper demonstrate that such kernel bases exist and that each basis function can be computed in nearly fixed time. Viz., an individual basis function can be computed in \( O((\log N)^d) \) time while the full basis can be computed in \( O(N(\log N)^d) \) time with \( N = \#\Xi = Ch^{-d} \). Moreover the basis is \( L^p \) stable.

The main tool employed is Theorem 4.3 which allows one to bound the rate of decay of Lagrange coefficients in terms of the corresponding decay rate of the Lagrange functions. In particular if the Lagrange functions have exponential decay so too do their corresponding coefficients. While, as we mentioned earlier, this fact was previously known in the scaled lattice case, but no such estimates have been available in the scattered case.

The sphere \( S^2 \) As an example of our main results for the sphere \( S^2 \) and the restricted surface splines of order \( s + 1 \), given by

\[ k_{s+1}(x, \alpha) := (1 - x \cdot \alpha)^s \log(1 - x \cdot \alpha), \]

for \( s = 1, 2, 3, \ldots \), we have the following theorem, which is a corollary of Theorem 5.1 in Section 5.

**Theorem.** For a sufficiently dense set of centers \( \Xi \), and for a sufficiently large constant \( \tau \) there is basis \( (b_\xi)_{\xi \in \Xi} \) for the space \( S(k, \Xi) \) satisfying the following.

- Each basis element \( b_\xi = \sum_{\zeta \in \Xi} A_{\xi, \zeta} k(\cdot, \zeta) \) is composed of at most \( M := \tau(\log N)^2 \) kernels. I.e.
  \[ c(\#\Xi) = M = \tau(\log N)^2. \]
- Each basis element exhibits polynomial decay: there exist constants \( C \) and \( J \) for which
  \[ |b_\xi(x)| \leq C \left( 1 + \frac{\text{dist}(x, \xi)}{h} \right)^{-J}. \]
- The rate of polynomial decay \( J \) depends linearly on the constant of proportionality \( \tau \) by
  \[ J = O(\sqrt{\tau}). \]
- The basis is \( L^p \) stable: there are constants \( 0 < c_1 \leq c_2 < \infty \) depending on \( \tau \) so that for all sequences \( a = (a_\xi)_{\xi \in \Xi} \in \mathbb{R}^\Xi \) the following holds:
  \[ c_1 q^{2/p} \|a\|_{\ell_p(\Xi)} \leq \left\| \sum_{\xi \in \Xi} a_\xi b_\xi \right\|_{L_p(S^2)} \leq c_2 q^{2/p} \|a\|_{\ell_p(\Xi)}. \]

The restricted surface splines on \( S^2 \) are of special importance, largely because the sphere is the setting of many problems of scientific interest, but also because the kernels themselves have a convenient, closed form representation and their approximation power that is well understood and optimal in the sense that approximation rates are in line with smoothness assumptions for the target functions (i.e., approximands in Sobolev classes \( W^s_p \) or Besov classes \( B^s_{p,\infty} \) are approximated by functions in \( S(k, \Xi) \) with error decaying like \( h^s \)).
We give some numerical examples for $S^2$ with such kernels (and others) in Section 6.

**Preconditioners** Over the years practical implementation of kernel approximation has progressed despite the ill-conditioning of kernel bases. This has happened with the help of clever numerical techniques like multipole methods [3, 8, 6] and often with the help of preconditioners [5, 12, 23, 35] of which [7, 27, 28] are of special interest to us, because these involve attempts to construct local bases. Indeed, another offshoot of our results in that one can now estimate, a priori, the number of coefficients needed to guarantee good preconditioners. Many results already exist in the RBF literature concerning preconditioners and “better” bases. For a good list of references and further discussion, see [11]. Several of these papers use “local Lagrange” functions in their efforts to efficiently construct interpolants. The number of points chosen to localize the Lagrange functions are ad hoc and seem to be based on experimental evidence. For example, Faul and Powell, in [13], devise an algorithm which converges to a given RBF interpolant that is based on local Lagrange interpolants using about thirty nearby centers. Beatson–Cherrie–Mouat, in [4], use fifty local centers (p. 260, Table 1) in their construction along with a few “far away” points to control the growth of the local interpolant at a distance from the center. In other work, Ling and Kansa [23] and co-workers have studied approximate cardinal basis functions based on solving least squares problems. Thus one goal of this paper is to provide some theoretical groundwork that may yield future improvements in preconditioner algorithms and better bases for kernel spaces.

**Organization** We devote Section 2 to treating some pertinent results and definitions for Riemannian manifolds. In Section 3 we consider the stable, local bases constructed in [20, 18, 19], which have many desirable properties but are computationally infeasible due to their cumbersome construction—each basis function of this type requires $\# \Xi$ nonzero kernel coefficients in its construction and, moreover, computing these requires $O((\# \Xi)^3)$ operations. An analysis of these coefficients show that they drop off rapidly—this is demonstrated in Section 3. The rapid decay of these coefficients leads to the (theoretical) existence of efficiently constructed bases, but sadly does not indicate the desired construction – this is treated in Section 5. In Section 6 we give numerical evidence to bolster the results of the previous sections, by giving results of experiments that show how rapidly the Lagrange basis and the coefficients decay. In this section we give some examples of techniques that fail to deliver, and provide some examples of families that seem to have the desired properties which have not been validated theoretically.

## 2 Geometric background

Throughout this paper, $\mathbb{M}$ denotes a compact, complete $d$-dimensional Riemannian manifold. The Riemannian metric for $\mathbb{M}$ is $g$, which defines an inner product $g_p(\cdot, \cdot) = \langle \cdot, \cdot \rangle_{g, p}$ on each tangent space $T_p \mathbb{M}$; the corresponding norm is $| \cdot |_{g, p}$.

The Riemannian metric is employed to measure arc length of a curve $\gamma$ via $\int_a^b |\dot{\gamma}|_{g, p} \, dt$. Geodesics are curves $\gamma : \mathbb{R} \to \mathbb{M}$ that locally minimize the arc length functional giving rise to
a distance function
\[
\text{dist}(p, q) = \min_{\gamma(0)=p} \int_0^1 |\gamma|_{g,p} dt.
\]

We denote the ball in \( M \) centered at \( x \) having radius \( r \) by \( B(x,r) \). Given a finite set \( \Xi \subset M \), we define its mesh norm (or fill distance) \( h \) and the separation radius \( q \) to be:
\[
h := \sup_{x \in M} \text{dist}(x, \Xi) \quad \text{and} \quad q := \frac{1}{2} \inf_{\xi, \zeta \in \Xi, \xi \neq \zeta} \text{dist}(\xi, \zeta). \tag{2.1}
\]
The mesh norm measures the density of \( \Xi \) in \( M \). The separation radius determines the spacing of \( \Xi \). The mesh ratio \( \rho := h/q \) measures the uniformity of the distribution of \( \Xi \) in \( M \). We say that the point set \( \Xi \) is quasi-uniformly distributed, or simply that \( \Xi \) is quasi-uniform if \( \Xi \) belongs to a class of finite subsets with mesh ratio bounded by a constant \( \rho_0 \).

The metric \( g \) also induces an invariant volume measure \( d\mu \) on \( M \). The local form of the measure is \( d\mu(x) = \sqrt{\det(g)} dx^1 \cdots dx^d \), where \( \det(g) = \det(g_{ij}) \). We indicate the measure of subsets \( \Omega \subset M \) by \( \text{vol}(\Omega) \). The integral, and the \( L_p \) spaces for \( 1 \leq p \leq \infty \), are defined with respect to this measure. The embeddings
\[
C(M) \subset L_p(M) \quad \text{for} \quad 1 \leq p \leq \infty \quad \text{and} \quad L_p(M) \subset L_q(M) \quad \text{for} \quad 1 \leq q \leq p \leq \infty
\]
hold. In addition, \( L_2 \) is a Hilbert space equipped with the inner product \( \langle \cdot, \cdot \rangle : \langle f, g \rangle \mapsto \langle f, g \rangle \).

**Sobolev spaces on subsets of \( M \)** Sobolev spaces on subsets of a Riemannian manifold can be defined in an invariant way, using the covariant derivative (or connection) \( \nabla \) (cf. [1]) which maps tensor fields of rank \( j \) to tensor fields of rank \( j + 1 \). The \( k \)th covariant derivative of a function is a rank \( k \) tensor field and is denoted \( \nabla^k f \). For \( k = 1 \), the covariant derivative in local coordinates is simply the usual expression for the “gradient” – it can be written simply as \( (\nabla f(x))_j = \frac{\partial f}{\partial x^j}(x) \). For \( k = 2 \), the “Hessian” tensor involves Christoffel symbols \( \Gamma^m_{ij} \) and can be expressed as \( (\nabla^2 f(x))_{ij} = \frac{\partial^2 f}{\partial x^i \partial x^j}(x) - \sum_{m=1}^4 \Gamma^m_{ij}(x) \frac{\partial f}{\partial x^m}(x) \). Higher order covariant derivatives have an analogous expression, using higher order derivatives of the Christoffel symbols – see [20 Eqn. (3)].

**Definition 2.1 ([1] p. 32).** Let \( \Omega \subset M \) be a measurable subset. We define the Sobolev space \( W^m_2(\Omega) \) to be all \( f : M \to \mathbb{R} \) such that, for \( 0 \leq k \leq m \), \( |\nabla^k f|_{g,p} \) in \( L_2(\Omega) \) with associated norm
\[
\|f\|_{m,\Omega}^2 := \|f\|_{W^m_2(\Omega)}^2 := \left( \sum_{k=0}^m \int_\Omega |\nabla^k f|_{g,p}^2 d\mu(p) \right)^{1/2}, \tag{2.2}
\]
coming from the Sobolev inner product
\[
\langle f, g \rangle_{m,\Omega} := \langle f, g \rangle_{W^m_2(\Omega)} := \sum_{k=0}^m \int_\Omega \langle \nabla^k f, \nabla^k g \rangle_{g,p} d\mu(p). \tag{2.3}
\]

When \( \Omega = M \), we may suppress the domain: \( \langle f, g \rangle_m = \langle f, g \rangle_{m,M} \) and \( \|f\|_m = \|f\|_{m,M} \).
Metric equivalence  The exponential map allows us to compare the Sobolev norms we’ve just introduced, to standard Euclidean Sobolev norms as follows:

**Lemma 2.2** ([20 Lemma 3.2]). For \( m \in \mathbb{N} \) and \( 0 < r < r_M/3 \), there are constants \( 0 < c_1 < c_2 \) so that for any measurable \( \Omega \subset B_r \), for all \( j \in \mathbb{N}, j \leq m \), and for any \( p_0 \in \mathbb{M} \), the equivalence

\[
c_1 \| u \circ \text{Exp}_{p_0} \|_{W^j_2(\Omega)} \leq \| u \|_{W^j_2(\text{Exp}_{p_0}(\Omega))} \leq c_2 \| u \circ \text{Exp}_{p_0} \|_{W^j_2(\Omega)}
\]

holds for all \( u : \text{Exp}_{p_0}(\Omega) \to \mathbb{R} \). The constants \( c_1 \) and \( c_2 \) depend on \( r \) and \( m \) but they are independent of \( \Omega \) and \( p_0 \).

3 The Lagrange basis

For a manifold \( \mathbb{M} \), a positive definite kernel \( k : \mathbb{M} \times \mathbb{M} \to \mathbb{R} \) and a set of centers \( \Xi \subset \mathbb{M} \), we are concerned with the robustness of the Lagrange basis \((\chi_\zeta)_{\zeta \in \Xi}\) for \( S(k, \Xi) \), where \( \chi_\zeta(\xi) = \delta_{\xi, \zeta} \) for all \( \xi \in \Xi \). The Lagrange basis plays a central role in most interpolation problems, and certainly this is the case for radial basis function and kernel interpolation. Decay of the Lagrange basis and analytic consequences have notably been considered in [24, 9, 29, 14].

3.1 The kernels considered

More recently, [20, 18, 19], develop a theory for fast decay and stability of the Lagrange basis associated with certain positive definite and conditionally positive definite kernels.

- "Sobolev kernels" denoted by \( \kappa_m \) were introduced in [20] for any smooth, complete, compact and connected Riemannian manifold. These are the reproducing kernels for the Sobolev inner product\(^1\) for \( W^m_2(M) \) when \( m > d/2 \):

\[
(u, v) \mapsto (u, v)_{W^m_2(\Omega)} = \sum_{k=0}^{m} \int_{\Omega} \langle \nabla^k f, \nabla^k g \rangle_{g,p} \, d\mu(p).
\]

- This was extended in [19] to treat a broader class of kernels on certain manifolds called kernels of polyharmonic and related type (these are discussed in Section 4.2).

- Included in the class considered in [19] are restricted surface splines on \( S^d \) which are kernels of the form \( k_m(x, \alpha) = \phi(x \cdot \alpha) \) where

\[
\phi(t) = \begin{cases} 
(1 - t)^{m-d/2} & \text{for } d \text{ odd} \\
(1 - t)^{m-d/2} \log(1 - t) & \text{for } d \text{ even}
\end{cases}
\]

These kernels are conditionally positive definite, meaning that interpolants are constructed by adding an auxiliary function. (In this case, a low degree spherical harmonic.) See Section 4.1 below.

\(^1\)In fact, the inner product can be weighted as \( \sum_{k=0}^{m} c_k \int_{\Omega} \langle \nabla^k f, \nabla^k g \rangle_{g,p} \, d\mu(p) \) with non-negative weights \( c_k \) for which \( c_0 \) and \( c_m \) are positive – each such reweighting gives a different inner product and a different Sobolev spline.
The expansion of the functions $\phi$ in terms of Gegenbauer polynomials by Baxter and Hubbert, [2], leads to Fourier (spherical harmonic) expansions of the kernels, and from there to their characterization as Green’s functions for elliptic differential operators. These operators are of polyharmonic type – they are of the form $Q(\Delta) = \prod_{j=1}^{m}(\Delta - r_j)$ for some real numbers $r_1, \ldots, r_m$. This, in turn, permits an understanding of the approximation power of the kernel, as investigate in [25, 17]: for functions having $L_p$ smoothness $s$ up to order $2m$ (namely, for target functions in smoothness spaces including $B_{p,q}^s(S^d), W_p^s(S^d), C^s(S^d)$ with $s \leq 2m$),

$$\text{dist}_p(f, S(k, \Xi)) = O(h^s).$$

Here, the space $S(k, \Xi)$ is modified by addition of low degree spherical harmonic terms $\Pi = \{Y_{\ell,m} \mid \ell \leq \lfloor m - d/2 \rfloor\}$ (this is described in Section 4.3 below).

- Surface splines on $SO(3)$ which are of the form $k(x, \alpha) = \phi(\omega(\alpha^{-1}x))$ with $\omega(x)$ the angle of rotation of $x$ (which is a left and right invariant metric on the group) and

$$\phi(t) = (\sin(t/2))^{m-3/2}.$$

In [21], an expansion of $\phi$ in even Chebyshev polynomials of the second kind leads to a Fourier (Wigner D-function) expansion of the kernel $k$. As in the spherical case, this leads to its characterization as a Green’s function for an operator of polyharmonic type on $SO(3)$, and to a realization of its approximation power: again, for $f \in B_{p,q}^s(SO(3)), W_p^s(SO(3)), C^s(SO(3))$ with $s \leq 2m$ we have

$$\text{dist}_p(f, S(k, \Xi)) = O(h^s).$$

**Restricted kernels** An alternative approach, taken in [15], is to consider the manifold $M$ as embedded in an ambient Euclidean space $\mathbb{R}^n$, and to use the restriction of a radial basis function – a Euclidean (conditionally) positive definite kernel satisfying rotational symmetry (of which there are many prominent examples) – as a (conditionally) positive definite kernel on $M$. In a sense, this is a completely different approach, in the sense that such kernels are almost never fundamental solutions to differential operators, a key point of [19]. On the other hand, such kernels may be easily localized in the ambient space $\mathbb{R}^n$, which may lead to an effective way of localizing and preconditioning the restricted kernels. Although the theory developed in Sections 3 and 4 does not address such kernels, we include a numerical example in Section 6.

### 3.2 Analytic properties of the Lagrange basis

The theory developed in [20] [15] [19] addresses analytic properties of bases for $S(k, \Xi)$, related to locality, stability of approximation and interpolation. In particular the following are shown.

**Locality.** The Lagrange basis is a local bases for $S(k, \Xi)$. That is,

$$|\chi_\xi(x)| \leq C \exp(-\nu \frac{\text{dist}(x, \xi)}{h}).$$
Stability of interpolation. Interpolation is stable: the Lebesgue constant is bounded (and more generally, is the $p$ norm of the interpolant is controlled by the $\ell_p$ norm of the data).

$L_p$ conditioning. There are constants depending only on $c_1, c_2$ such that $c_1 \|a\|_{\ell_p} \leq \|\sum_{j=1}^{N} a_\xi \chi_\xi\|_{L_p} \leq c_2 \|a\|_{\ell_p}$, with $c_1, c_2$ depending only on $m, M$ and the mesh ratio $\rho$. In particular, they are independent of $\#\Xi = \dim(S(k,\Xi))$, and, after a suitable normalization, independent of $p$.

Marcinkiewicz-Zygmund property. The space $S(k,\Xi)$ possess a Marcinkiewicz-Zygmund property relating samples to the size of the function. For $s \in S(k,\Xi)$, this means that the norms $\|\xi \mapsto s(\xi)\|_{\ell_p(\Xi)}$ and $\|s\|_{L_p}$ are equivalent, with constants involved independent of $\#\Xi$.

Stability of approximation in $L_p$. Approximation by $L_2$ projection is stable in $L_p$ for $1 \leq p \leq \infty$. In particular, the orthogonal projector with range $S(k,\Xi)$ can be continuously extended to each $L_p$, and it has bounded operator norm independent of $\#\Xi$.

4 Lagrange function coefficients

In this section we give theoretical results for the coefficients in the kernel expansion of Lagrange functions. In the first part we give a formula, relating these coefficients to native space inner products of the Lagrange functions themselves (this is Proposition 4.2). We then obtain estimates on the decay of these coefficients for a class of kernels on certain compact Riemannian manifolds (two point homogeneous spaces).

4.1 Interpolation with conditionally positive definite kernels

The kernels we consider in this article are conditionally positive definite on the compact Riemannian manifold. As a reference on this topic, we suggest [10, Section 4].

Definition 4.1. A kernel is conditionally positive definite with respect to a finite dimensional space $\Pi$ if, for any set of centers $\Xi$, the matrix $C_\Xi := (k(\xi,\zeta))_{\xi,\zeta \in \Xi}$ is positive definite on the subspace of all vectors $\alpha \in \mathbb{C}^\Xi$ satisfying $\sum_{\xi \in \Xi} \alpha_\xi p(\xi) = 0$ for $p \in \Pi$.

This is a very general definition which we will make concrete in the next subsections. Given a complete orthonormal basis $(\phi_j)_{j \in \mathbb{N}}$, of continuous functions (i.e., $\|\phi_j\|_\infty = 1$) any kernel

$$k(x, y) := \sum_{j \in \mathbb{N}} \tilde{k}(j) \varphi_j(x) \varphi_j(y)$$

with coefficients $\tilde{k} \in \ell_2(\mathbb{N})$ for which all but finitely many coefficients $\tilde{k}(j)$ are positive (negative) is conditionally positive definite with respect to $\Pi_\mathcal{J} = \text{span}(\phi_j \mid j \in \mathcal{J})$, where $\mathcal{J} = \{j \mid \tilde{k}(j) \leq 0\}$, since, evidently,

$$\sum_{\xi \in \Xi} \sum_{\zeta \in \Xi} \alpha_\xi k(\xi, \zeta) \alpha_\zeta = \sum_{\xi \in \Xi} \sum_{\zeta \in \Xi} \alpha_\xi \alpha_\zeta \left( \sum_{j \in \mathbb{N}} \tilde{k}(j) \phi_j(\xi) \phi_j(\zeta) \right) = \sum_{j \in \mathcal{J}} \tilde{k}(j) \sum_{\xi, \zeta \in \Xi} \alpha_\xi \phi_j(\xi) \alpha_\zeta \phi_j(\zeta) = \sum_{j \notin \mathcal{J}} \tilde{k}(j) \|\alpha \phi_j\|_{\ell_2(\Xi)}^2 > 0$$
provided $\sum_{\xi} \alpha_{\xi} \phi_j(\xi) = 0$ for $j$ satisfying $\tilde{k}(j) \leq 0$.

In this case if the set of centers $\Xi \subset M$ is unisolvent with respect to $\Pi_J = \text{span}(\varphi_j \mid j \in J)$ (meaning that $p \in \Pi_J$ and $p(\xi) = 0$ for $\xi \in \Xi$ implies that $p = 0$) then the system of equations

$$
\begin{align*}
\sum_{\xi \in \Xi} a_{\xi} k(\xi, \cdot) + \sum_{j \in J} b_j \varphi_j(\xi) &= y_{\xi} \quad \xi \in \Xi \\
\sum_{\xi \in \Xi} a_{\xi} \varphi_j(\xi) &= 0 \quad j \in J
\end{align*}
$$

(4.1)

has a unique solution in $C^{\Xi} \times C^J$ for each data sequence $y = (y_{\xi})_{\xi \in \Xi} \in C^{\Xi}$.

By writing the same system in matrix form, with collocation matrix $K_\Xi = (k(\xi, \zeta))_{\xi, \zeta \in \Xi}$ and auxiliary matrix $\Phi = (\varphi_j(\xi))_{\xi, j \in J}$.

$$
\begin{pmatrix}
K_\Xi & \Phi \\
\Phi^* & 0
\end{pmatrix}
\begin{pmatrix}
a \\
b
\end{pmatrix}
=
\begin{pmatrix}
y \\
0
\end{pmatrix}
$$

(4.2)

When data is sampled from a continuous function at points $\Xi$ (i.e., $y_{\xi} = f(\xi)$) that are unisolvent for $\Pi_J$ this solution generates a continuous interpolant:

$$
I_\Xi f = I_{k, J, \Xi} f = \sum_{\xi \in \Xi} a_{\xi} k(\cdot, \xi) + p_f
$$

where $p_f = \sum_{j \in J} b_j \varphi_j \in \Pi_J$ and $\sum_{\xi \in \Xi} a_{\xi} p(\xi) = 0$ for all $p \in \Pi_J$. Indeed, this interpolant is unique among functions from the space

$$
S(k, \Xi) := S(k, \Xi, \Pi_J) := \left\{ \sum_{\xi \in \Xi} a_{\xi} k(\cdot, \xi) + p_f \bigg| p_f \in \Pi_J, \quad \sum_{\xi \in \Xi} a_{\xi} p(\xi) = 0 \right\}
$$

(4.3)

It has a dual role as the minimizer of the semi-norm $\| \cdot \|_{k, J}$ induced in the usual way from the “native space” semi-inner product

$$
\langle u, v \rangle_{k, J} = \sum_{j \in \Xi} \hat{u}(j) \varphi_j \sum_{j \in \Xi} \hat{v}(j) \varphi_j = \sum_{j \in J} \frac{\hat{u}(j) \hat{v}(j)}{\tilde{k}(j)}.
$$

(4.4)

When $u, v \in S(k, \Xi)$ – meaning that they have the expansion $u = \sum_{\xi \in \Xi} a_{1, \xi} k(\cdot, \xi) + p_u$ and $v = \sum_{\xi \in \Xi} a_{2, \xi} k(\cdot, \xi) + p_v$ with coefficients $(a_{j, \xi})_{\xi \in \Xi} \perp (\Pi_J)_{\Xi}$ for $j = 1, 2$ – then the semi-inner product is

$$
\langle u, v \rangle_{k, J} = \sum_{\xi \in \Xi} \sum_{\xi \in \Xi} a_{1, \xi} a_{2, \xi} k(\xi, \zeta)
$$

We can use this expression of the inner product to investigate the kernel expansion of the Lagrange function.

---

\[\text{Meaning that } p|_\Xi = 0 \text{ for } p \in \Pi_J \text{ implies that } p = 0.\]
Lemma 4.2. Let $k = \sum_{j \in \mathbb{N}} \bar{\phi}(j) \varphi_j \bar{\varphi}_j$ be a conditionally positive definite kernel with respect to the space $\Pi_{\mathcal{J}} = \text{span}_{j \in \mathcal{J}} \varphi_j$, and let $\Xi$ be unisolvent for $\Pi_{\mathcal{J}}$. Then $\chi_\eta \in S(k, \Xi)$ (the Lagrange function centered at $\eta$) has the kernel expansion $\chi_\eta(x) = \sum_{\xi \in \Xi} A_{\eta, \xi} k(\cdot, \xi) + p_\zeta$ with coefficients

$$A_{\eta} = (A_{\eta, \xi})_{\xi \in \Xi} = ((\chi_\zeta(x), \chi_\eta(x))_{k, \mathcal{J}})_{\xi \in \Xi}.$$ 

**Proof.** Select two centers $\zeta, \eta \in \Xi$ with corresponding Lagrange functions $\chi_\zeta$ and $\chi_\eta \in S(k, \Xi)$. Because $A_\zeta$ and $A_\eta$ are both orthogonal to $(\Pi_{\mathcal{J}})|_{\Xi}$, we have

$$\langle \chi_\zeta, \chi_\eta \rangle_{k, \mathcal{J}} = \sum_{\xi_1, \xi_2 \in \Xi} \sum_{\ell_1, \ell_2 \in \Xi} A_{\zeta, \xi_1} A_{\eta, \xi_2} k(\xi_1, \xi_2) = \langle K_{\Xi} A_\zeta, A_\eta \rangle_{\ell_2(\Xi)}.$$

Now define $P := \Phi(\Phi^*\Phi)^{-1}\Phi^* : \ell_2(\Xi) \to (\Pi_{\mathcal{J}})|_{\Xi} \subset \ell_2(\Xi)$ to be the orthogonal projection onto the subspace of samples of $\Pi_{\mathcal{J}}$ on $\Xi$ and let $P^\perp = \text{Id} - P$ be its complement. Then for any data $y$, (4.11) yields coefficient vectors $A$ and $b$ satisfying $P^\perp A = A$ and $P^\perp b = 0$, hence $P^\perp K_{\Xi} P^\perp A = P^\perp K_{\Xi} A = P^\perp y$. Because $P^\perp : \ell_2(\Xi) \to \ell_2(\Xi)$ is also an orthogonal projector, and therefore self-adjoint, it follows that

$$\langle \chi_\zeta, \chi_\eta \rangle_{k, \mathcal{J}} = \langle K_{\Xi} A_\zeta, A_\eta \rangle_{\ell_2(\Xi)} = \langle K_{\Xi} A_\zeta, P^\perp A_\eta \rangle_{\ell_2(\Xi)} = \langle P^\perp K_{\Xi} A_\zeta, A_\eta \rangle_{\ell_2(\Xi)} = \langle P^\perp e_\zeta, A_\eta \rangle_{\ell_2(\Xi)}.$$ 

In the last line, we have introduced the sequence $e_\zeta = (\delta_{\zeta, \xi})_{\xi \in \Xi}$ for which $K_{\Xi} A_\zeta + p_\zeta|_{\Xi} = e_\zeta$ which implies that $P^\perp K_{\Xi} A_\zeta = P^\perp e_\zeta$. Using once more the fact that $P^\perp$ is self-adjoint, and that $A_\eta$ is in its range, we have

$$\langle \chi_\zeta(x), \chi_\eta(x) \rangle_{k, \mathcal{J}} = \langle P^\perp e_\zeta, A_\eta \rangle = \langle e_\zeta, P^\perp A_\eta \rangle = \langle e_\zeta, A_\eta \rangle$$

and the lemma follows. \qed

### 4.2 Estimating Lagrange function coefficients

In [20, 19], it has been shown that Lagrange functions decay rapidly away from the center. We can use this characterization of the Lagrange function to estimate the decay of its coefficients. In this section we use Proposition 4.2 to estimate the size of coefficients first for the class of strictly positive definite functions developed in [20]. Then we attempt to do the same for the more general class of kernels of polyharmonic and related type of [19].

**Sobolev kernels on compact Riemannian manifolds** We begin by considering kernels $k = \kappa_m$ with native space inner product given by an expression like

$$\langle u, v \rangle_{\kappa_m} := \langle u, v \rangle_{\kappa_m, M} := \int_M \beta(u, v) x d\mu$$

where $\beta$ is a pointwise bilinear form $\beta(u, v) x := \sum_{j=0}^m c_j \langle \nabla^j u, \nabla^j v \rangle x$ with the condition that $m > d/2$ and $c_0, c_m \neq 0$ and $c_j \geq 0$ for all $j = 0, \ldots, m$. Such kernels were considered in

\[ \text{in such cases, the kernel } \kappa_m \text{ is naturally (strictly) positive definite and moreover, } \kappa_m \text{ is the fundamental solution to the elliptic operator } L_m = \sum_{j=0}^m c_j (\nabla^j)^* \nabla^j \]
eralization to inner products for subsets $\Omega$: namely $\langle L \rangle$ the spectrum of expressed as $\sum_{k}$. It is possible to estimate $|\langle \chi_{\xi}, \chi_{\zeta} \rangle_{\kappa_{m}}|$: 

$$
|\langle \chi_{\xi}, \chi_{\zeta} \rangle_{\kappa_{m}}| \leq \|\chi_{\xi}\|_{W^{m}_{2}(\mathbb{M}\setminus b(\zeta, \frac{\text{dist}(x,\zeta)}{\kappa}))} \|\chi_{\zeta}\|_{W^{m}_{2}(\mathbb{M})} + \|\chi_{\xi}\|_{L^{2}(\mathbb{M})} \|\chi_{\zeta}\|_{L^{2}(\mathbb{M}\setminus b(\zeta, \frac{\text{dist}(x,\zeta)}{\kappa}))}
$$

By way of [20, Corollary 4.4] we have that $|A_{\xi,\zeta}| = |\langle \chi_{\xi}, \chi_{\zeta} \rangle_{\kappa_{m}}| \leq C_{m,M} q^{d-2m} e^{-\nu \frac{\text{dist}(x,\zeta)}{\kappa}}$.

Unfortunately, this family of kernels is not suitable for treating practical problems. In particular, the kernels having native space inner products of the form (1.5), even when $\mathbb{M}$ is the sphere, are not known to have closed form representations in terms of the spatial variable (despite being zonal and having a simple and satisfying Fourier-Legendre expansion).

To remedy this, we remove the restriction that the coefficients $c_{j}$ are non-negative (although $c_{m}$ must be positive). An immediate consequence of this is that we must contend with a conditionally positive definite kernel. The upshot is that, for a large class of interesting manifolds (including spheres and projective spaces) we can write the Dirichlet form (1.5) as linear combinations of powers of the Laplace–Beltrami operators. The motivation for this approach is that the restricted surface splines are fundamental solutions for operators of this type. We now describe this.

**Polyharmonic and related kernels on 2 point homogeneous spaces** Let $\mathbb{M}$ be a compact, two point homogeneous space. Included among these are spheres, $SO(3)$ and various projective spaces. For our purposes, this is a metric space with distance function $\text{dist}(x, y)$ and measure $\mu$ for which $\mu(b(x, r)) = \mu\{y \mid \text{dist}(x, y) \leq r\} \sim r^{d}$. Because it is compact, there is a Laplace–Beltrami operator $\Delta$ with countable spectrum $\sigma(\Delta) = \{\lambda_{0}, \lambda_{1}, \ldots\}$. Denote the corresponding orthogonal, $L_{2}$ normalized eigenfunctions for $\Delta$ by $(\psi_{j})_{j \in \mathbb{N}}$.

For such a manifold and for any $k \in \mathbb{N}$, the operator $(\nabla^{k})^{\ast} \nabla^{k}$ can be expressed as $\sum_{j=0}^{k} b_{j} \Delta^{j}$ with $b_{k} = (-1)^{k}$. Consequently, any operator of the form $\sum_{j=0}^{k} c_{j} (\nabla^{j})^{\ast} \nabla^{j}$ can be expressed as $\sum_{j=0}^{k} b_{j} \Delta^{j}$ with $b_{k} = (-1)^{k} c_{k}$ and vice-versa:

$$
\forall (b_{0}, \ldots, b_{m}) \exists (c_{0}, \ldots, c_{m}) \text{ with } b_{m} = (-1)^{m} c_{m} \text{ and } \sum_{j=0}^{m} b_{j} \Delta^{j} = \sum_{j=1}^{m} c_{j} (\nabla^{j})^{\ast} \nabla^{j}. \quad (4.6)
$$

Suppose that the kernel $k_{m} : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ acts as the Green’s function for the elliptic operator $L_{m} := \sum_{j=0}^{m} b_{j} \Delta^{j} = Q(\Delta)$, in the sense that

$$
f = \int_{\mathbb{M}} k_{m}(\cdot, \alpha)L_{m}[f(\alpha) - p_{f}(\alpha)] \, d\alpha + p_{f}
$$

where $p_{f}(x)$ is the projection on $\Pi J$, $p_{f} = \sum (f, \psi_{j})_{L_{2}(\mathbb{M})} \psi_{j}$, and the complementary part of the spectrum of $L_{m}$, $\{Q(\lambda_{j}) \mid j \notin J\} \subset \sigma(L_{m})$, is real and lies to one side of 0 (without loss,
we can take $\sigma(\mathcal{L}_m) \subset (0, \infty)$ – namely, by considering $-k$ if needed; this is equivalent to taking $b_m > 0$ since the spectrum of $\Delta$ has $\infty$ as an accumulation point and $(\lambda_k)^m > |\sum_{j=0}^{m-1} b_j (\lambda_k)^j|$ for all but finitely many $k$). Such a kernel is said to be of polyharmonic or related type.

The native space “inner product” on subsets It follows directly that $k_m$ is conditionally positive definite with respect to $\Pi_j$. What’s more, when $\mathcal{L}_m \Pi_j = \{0\}$, the native space semi-inner product can be expressed as

$$\langle u, v \rangle_{k_m, \mathcal{J}} = (\mathcal{L}_m u, v)_{L_2(\mathcal{M})} = \int_\mathcal{M} \beta(u, v)_x d\mu(x)$$

with $\beta(u, v)_x = \sum_{k=0}^{m} c_k (\nabla^k u, \nabla^k v)_x$ and $c_0, \ldots, c_m$ guaranteed by (4.6). The latter expression allows us to extend naturally the native space inner product to measurable subsets $\Omega$ of $\mathcal{M}$. Namely,

$$\langle u, v \rangle_{\Omega, k_m, \mathcal{J}} := \int_\Omega \beta(u, v)_x d\mu(x).$$

This has the desirable property of set additivity: for sets $A$ and $B$ with $\mu(A \cap B) = 0$, we have $\langle u, v \rangle_{A \cup B, k_m, \mathcal{J}} = \langle u, v \rangle_{A, k_m, \mathcal{J}} + \langle u, v \rangle_{B, k_m, \mathcal{J}}$. Unfortunately, since some of the coefficients $c_k$ may be negative, $\beta(u, u)$ and $\langle u, v \rangle_{\Omega, k_m, \mathcal{J}}$ may assume negative values for some $u$: in other words, the bilinear form $(u, v) \mapsto \langle u, v \rangle_{\Omega, k_m, \mathcal{J}}$ is only an indefinite inner product.

However, when $\Omega$ has Lipschitz boundary and $u$ has many zeros, we can relate the quadratic form $\|u\|^2_{\Omega, k_m, \mathcal{J}} = \langle u, u \rangle_{\Omega, k_m, \mathcal{J}}$ to a Sobolev norm $\|u\|^2_{W^m_2(\Omega)}$. Arguing as in [19 (4.2)], we see that

$$c_m \|u\|^2_{W^m_2(\Omega)} - \left( \max_{j \leq m-1} |c_j| \right) \|u\|^2_{W^{m-1}_2(\Omega)} \leq \int_\Omega \beta(u, u)_x d\mu(x) \leq \left( \max_{j \leq m} |c_j| \right) \|u\|^2_{W^m_2(\Omega)}.$$

If $u|_{\Xi} = 0$ on a set $\Xi$ with $h(\Xi, \Omega) \leq h_0$ with $h_0$ determined only by the boundary of $\Omega$ (specifically the radius and aperture of an interior cone condition satisfied by $\partial \Omega$), Theorem A.11 of [19] guarantees that $\|u\|^2_{W^{m-1}_2(\Omega)} \leq C h^2 \|u\|^2_{W^m_2(\Omega)}$ with $C$ depending only on the order $m$, the global geometry of $\mathcal{M}$ and the roughness of the boundary (in this case, depending only on the aperture of the interior cone condition). Thus, by choosing $h$ sufficiently small, $h \leq h^*$, where $h^*$ satisfies the two conditions

$$h^* \leq h_0 \quad \text{and} \quad C(h^*)^2 \times \left( \max_{j \leq m} |c_j| \right) \leq \frac{|c_m|}{2}, \quad (4.7)$$

we have

$$\frac{c_m}{2} \|u\|^2_{W^m_2(\Omega)} \leq \|u\|^2_{\Omega, k_m, \mathcal{J}} \leq \left( \max_{j \leq m} |c_j| \right) \|u\|^2_{W^m_2(\Omega)}.$$

The threshold value $h^*$ depends on the coefficients $c_j$ as well as the radius $R_\Omega$ and aperture $\phi_\Omega$ of the cone condition for $\Omega$. When $\Omega$ is an annulus of sufficiently small inner radius, the cone parameters can be replaced by a single global constant, and $h^*$ can be taken to depend only on $c_0, \ldots, c_m$. In other words, only on $k_m$ – cf. [19 Corollary A.16].
A direct consequence of this is positive definiteness for such functions, \( \|u\|_{\Omega,k_m,J} \geq 0 \) with equality only if \( u|_{\Omega} = 0 \). From this, we have a version of the Cauchy-Schwarz inequality: if \( u \) and \( v \) share a set of zeros \( Z \) (i.e., \( u|_Z = v|_Z = \{0\} \)) that is sufficiently dense in \( \Omega \), then

\[
|\langle u, v \rangle_{\Omega,k_m,J}| \leq \|u\|_{\Omega,k_m,J} \|v\|_{\Omega,k_m,J}
\]

follows (sufficient density means that \( h(Z, \Omega) < h^* \) as above).

**Decay of coefficients for kernels of polyharmonic and related type** Fortunately, Lagrange functions have many zeros, and [19, Lemma 5.1] guarantees that the Lagrange function \( \chi_\xi \) satisfies the bulk chasing estimate there is a fixed constant \( 0 \leq \epsilon < 1 \) so that for radii \( r \) less than a constant \( r_M \) depending on \( \mathbb{M} \) (for a compact, 2-point homogeneous space, the injectivity radius is \( r_M = \text{diam}(\mathbb{M})/2 \)) the estimate \( \|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M} \setminus b(\xi,r))} \leq \epsilon \|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M} \setminus b(\xi,r-\frac{h}{4h_0}))} \) holds. In other words, a fraction of \( 1-\epsilon \) of the bulk of the tail \( \|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M} \setminus b(\xi,r))} \) is to be found in the annulus \( b(\xi,r) \setminus b(\xi,r-\frac{h}{4h_0}) \) of width \( \frac{h}{4h_0} \) (with constant of proportionality \( \frac{h}{4h_0} \) depending only on \( \mathbb{M} \), \( m \) and the boundary of \( \Omega \)). Provided \( r \leq r_M \), it is possible to iterate this \( n \) times for \( \frac{nh}{4h_0} \leq r \). It follows that there is \( \nu = -4h_0 \log \epsilon > 0 \) so that

\[
\|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M} \setminus b(\xi,r))} \leq \epsilon^n \|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M})} \leq Ce^{-\nu r/h} \|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M})}.
\]

By [19, (5.1)] (a simple comparison of \( \chi_\xi \) to a smooth “bump” \( \phi_\xi \) of radius \( q \) – also an interpolant to the delta data \( (\delta_\xi) \), but worse in the sense that \( \|\chi_\xi\|_{k_m,J} \leq \|\phi\|_{k_m,J} \) – see the proof of Theorem 4.3 below) we have

\[
\|\chi_\xi\|_{L^p_{\Omega}(\mathbb{M} \setminus b(\xi,r))} \leq Cq^{d/2-m}e^{-\nu r/2}.
\]

This leads us to our main result.

**Theorem 4.3.** Let \( \mathbb{M} \) be a compact, 2-point homogeneous manifold and let \( k_m \) be a kernel of polyharmonic or related type, so that the associated elliptic operator \( \mathcal{L}_m \) annihilates the polynomial space \( \Pi_J \). Let \( \rho > 0 \) be a fixed mesh ratio.

There exist constants \( h^*, \nu \) and \( C \) depending only on \( \mathbb{M} \) and \( k_m \) if \( \Xi \subset \mathbb{M} \) is sufficiently dense (i.e., \( h(\Xi, \mathbb{M}) \leq h^* \)) then the coefficients of the Lagrange function \( \chi_\xi = \sum_{\xi \in \Xi} A_{\xi,\xi} k_m(\cdot, \xi) + p_\xi \in S(\Xi, J) \) satisfy

\[
|A_{\xi,\xi}| \leq Cq^{d/2-m} \exp\left(-\frac{\nu \text{dist}(\xi, \zeta)}{h}\right).
\]

**Proof.** By Proposition 4.2 and set additivity, we have that

\[
|A_{\xi,\xi}| = \langle \chi_\xi, \chi_\zeta \rangle_{k_m,J} = \langle \chi_\xi, \chi_\zeta \rangle_{\Omega_\zeta, k_m,J} + \langle \chi_\xi, \chi_\zeta \rangle_{\Omega_\xi, k_m,J},
\]

where \( \Omega_\zeta = \{ \alpha \in \mathbb{M} \mid \text{dist}(\alpha, \zeta) < \text{dist}(\alpha, \xi) \} \), \( \Omega_\xi = \{ \alpha \in \mathbb{M} \mid \text{dist}(\alpha, \xi) < \text{dist}(\alpha, \zeta) \} \), and (modulo a set of measure zero) \( \mathbb{M} \setminus \Omega_\zeta = \Omega_\xi \). For a compact, 2-point homogeneous space, \( \Omega_\zeta \) and \( \Omega_\xi \) are two balls of radius \( \text{diam}(\mathbb{M})/2 \).
We can apply the Cauchy–Schwarz type inequality \( (4.8) \) to obtain
\[
|A_{\xi,\zeta}| \leq \|\chi_{\xi}\|_{W^{m}(\Omega,\mathcal{J})} \|\chi_{\zeta}\|_{W^{m}(\Omega,\mathcal{J})} + \|\chi_{\xi}\|_{\Omega} \|\chi_{\zeta}\|_{\Omega}
\]
\[
\leq \sqrt{\max_{j\leq m} |c_{j}| \left( \|\chi_{\xi}\|_{W^{m}(\Omega,\mathcal{J})} \|\chi_{\zeta}\|_{W^{m}(\Omega,\mathcal{J})} + \|\chi_{\xi}\|_{\Omega} \|\chi_{\zeta}\|_{\Omega} \right)}
\]

Since \( \Omega_{\xi} \subset b^c(\zeta, r) := M \setminus b(\zeta, \frac{1}{2}\text{dist}(\xi, \zeta)) \) and \( \Omega_{\zeta} \subset b^c(\xi, r) := M \setminus b(\xi, \frac{1}{2}\text{dist}(\xi, \zeta)) \), we can again employ set additivity and positive definiteness (this time \( \|\chi_{\xi}\|_{\Omega_{\xi},\mathcal{J}} \leq \|\chi_{\xi}\|_{M_{\zeta},\mathcal{J}} \)), which follows from the fact that \( M = \Omega_{\xi} \cup \Omega_{\zeta} \) and that \( \chi_{\xi} \) vanishes to high order in \( \Omega_{\zeta} \) — the same holds for \( \chi_{\zeta} \) to obtain
\[
|A_{\xi,\zeta}| \leq \sqrt{\max_{j\leq m} |c_{j}| \left( \|\chi_{\xi}\|_{W^{m}(b^c(\zeta, r))} \|\chi_{\zeta}\|_{b^c(\zeta, r)} + \|\chi_{\xi}\|_{b^c(\zeta, r)} \|\chi_{\zeta}\|_{b^c(\zeta, r)} \right)}
\]

The full energy of the Lagrange function can be bounded by comparing it to the energy of a bump function – for \( \chi_{\xi} \) this is \( \phi_{\delta} \), which can be defined on the tangent space by using a fixed, smooth, radial cutoff function \( \sigma: \phi_{\delta} \circ \text{Exp}_{x}(x) = \sigma(|x|/\delta) \). This is done in \( [19] \ (5.1) \) and we have that \( \|\chi_{\xi}\|_{M_{\zeta},\mathcal{J}} \) and \( \|\chi_{\xi}\|_{M_{\zeta},\mathcal{J}} \) are bounded by \( C_{q}^{d/2-m} \).

On the other hand, we can employ \( (4.9) \) to treat \( \|\chi_{\xi}\|_{W^{m}_{2}(b^c(\zeta, r))} \) and \( \|\chi_{\xi}\|_{W^{m}_{2}(b^c(\zeta, r))} \), which gives
\[
\|\chi_{\xi}\|_{W^{m}_{2}(b^c(\zeta, r))}, \|\chi_{\xi}\|_{W^{m}_{2}(b^c(\zeta, r))} \leq C_{q}^{d/2-m} e^{-\nu_{r}^{2} \frac{r_{\zeta}}{h}} = C_{q}^{d/2-m} e^{-\nu_{r}^{2} \frac{\text{dist}(\xi, \zeta)}{h^{2}}}.
\]
From this, the result follows.

**Note 1.** A similar argument shows that, on a compact, 2-point homogeneous manifold, the Lagrange function coefficients \( A_{\xi,\zeta} \) for a general kernel \( k_{m} \) of polyharmonic and related type (regardless of whether \( L_{m} \) annihilates \( \Pi_{\mathcal{J}} \)) decay like
\[
|A_{\xi,\zeta}| \leq C_{q}^{d-2m} \max(\exp(-\nu_{r}^{2} \frac{r_{\zeta}}{h}), h^{2m}).
\]
In such cases, the theory developed here and in \( [19] \) indicate a slower decay for Lagrange functions and coefficients (although it remains an open problem to determine if these rates can be improved, and by how much). In particular, this holds for the restricted surface splines \( k_{m}(x, \alpha) = (1 - x \cdot \alpha)^{m-d/2} \) on odd dimensional spheres \( (d \in 2\mathbb{N} + 1 \) and \( m > d/2) \) as well as the surface splines on \( \text{SO}(3) \) (see Section 3).

**Note 2.** Theorem \( [4, 3] \) holds for restricted surface splines \( k_{m}(x, \alpha) = (1-x \cdot \alpha)^{m-d/2} \log(1-x \cdot \alpha) \) on spheres of even dimension \( (d \in 2\mathbb{N} \) and \( m > d/2) \) – in particular for \( \mathbb{S}^{2} \). See Section \( 4 \) below for some numerical examples in this setting.

## 5 A better basis: truncating the Lagrange basis

We now want to show the existence of a good approximation to the Lagrange function \( \chi_{\xi} \) that uses many fewer elements in its kernel expansion than the \( N \) needed for \( \chi_{\xi} \) itself. To do
this, we will start with the expansion \( \chi_\xi = \sum_{\zeta \in \Xi} A_{\xi,\zeta} \kappa(\cdot, \zeta) \) and approximate it by a truncated expansion of the form

\[
\tilde{\chi}_\xi = \sum_{\zeta \in \Upsilon(\xi)} A_{\xi,\zeta} \kappa(\cdot, \zeta) = \chi_\xi - \sum_{\zeta \not\in \Upsilon(\xi)} A_{\xi,\zeta} \kappa(\cdot, \zeta),
\]

where \( \Upsilon(\xi) \subset \Xi \cap B(\xi, r(h)) \). Our goal is to show that, under the assumptions listed below, which apply to a wide class of kernels, we may take \( r(h) = K h |\log(h)|, \) where \( K > \frac{2m}{\nu} \), while maintaining \( \|\tilde{\chi}_\xi - \chi_\xi\|_\infty \leq Ch^J, \) \( J := K \nu - 2m \).

How many basis elements are used in expanding \( \tilde{\chi}_\xi \)? Doing a simple volume estimate shows that the number required is

\[
\#\Upsilon(\xi) = O((Kh |\log h|)^d/q^d) = O(|\log h|^d) = O((|\log N|^d) \ll N,
\]

where we have used \( h/q = \rho \) and \( N = O(h^{-d}) \).

One final remark before proceeding with the analysis. Finding \( \tilde{\chi}_\xi \) requires knowing the expansion for \( \chi_\xi \) and carrying out the truncation above. This is expensive, although it does have utility in terms of speeding up evaluations for interpolation when the same set of centers is to be used repeatedly. The main point is that we now know roughly how many basis elements are required to obtain a good approximation to \( \chi_\xi \). We are currently engaged in investigating cost effective algorithms to obtain approximate Lagrange functions similar to \( \tilde{\chi}_\xi \).

**First assumptions** We make the following three assumptions

1. The Lagrange functions decay at a rate \( |\chi_\xi| \leq C_L \exp\left( -\nu_L \frac{\text{dist}(x, \xi)}{h} \right) \).

2. The kernel coefficients of the Lagrange function decay like \( |A_{\xi,\zeta}| \leq C_c \exp\left( -\nu_c \frac{\text{dist}(x, \xi)}{h} \right) \).

3. The Lagrange basis is \( L_p \) stable in the sense that

\[
c_1 q^{d/p} \|a\|_{\ell_p(\Xi)} \leq \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{L_p(M)} \leq c_2 q^{d/p} \|a\|_{\ell_p(\Xi)}.
\]

We note that the family of restricted surface splines on \( S^d \) when \( d \in 2\mathbb{N} \) satisfy these three conditions (conditions 1 and 3 are in [19], while condition 2 follows from Theorem 4.3), as do the Sobolev splines on any compact Riemannian manifold \( M \) (condition 1 follows from [20], condition 3 from [18] and condition 2 from Theorem 4.3 again).

**Decay** By the estimate of coefficients in Theorem 4.3, it suffices to retain only the part of \( \Xi \) that is within \( Kh |\log h| \) from \( \xi \), since the coefficients we cut out have size roughly \( Ch^{d-2m} h^{K \nu} \). There are no more than \( \#\Xi \leq C_d h^{-d} \) of them on the \( d \)-sphere, and the kernel is uniformly bounded, so we have that

\[
|\tilde{\chi}_\xi(x) - \chi_\xi(x)| \leq Ch^{K \nu - 2m}.
\]
and we should choose $K > \frac{2m}{\nu}$ at least. Indeed, the pointwise estimate above shows that

$$|\tilde{\chi}_\xi(x)| \leq C \left( e^{-\nu \frac{\text{dist}(x, \xi)}{h}} + h^{K\nu - 2m} \right) \leq C \left( 1 + \frac{\text{dist}(x, \xi)}{h} \right)^{2m - K\nu},$$

which indicates that we may wish to choose $K$ even larger. This is at our discretion, but to preserve stability, we choose $K > \frac{2m + d}{\nu}$.

**Computational efficiency** Since we retain only the coefficients centered at a distance of $O(h|\log h|)$ from $\xi$, we use

$$\#\Upsilon(\xi) = O\left( \frac{(h|\log h|)^d}{q^d} \right) = O((\log N)^d)$$

coefficients (when centers are quasiuniform) to compute each basis function $\tilde{\chi}_\xi$.

**Stability** By the $L_\infty$ stability of the Lagrange basis, for $s \in S(k, \Xi)$, the samples $s|_\Xi =: (A_\xi)_{\xi \in \Xi}$ are bounded in the $\ell_\infty$ norm by $\|s\|_\infty / c_1$. Using the same coefficients but in the new basis $\tilde{\chi}_\xi$, we form $\tilde{s} = \sum A_\xi \tilde{\chi}_\xi$. The difference between the original and new function is

$$\|\tilde{s} - s\|_\infty \leq \|A\|_{\ell_\infty} \sum_{\xi} |\chi_\xi(x) - \tilde{\chi}_\xi(x)| \leq C \|A\|_{\ell_\infty} h^{K\nu - 2m} q^{-d} \leq Ch^{K\nu - 2m - d} \|s\|_\infty,$$

so

$$\|\tilde{s}\|_\infty \geq (1 - Ch^{K\nu - 2m - d}) \|s\|_\infty \geq c_1 (1 - Ch^{K\nu - 2m - d}) \|A\|_{\ell_\infty} \tag{5.2}$$

$$\|\tilde{s}\|_\infty \leq (1 + Ch^{K\nu - 2m - d}) \|s\|_\infty \leq c_2 (1 + Ch^{K\nu - 2m - d}) \|A\|_{\ell_\infty}. \tag{5.3}$$

This can be viewed in two ways.

- Provided $h$ is small enough, the family is a basis. There are $\#\Xi$ elements and they are linearly independent. (In particular, if the function is zero, all the coefficients are zero.)
- The family $(\tilde{\chi}_\xi)$ is stable in $L_\infty$, since the map $(A_\xi) \mapsto \tilde{s}$ is boundedly invertible.

We stress that it remains to be determined how actually to compute the basis – we have simply shown that a preconditioner exists that has complexity $O((\log N)^d)$.

**Theorem 5.1.** Let $\mathbb{M}$ be a compact, 2-point homogeneous manifold and let $k_m$ be a kernel of polyharmonic or related type, so that the associated elliptic operator $L_m$ annihilates the polynomial space $\Pi_J$. Let $\rho > 0$ be a fixed mesh ratio.

For sufficiently dense $\Xi$, with $h = h(\Xi, \mathbb{M}) \leq H$, with $H$ a constant depending only on $\rho$, $\mathbb{M}$ and $k_m$, there is a basis $(b_\xi)_{\xi \in \Xi}$ where each basis element $b_\xi = \sum_{\xi \in \Xi} A_{\xi, \xi} k_m(\cdot, \xi)$ is composed of kernels centered in the ball $B(\xi, -K|\log h|)$. The following are satisfied:

- The cost of constructing each $b_\xi$ is $\#\{\xi \mid A_{\xi, \xi} \neq 0\} \leq \tau (\log \#\Xi)^d$ with $\tau \leq CK^d$. 


• Each basis element exhibits polynomial decay of degree \( J := K \nu - 2m \): there exists \( C \) for which
\[
|b_\xi(x)| \leq C \left(1 + \frac{\text{dist}(x, \xi)}{h}\right)^{-J}.
\]

• The basis is \( L_p \) stable: there are \( c_1, c_2 \) for which
\[
c_1 q^{d/p} \|a\|_{\ell_p(\Xi)} \leq \left\| \sum_{\xi \in \Xi} a_\xi b_\xi \right\|_{L_p(M)} \leq c_2 q^{d/p} \|a\|_{\ell_p(\Xi)}.
\]

**Proof.** It remains to demonstrate the \( L_p \) stability of \( (b_\xi) = (\tilde{\chi}_\xi) \) for \( 1 \leq p < \infty \).

When \( p = 1 \), we consider a sequence \( a = (a_\xi)_{\xi \in \Xi} \in \ell_1(\Xi) \) and set \( \tilde{s} := \sum a_\xi \tilde{\chi}_\xi \). Hölder’s inequality gives
\[
\|\tilde{s} - s\|_{L_1(M)} \leq C \|a\|_{\ell_1(\Xi)} \mathrm{vol}(M) h^{K \nu - 2m}
\]
since for each \( x \) we have the estimate \( |\tilde{s}(x) - s(x)| \leq \left( \sum_{\xi \in \Xi} |a_\xi| \right) \max_{\xi \in \Xi} |\chi_\xi(x) - \tilde{\chi}_\xi(x)|. \)

Interpolating between (5.4) and (5.1) (i.e., interpolating the finite rank operator \( a \mapsto (s - \tilde{s}) \)) gives
\[
\|s - \tilde{s}\|_{L_p(M)} \leq Ch^{K \nu - 2m - d}(1 - 1/p) \|a\|_{\ell_p(\Xi)}.
\]

Therefore,
\[
\|s\|_{L_p(M)} - Ch^{K \nu - 2m - d}(1 - 1/p) \|a\|_{\ell_p(\Xi)} \leq \|\tilde{s}\|_{L_p(M)} \leq \|s\|_{L_1(M)} + Ch^{K \nu - 2m - d}(1 - 1/p) \|a\|_{\ell_p(\Xi)}
\]
and we have
\[
c_1 q^{d/p} \|a\|_{\ell_p(\Xi)} \sum_{J=0}^{d} (1 - Ch^{K \nu - 2m - d})^{J} \leq \|\tilde{s}\|_{L_p(M)} \leq c_2 q^{d/p} \|a\|_{\ell_p(\Xi)} \sum_{J=0}^{d} (1 - Ch^{K \nu - 2m - d})^{J}.
\]
Figure 1: Maximum latitudinal values of the Lagrange function for the kernel $k_2(x, \alpha) = (1 - x \cdot \alpha) \log(1 - x \cdot \alpha)$. This experiment was carried out in double precision arithmetic and the plateau at roughly $10^{-11}$ occurs due to ill conditioning of the collocation matrices and truncation error.

Example 1. We demonstrate the decay of Lagrange functions and their coefficients for the second order restricted surface spline (also known as the thin plate spline) $k_2(x, \alpha) = (1 - x \cdot \alpha) \log(1 - x \cdot \alpha)$. The interpolant takes the form $\chi_\xi = \sum_{\zeta \in \Xi} A_{\zeta, \xi} k(\cdot, \zeta) + p_\xi$, where $p_\xi$ is a degree 1 spherical harmonic. In this example, we use the “minimal energy points” of Womersley for the sphere – these are described and distributed at the website [36]. The value of these point sets is as benchmarks. Each set of centers has a nearly identical mesh ratio. Furthermore, the important geometric properties (e.g., fill distance and separation distance) are explicitly documented. Their potential theoretic properties and their importance in constructing quadrature rules and spherical designs, which are discussed in [32, 33], are not pertinent to this work. Because of the nice geometric properties of the minimal energy point sets, it is sufficient to consider the Lagrange function $\chi_\xi$ centered at the north pole $\xi = (0, 0, 1)$.

Figure 1 displays the maximal latitudinal values $\log_{10} |\chi_\xi|$ of $\log_{10} |\chi_\xi|$. We clearly observe the

---

4The function $\chi_\xi$ is evaluated on a set of points $(\phi, \theta)$ with $n_0$ equispaced latitudes $\phi \in [0, \pi]$ and $n_1$ equispaced longitudes $\theta \in [0, 2\pi]$. 

18
exponential decay of the Lagrange function\(^5\)

\[
|\chi_{\xi}(x)| \leq C_L \exp \left( -\nu_L \frac{d(x, \xi)}{h} \right)
\]
guaranteed by [19, Theorem 5.3]. From this figure, the value of \(\nu_L\), which measures the rate of exponential decay is observed to be close to 1.35.

We can visualize the decay of the corresponding coefficients in the same way. We again take the Lagrange function centered at the north pole: for each \(\zeta' \in \Xi\), the coefficient \(|A_{\zeta',\xi}|\) of the kernel \(k(\cdot, \zeta')\) in the expansion \(\chi_{\xi} = \sum A_{\zeta',\xi} k(\cdot, \zeta') + p_{\xi}\) is plotted with horizontal coordinate \(\sin(\zeta')\). The results for sets of centers of size \(N = 900, 2500\) and \(10000\) are given in Figure 2. The exponential decay

\[
|A_{\zeta',\xi}| \leq C_c q^{d-2m} \exp \left( -\nu_c \frac{\text{dist}(\xi, \zeta')}{h} \right).
\]
guaranteed by Theorem 4.3 is clearly in force, and we can estimate the constants \(\nu_c\) and \(C_c\) for the decay of the coefficients following the method used for the Lagrange functions themselves – we note that coefficients are shifted vertically, which is a consequence of the factor of \(q^{d-2m} = q^{-2}\) in the estimate (4.10). Table 1 gives more results with some added detail, including estimates of the constants \(C_L\) and \(C_c\).

\footnote{At least until a terminal value of roughly \(10^{-11}\), at which point there is a plateau beyond which the values no longer decay – see below and Figure 3 for an explanation of this.}

Figure 2: Plot of coefficients for a Lagrange function in the kernel space \(S(k_2, \Xi)\). This experiment was carried out in double precision arithmetic.
Table 1: Estimates of fill distance $h$, mesh ratio $\rho$ for some minimum energy point sets on the sphere and $\nu$ and $C$ values for the kernel $k_2(x,\alpha) = (1 - x \cdot \alpha) \log(1 - x \cdot \alpha)$.

| $N$  | $h_X$  | $\rho_X$ | $\nu_L$  | $C_L$  | $\nu_c$ | $C_c$  |
|------|--------|----------|----------|--------|---------|--------|
| 400  | 0.1136 | 1.2930   | 1.1119   | 0.8382 | 1.0997  | 69.9891|
| 900  | 0.0874 | 1.5302   | 1.3566   | 1.0982 | 1.3445  | 231.5573|
| 1600 | 0.0656 | 1.5333   | 1.3513   | 1.2170 | 1.3216  | 324.8534|
| 2500 | 0.0522 | 1.5278   | 1.3345   | 0.9618 | 1.3117  | 470.6483|
| 5041 | 0.0365 | 1.5304   | 1.3395   | 1.1080 | 1.3158  | 1087.8  |
| 10000| 0.0260 | 1.5421   | 1.3645   | 1.1934 | 1.3369  | 2564.9  |

The perceived plateau present in the Lagrange function values as well as the coefficients shown in Figures 1 and 2 is due purely to round-off error related to the conditioning of kernel collocation and evaluation matrices. These results were produced using double-precision (approximately 16 digits) floating point arithmetic. To illustrate this point, we plot the decay rate of the Lagrange coefficients for the 900 and 1600 point node sets as computed using high-precision (40 digits) floating point arithmetic in Figure 3. The figure clearly shows the exponential decay does not plateau and continues as the theory predicts.

Example 2. In this example we construct a basis $(\tilde{\chi}_\xi)_{\xi \in \Xi}$ for the kernel space $S(k_2, \Xi) \subset C(S^2)$ by using $O((\log N)^2)$ centers to construct each $\tilde{\chi}_\xi$.

With this basis, we use an equivalent representation in the form

$$I_\Xi f = \sum_{\xi \in \Xi} c_\xi \tilde{\chi}_\xi(\cdot),$$

(6.1)

where each $\tilde{\chi}_\xi$ is a local Lagrange function about the node $\xi$ formed by $M \ll N$ basis elements of $S(k_2, \Xi)$. Specifically, let $Y(\xi) \subset \Xi$ such that $\xi \in Y(\xi)$, $\#Y(\xi) = M$, and $\bigcup_{\xi \in \Xi} Y(\xi) = \Xi$, then

$$\tilde{\chi}_\xi = \sum_{\zeta \in Y(\xi)} a_{\xi,\zeta} k(\cdot, \zeta) + \sum_{j=1}^4 b_{\xi,j} \phi_j.$$  

(6.2)

The coefficients $a_{\xi,\zeta}$ and $b_{\xi,j}$ are determined from the conditions

$$\tilde{\chi}_\xi(\zeta) = \begin{cases} 1 & \text{if } \zeta = \xi, \\ 0 & \text{if } \zeta \in Y(\xi) \setminus \xi, \quad \text{and} \quad \sum_{\zeta \in Y(\xi)} a_{\xi,\zeta} \varphi_j(\zeta) = 0. \end{cases}$$

The linear system for determining the interpolation coefficients $c_\xi$ in (6.1) can be written as:

$$[K_\Xi \Phi] \begin{bmatrix} A_T \\ B_T \end{bmatrix} [c] = [f],$$

(6.3)

where $(K_\Xi)_{i,j} = k_2(\xi_i, \xi_j)$ and $B_{i,j} = \phi_j(\xi_i)$, $i, j = 1, \ldots, N$. The matrix $A_T$ is a $N$-by-$N$ sparse matrix where each column contains $M$ entries corresponding to the values of $a_{\xi,\zeta}$ in
Figure 3: Plot of coefficients for a Lagrange function in the kernel space $S(k_2, \Xi)$. This experiment was carried out in Maple with 40 digit arithmetic.

The matrix $B_T$ is a 4-by-$N$ matrix with each column containing the values of $b_{\xi,j}$ in (6.2). With the linear system written in this way, one can view the matrix $[A_T \ B_T]^T$ as a right preconditioner for the standard kernel interpolation matrix.

If the sets $\Upsilon(\xi)$ are chosen appropriately then the linear system (6.3) should be “numerically nice” in the sense that the matrix $K_\Xi A_T + \Phi B_T$ should have decaying elements from its diagonal and should be well conditioned. In the example below, each $\Upsilon(\xi)$ is chosen as the $M$ nearest nodes to $\xi$. Section 5 suggests taking $M = O((\log N)^2)$. Through trial and error we found that choosing $M = 7\lceil(\log_{10} N)^2 \rceil$ gave very good results over several decades of $N$. Each set $\Upsilon(\xi)$ can be determined $O(\log N)$ operations by using a KD-tree algorithm for sorting and searching through the nodes $\Xi$. The cost for constructing the KD-tree is $O(N(\log N)^2)$. Thus, constructing all the sets $\Upsilon$ takes $O(N(\log N)^2)$ operations.

To solve this linear system we will use the generalized minimum residual method (GMRES) [31]. This is a Krylov subspace method which is applicable to non-symmetric linear systems and only requires computing matrix-vector products. Ideally, there should be a method for computing these matrix vector products in $O(N)$ or $O(N \log N)$ operations to make GMRES more efficient. Keiner et. al. have shown that this can be done in the case of the kernel matrix $K_\Xi$ using fast algorithms for spherical Fourier transforms [22]. In the results that follow, we have not used this algorithm, but have instead just computed the matrix vector products directly. We will investigate the use of these fast algorithms in a follow up study.
For the numerical tests we use icosahedral node sets $\Xi \subset S^2$ of increasing cardinality. These were chosen because of their popularity in atmospheric fluid dynamics (see, for example, [16, 34, 30, 25]) where interpolation between node sets is often required. The values of $f$ were chosen to take on random values from a uniform distribution between $[-1, 1]$. Table 2 displays the number of GMRES iterations to compute an approximate solution to the resulting linear systems (6.3). As we can see, the number of iterations stays relatively constant as $N$ increases and does not appear to increase with $N$.

| $N$  | $m$  | $\text{Number GMRES iterations}$ |
|------|------|----------------------------------|
| 2562 | 84   | 7 $\text{tol} = 10^{-6}$       |
|      |      | 5 $\text{tol} = 10^{-8}$       |
| 10242| 119  | 5                                 |
|      |      | 7                                 |
| 23042| 140  | 6                                 |
|      |      | 7                                 |
| 40962| 154  | 5                                 |
|      |      | 7                                 |
| 92162| 175  | 6                                 |
|      |      | 8                                 |
| 163842| 196 | 5                                 |
|      |      | 7                                 |

Table 2: Number of GMRES iterations required for computing an approximate solution to (6.3) using icosahedral node sets of cardinality $N$. $m$ corresponds to the number of nodes used to construct the local basis and $\text{tol}$ refers to the tolerance on the relative residual in the GMRES method. The right hand side was set to random values uniformly distributed between $[-1, 1]$ and the initial guess for GMRES was set equal to the function values.

**Example 3.** A second example shows similar results for Lagrange functions for the kernel $k(x, \alpha) = (x - \alpha)^2 \log(x - \alpha)$ restricted to a torus of outer radius 4 and inner radius 2. In other words, the surface parametrized by

$$
\begin{align*}
  x &= (3 + \cos v) \cos u \\
  y &= (3 + \cos v) \sin u \\
  z &= \sin v,
\end{align*}
$$

with $u, v \in [0, 2\pi]$. This combination of kernel and manifold is not treated in [19] (the curved torus is not even a symmetric space) although it is considered in [15], as a subset of $\mathbb{R}^3$ that is $\Pi_1$ unisolvent. Indeed the torus is the zero set of a degree 4 polynomial in $\mathbb{R}^d$, a sufficiently dense subset will also be $\Pi_1$ unisolvent. Hence, interpolation on such sets by shifts of $k$ is well posed.

For this experiment, we consider “minimum energy” point sets $\Xi$ produced by Ayla Gafni, Doug Hardin and Ed Saff which have previously been used in [15]. In each case we fix the point $\xi = (4, 0, 0)$ and we examine coefficients $(A_{\xi, \zeta})_{\zeta \in \Xi}$ of the Lagrange function $\chi_\xi = \sum_{\zeta \in \Xi} A_{\xi, \zeta} k(\cdot, \zeta) + p$. To solve the interpolation problem, we add a linear polynomial, $p \in \Pi_1(\mathbb{R}^3)$, and require the coefficients to satisfy the side conditions $\sum_{\zeta \in \Xi} A_{\xi, \zeta} q(\zeta) = 0, \forall q \in \Pi_1(\mathbb{R}^3)$ – in other words, the coefficients annihilate linear polynomials.

Despite the fact that Theorem 4.3 does not apply, a certain exponential decay is observed
for these coefficients as well. As before, the rate of decay seems to be independent of $h$ as well as $N = \#\Xi$.

We provide two ways to visualize the decay of the coefficients: first by arranging them latitudinally with horizontal axis representing the distance from $\xi$ in the $u$ direction and then longitudinally with horizontal axis representing the distance from $\xi$ in the $u$ direction. For the specific choice of $\xi = (4, 0, 0)$, both correspond to geodesic distances. (Other choices of $\xi$ are observed to have the same rate of coefficient decay.)

![Figure 4: Log plot of coefficients (in absolute value) of Lagrange functions centered at (4,0,0) for the kernel $k(x, \alpha) = |x-\alpha|^2 \log |x-\alpha|$ using minimal energy sets on the torus of cardinality on 500, 1000 and 4000.](image)

A more complete account of this experiment is given in Table 3, where estimates of the exponential decay rates $\nu_{\text{Lat}}$ and $\nu_{\text{Long}}$ (in the $v$ and $u$ directions, respectively) are given. We note that they are between 1 and 1.3 in either direction.

| $N$  | $h_X$  | $\rho_X$ | $\nu_{\text{Lat}}$ | $C_{\text{Lat}}$ | $\nu_{\text{Long}}$ | $C_{\text{Long}}$ |
|------|--------|----------|---------------------|------------------|---------------------|------------------|
| 500  | 0.3383 | 1.5242   | 1.2312              | 6.9946           | 1.1504              | 7.3231           |
| 750  | 0.2737 | 1.5024   | 1.1556              | 9.4737           | 1.2122              | 15.4474          |
| 1000 | 0.2375 | 1.5014   | 1.2870              | 15.6376          | 1.2401              | 20.4220          |
| 1999 | 0.1639 | 1.4725   | 1.1213              | 25.3917          | 1.2719              | 49.0776          |
| 3000 | 0.1333 | 1.4479   | 1.0793              | 36.1593          | 1.2421              | 58.7687          |
| 4000 | 0.1151 | 1.4498   | 1.1836              | 57.4460          | 1.2738              | 105.2720         |

Table 3: Results for a Lagrange function experiment on the torus using the thin plate spline kernel $k(x, \alpha) = |x-\alpha|^2 \log |x-\alpha|$. In addition to fill distance $h$, mesh ratio $\rho$ for minimum energy set of Saff and Hardin, estimates of the latitudinal and longitudinal exponential decay rate and constant $\nu$ and $C$ values are given.
References

[1] T. Aubin, *Nonlinear analysis on manifolds. Monge-Ampère equations*, vol. 252 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, New York, 1982.

[2] B. J. C. Baxter and S. Hubbert, *Radial basis functions for the sphere*, in Recent progress in multivariate approximation (Witten-Bommerholz, 2000), vol. 137 of Internat. Ser. Numer. Math., Birkhäuser, Basel, 2001, pp. 33–47.

[3] R. Beatson and L. Greengard, *A short course on fast multipole methods*, in Wavelets, multilevel methods and elliptic PDEs (Leicester, 1996), Numer. Math. Sci. Comput., Oxford Univ. Press, New York, 1997, pp. 1–37.

[4] R. K. Beatson, J. B. Cherrie, and C. T. Mouat, *Fast fitting of radial basis functions: methods based on preconditioned GMRES iteration*, Adv. Comput. Math., 11 (1999), pp. 253–270. Radial basis functions and their applications.

[5] R. K. Beatson, G. Goodsell, and M. J. D. Powell, *On multigrid techniques for thin plate spline interpolation in two dimensions*, in The mathematics of numerical analysis (Park City, UT, 1995), vol. 32 of Lectures in Appl. Math., Amer. Math. Soc., Providence, RI, 1996, pp. 77–97.

[6] R. K. Beatson and W. A. Light, *Fast evaluation of radial basis functions: methods for two-dimensional polyharmonic splines*, IMA J. Numer. Anal., 17 (1997), pp. 343–372.

[7] R. K. Beatson and M. J. D. Powell, *An iterative method for thin plate spline interpolation that employs approximations to Lagrange functions*, in Numerical analysis 1993 (Dundee, 1993), vol. 303 of Pitman Res. Notes Math. Ser., Longman Sci. Tech., Harlow, 1994, pp. 17–39.

[8] R. K. Beatson, M. J. D. Powell, and A. M. Tan, *Fast evaluation of polyharmonic splines in three dimensions*, IMA J. Numer. Anal., 27 (2007), pp. 427–450.

[9] M. D. Buhmann, *Multivariate cardinal interpolation with radial-basis functions*, Constr. Approx., 6 (1990), pp. 225–255.

[10] N. Dyn, F. J. Narcowich, and J. D. Ward, *Variational principles and Sobolev-type estimates for generalized interpolation on a Riemannian manifold*, Constr. Approx., 15 (1999), pp. 175–208.

[11] G. E. Fasshauer, *Meshfree approximation methods with MATLAB*, vol. 6 of Interdisciplinary Mathematical Sciences, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007. With 1 CD-ROM (Windows, Macintosh and UNIX).

[12] A. C. Faul, G. Goodsell, and M. J. D. Powell, *A Krylov subspace algorithm for multiquadric interpolation in many dimensions*, IMA J. Numer. Anal., 25 (2005), pp. 1–24.
[13] A. C. Faul and M. J. D. Powell, Proof of convergence of an iterative technique for thin plate spline interpolation in two dimensions, Adv. Comput. Math., 11 (1999), pp. 183–192. Radial basis functions and their applications.

[14] B. Fornberg, N. Flyer, S. Hovde, and C. Piret, Locality properties of radial basis function expansion coefficients for equispaced interpolation, IMA J. Numer. Anal., 28 (2008), pp. 121–142.

[15] E. Fuselier and G. Wright, Scattered data interpolation on embedded submanifolds with restricted positive definite kernels: Sobolev error estimates. arXiv:1007.2825.

[16] F. X. Giraldo, Lagrange-Galerkin methods on spherical geodesic grids, J. Comput. Phys., 136 (1997), pp. 197–213.

[17] T. Hangelbroek, Polyharmonic approximation on spheres, Constr. Approx., 33 (2011), pp. 77–92.

[18] T. Hangelbroek, F. J. Narcowich, X. Sun, and J. D. Ward, Kernel approximation on manifolds II: The $L_\infty$ norm of the $L_2$ projector, SIAM Journal on Mathematical Analysis, 43 (2011), pp. 662–684.

[19] T. Hangelbroek, F. J. Narcowich, and J. D. Ward, Polyharmonic and related kernels on manifolds: Interpolation and approximation. arXiv:1012.4852.

[20] ———, Kernel approximation on manifolds I: Bounding the Lebesgue constant, SIAM Journal on Mathematical Analysis, 42 (2010), pp. 1732–1760.

[21] T. Hangelbroek and D. Schmid, Surface spline approximation on SO(3), Applied and Computational Harmonic Analysis, 31 (2011), pp. 169 – 184.

[22] J. Keiner, S. Kunis, and D. Potts, Fast summation of radial functions on the sphere, Computing, 78 (2006), pp. 1–15.

[23] L. Ling and E. J. Kansa, A least-squares preconditioner for radial basis functions collocation methods, Adv. Comput. Math., 23 (2005), pp. 31–54.

[24] W. R. Madych and S. A. Nelson, Polyharmonic cardinal splines, J. Approx. Theory, 60 (1990), pp. 141–156.

[25] D. Majewski, D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner, The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests, Mon. Wea. Rev., 130 (2002), pp. 319–338.

[26] H. N. Mhaskar, F. J. Narcowich, J. Prestin, and J. D. Ward, $L^p$ Bernstein estimates and approximation by spherical basis functions, Math. Comp., 79 (2010), pp. 1647–1679.
[27] C. Rabut, *Elementary m-harmonic cardinal B-splines*, Numer. Algorithms, 2 (1992), pp. 39–61.

[28] ———, *High level m-harmonic cardinal B-splines*, Numer. Algorithms, 2 (1992), pp. 63–84.

[29] S. D. Riemenschneider and N. Sivakumar, *On cardinal interpolation by Gaussian radial-basis functions: properties of fundamental functions and estimates for Lebesgue constants*, J. Anal. Math., 79 (1999), pp. 33–61.

[30] T. D. Ringler, R. P. Heikes, and D. A. Randall, *Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores*, Mon. Wea. Rev., 128 (2000), pp. 2471–2490.

[31] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Comput., 7 (1986), pp. 856–869.

[32] E. B. Saff and A. B. J. Kuijlaars, *Distributing many points on a sphere*, Math. Intelligencer, 19 (1997), pp. 5–11.

[33] I. H. Sloan and R. S. Womersley, *Extremal systems of points and numerical integration on the sphere*, Adv. Comput. Math., 21 (2004), pp. 107–125.

[34] G. R. Stuhne and W. R. Peltier, *New icosahedral grid-point discretizations of the shallow water equations on the sphere*, J. Comput. Phys., 148 (1999), pp. 23–53.

[35] T. Tran, Q. T. Le Gia, I. H. Sloan, and E. P. Stephan, *Preconditioners for pseudodifferential equations on the sphere with radial basis functions*, Numer. Math., 115 (2010), pp. 141–163.

[36] R. Womersley, *Minimum energy points on the sphere $S^2$*, http://web.maths.unsw.edu.au/~rsw/Sphere/Energy/index.html, 2003.