Surface spline approximation on *SO*(3)

Thomas Hangelbroek [a,1], Dominik Schmid [b,*,2]

[a] *Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA*
[b] *Institute of Biomathematics and Biometry, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstraße 1,*
D-85764 Neuherberg, Germany

A R T I C L E   I N F O

A B S T R A C T

The purpose of this article is to introduce a new class of kernels on *SO*(3) for approximation and interpolation, and to estimate the approximation power of the associated spaces. The kernels we consider arise as linear combinations of Green's functions of certain differential operators on the rotation group. They are conditionally positive definite and have a simple, closed-form expression, lending themselves to direct implementation via, e.g., interpolation or least-squares approximation. To gauge the approximation power of the underlying spaces, we introduce an approximation scheme providing precise L_p error estimates for linear schemes, namely with L_p approximation order conforming to the L_p smoothness of the target function.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Approximating a function $f : \mathbb{X} \to \mathbb{C}$ on various structures $\mathbb{X}$ by linear combinations of translates of a given single basis function is a widely used method. At the heart of this methodology is the creation of an approximant using a fixed kernel $\kappa : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$, by taking a scattered linear combination of copies of this kernel: $s_\Xi(\mathbf{x}) = \sum_{\xi \in \Xi} A_\xi \kappa(\mathbf{x}, \xi)$. The success of these methods derives from their ability to generate approximants from data having arbitrary geometry. The case where the underlying structure are Euclidean spaces $\mathbb{R}^d$ or Euclidean spheres $\mathbb{S}^d$ has been studied in great detail over the last decade, see [1,2]. Such kernel methods have found success in approximation theory (for treating high dimensional scattered data) and learning theory (in kernel based learning, support vector machines, neural networks, etc.).

However, in various applications we are confronted with the situation that the underlying set is a compact or locally compact group and one is asked to propose suitable approximation procedures on these structures (this is a setting with no natural dilation operator and no 'regular grids'). Such problems arise in biochemistry, crystallography and robotics to name only a few. The monograph [3] provides a great collection of scattered data approximation problems where different matrix groups are involved.

In view of applications, the rotation group *SO*(3) is, without doubt, one of the most important groups. There exist a wealth of scattered data approximation problems on *SO*(3), cf. [4–7]. Such problems have attracted significant recent

---

* Corresponding author.
E-mail addresses: thomas.c.hangelbroek@vanderbilt.edu (T. Hangelbroek), dominik.schmid@helmholtz-muenchen.de (D. Schmid).

attention from the mathematical community, which has investigated approximation via Wigner functions [8–12], and kernels [13,14].

The goal of this article is to introduce a family of kernels on the rotation group suitable for approximation and to investigate the approximation power of this family by estimating the performance of a (theoretical) approximation scheme. The kernels we present are notable in many respects:

- they are *conditionally positive definite*, have a simple, closed-form expression, and are consequently well-suited for implementation (for, say, interpolation or least-squares approximation);
- they invert iterated, perturbed Laplace–Beltrami operators, tying them directly to successful kernel based approximation in other settings (cf. *surface spline* approximation on $\mathbb{R}^d$ [15,16], restricted surface splines on spheres [17], and, of course, polynomial splines in one dimension);
- they have excellent scaling properties, much like the surface splines in the Euclidean case, suggesting that they can adjust themselves to data whose density varies spatially, as in [18].

In addition to these three points, we study the approximation power of the spaces generated by the kernels. We develop an approximation scheme and accompanying error estimates derived from the kernel's role as a Green's function. To be sure, the scheme we develop is theoretical and does not immediately lend itself to direct implementation, but the accompanying error estimates are of a type that is notoriously difficult for kernel based approximation when the kernel is not dilated. This is true even in the Euclidean case, as in radial basis function (RBF) approximation. The results we are after give precise error rates for functions at all appropriate levels of smoothness; we determine L_p rates of convergence commensurate with the L_p smoothness of the function being approximated. We note that the results we seek are well known for kernel approximation where dilations and regular grids are employed [19,20], and even without grids [21,22]. However, in the absence of dilation, error estimates are often known only for target functions residing in a reproducing kernel Hilbert space.

We have organized the article as follows. In the next section we give the necessary results of analysis on the rotation group. In Section 3 we introduce the kernels and give their expansions in terms of Wigner D-functions, the basic elements of Fourier analysis on the rotation group. From this we are able to show that the kernels are conditionally positive definite, and, hence are capable of solving interpolation and other scattered data fitting problems. The expansion also allows us to identify the integral identities these kernels satisfy. Section 4 establishes the basic strategy to localize a kernel: by replacing a given kernel by a linear combination of its nearby copies. A precise estimate of the error in making this exchange is then given. In Section 5 we propose an approximation scheme and present error estimates that demonstrate the approximation power of the kernels. Finally, Section 6 involves a discussion of potential generalizations of this method to other groups and homogeneous spaces. Recent work on bounding Lebesgue constants for kernel interpolation is discussed, which indicates that the theoretical error estimates may also hold for more practical schemes like interpolation and least-squares approximation.

2. Background

In this section we collect some basic material and necessary notation on the rotation group to keep the paper self-contained.

2.1. Basic facts

Let $SO(3) := \{\boldsymbol{x} \in \mathbb{R}^{3\times 3}:\ \boldsymbol{x}^T\boldsymbol{x} = \boldsymbol{e},\ \det \boldsymbol{x} = 1\}$ denote the non-Abelian compact group of proper rotations in the Euclidean space $\mathbb{R}^3$ and let μ denote the normalized Haar measure on $SO(3)$, i.e., we have $\mu(SO(3)) = 1$. There are various ways to parameterize this group. An important parametric form for the rotation group uses the so-called Euler angles $(\varphi_1, \theta, \varphi_2) \in [0, 2\pi) \times [0, \pi] \times [0, 2\pi)$. In the literature there are miscellaneous conventions of Euler angles. In this paper we follow the conventions made in [23, Section 1.2]. There the Euler angles are defined such that we can write every $\boldsymbol{x} \in SO(3)$ as

$$\boldsymbol{x} = \boldsymbol{x}(\varphi_1, \theta, \varphi_2) = \boldsymbol{s}_z[\varphi_1]\boldsymbol{s}_x[\theta]\boldsymbol{s}_z[\varphi_2], \tag{1}$$

where

$$\boldsymbol{s}_z[t] = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{s}_x[t] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}$$

are rotations about the z-axis and the x-axis, respectively.

Any integral over $SO(3)$ (or portions thereof) will always be assumed to be the Haar integral. Using Euler angles, the Haar integral of a function f on $SO(3)$ reads as

$$\int\limits_{SO(3)} f(\boldsymbol{x})\, \mathrm{d}\mu(\boldsymbol{x}) = \frac{1}{8\pi^2} \int\limits_0^{2\pi} \int\limits_0^{\pi} \int\limits_0^{2\pi} f(\varphi_1, \theta, \varphi_2) \sin\theta\, \mathrm{d}\varphi_1\, \mathrm{d}\theta\, \mathrm{d}\varphi_2. \tag{2}$$

Besides the well-known parameterization via Euler angles, the parameterization via the three-dimensional projective space is also of some importance to us. Let $\mathcal{K}_\pi$ be the closed ball in $\mathbb{R}^3$ of radius π centered at the origin and identify antipodal points. This is the three-dimensional projective space. An element $\boldsymbol{x} \in SO(3)$ is identified with a point in the projective space $\mathcal{K}_\pi$ by $\boldsymbol{x} \to \omega \mathbf{r}$ where $\mathbf{r}$, satisfying $\boldsymbol{x}\mathbf{r} = \mathbf{r}$ and $\|\mathbf{r}\| = 1$, is the *rotation axis* and ω, which can be chosen in $[0, \pi]$, is the *rotation angle* of $\boldsymbol{x}$.

Then it is easy to see that

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) := \omega\big(\boldsymbol{y}^{-1}\boldsymbol{x}\big) \tag{3}$$

defines a translation invariant metric on $SO(3)$. Of course, this is not the only metric one can place on $SO(3)$, see [24] for a discussion of other metrics and sampling in $SO(3)$, but it is compatible with the Riemannian structure of the group and we will use this metric throughout this article. This leads to the usual definition of the ball centered at $\boldsymbol{\alpha}$ having radius ρ: $B(\boldsymbol{\alpha}, \rho) = \{\boldsymbol{\zeta} \mid \text{dist}(\boldsymbol{\zeta}, \boldsymbol{\alpha}) < \rho\}$. A consequence of (2) is that we can estimate the volume of a ball $\mu(B(\boldsymbol{\alpha}, \rho)) = \int_{SO(3)} \mathbb{1}_{B(\boldsymbol{\alpha}, \rho)}(\boldsymbol{x}) \, d\mu(\boldsymbol{x})$, obtaining constants $b_1 := 2/(3\pi^3)$ and $b_2 := 1/(6\pi)$ so that for any $0 < \rho < \pi$ and any $\boldsymbol{\alpha} \in SO(3)$,

$$b_1 \rho^3 \leqslant \mu\big(B(\boldsymbol{\alpha}, \rho)\big) \leqslant b_2 \rho^3. \tag{4}$$

We utilize the notation ω in the context of the rotation angle and dist in the context of the metric. We can use this metric in order to quantify the distribution of the centers in the set

$$\boldsymbol{\Xi} := \big\{\boldsymbol{\xi}_j \in SO(3) \colon j = 0, \dots, M-1\big\}.$$

To do so we define two parameters. The first one is the separation distance

$$q = q(\boldsymbol{\Xi}) := \min_{0 \leqslant i < j \leqslant M-1} \text{dist}(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j)$$

which measures 'clustering trends' in $\boldsymbol{\Xi} \subset SO(3)$. On the other hand, the fill distance

$$h = h(\boldsymbol{\Xi}) := \max_{\boldsymbol{y} \in SO(3)} \min_{j=0, \dots, M-1} \text{dist}(\boldsymbol{\xi}_j, \boldsymbol{y})$$

describes the density of $\boldsymbol{\Xi}$ in $SO(3)$.

2.2. Harmonic analysis and kernels

The basic building blocks in harmonic analysis on topological groups are the irreducible unitary representations. On $SO(3)$, it is possible to compute a complete set of irreducible unitary representations, explicitly. More precisely, using the quasi-left regular representation of $SO(3)$ on $L_2(\mathbb{S}^2)$ one is able to compute for each $\ell \in \mathbb{N}_0$ a $(2\ell + 1) \times (2\ell + 1)$-matrix D^ℓ such that these operators $D^\ell, \ell \in \mathbb{N}_0$, form a complete set of irreducible unitary representations of $SO(3)$. The matrix elements $D^\ell_{k,m}$ in the canonical basis are often called Wigner D-functions of degree ℓ and orders k and m. In terms of Euler angles these functions are given by

$$D^\ell_{k,m}(\boldsymbol{x}) = D^\ell_{k,m}(\varphi_1, \theta, \varphi_2) = e^{-ik\varphi_1} P^\ell_{k,m}(\cos\theta) e^{-im\varphi_2}, \quad -\ell \leqslant k, m \leqslant \ell, \tag{5}$$

where the function $P^\ell_{k,m}$ is given by

$$P^\ell_{k,m}(t) = C(1-t)^{-\frac{m-k}{2}}(1+t)^{-\frac{m+k}{2}} \frac{d^{\ell-m}}{dt^{\ell-m}}\big((1-t)^{\ell-k}(1+t)^{\ell+k}\big)$$

with $C = \frac{(-1)^{\ell-k} i^{m-k}}{2^\ell (\ell-k)!} \sqrt{\frac{(\ell-k)!(\ell+m)!}{(\ell+k)!(\ell-m)!}}$.

We refer to [23] for all the details in these computations. By means of the Peter–Weyl theorem the Wigner D-functions, the matrix elements of the irreducible representations, constitute an orthogonal basis of the $L_2(SO(3))$. Hence, every $f \in L_2(SO(3))$ can be expanded in an $SO(3)$ Fourier series

$$f = \sum_{\ell \in \mathbb{N}_0} \sum_{k,m=-\ell}^{\ell} \sqrt{2\ell+1}\, \hat{f}^\ell_{k,m} D^\ell_{k,m}$$

with $SO(3)$ Fourier coefficients $\hat{f}^\ell_{k,m} = \sqrt{2\ell+1} \int_{SO(3)} f(\boldsymbol{x}) \overline{D^\ell_{k,m}(\boldsymbol{x})} \, d\mu(\boldsymbol{x})$. On the other hand, to each eigenvalue, $\nu_\ell := \ell(\ell+1)$, of the Laplace–Beltrami operator Δ on $SO(3)$, there corresponds an eigenspace of dimension $(2\ell+1)^2$ with orthogonal basis $(D^\ell_{k,m})_{k,m}$, i.e., we have for all $\ell \in \mathbb{N}_0$ and all $k, m = -\ell, \dots, \ell$ that

$$\Delta D^\ell_{k,m} = \ell(\ell+1) D^\ell_{k,m}. \tag{6}$$

We denote this eigenspace by $\mathcal{H}_\ell$, and we refer to $\boldsymbol{\Pi}_n = \bigoplus_{\ell=0}^n \mathcal{H}_\ell$ as the polynomials of degree no more than n.

The kernels we study in this article possess a certain symmetry; they are functions of the distance between the two arguments, $\text{dist}(\boldsymbol{x}, \boldsymbol{\alpha}) = \omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x}) \in [0, \pi]$ only. By the bijectivity of $\sin(\frac{\cdot}{2})$ and $\cos(\frac{\cdot}{2})$ on $[0, \pi]$ we are entitled to express these kernels as $k(\boldsymbol{x}, \boldsymbol{\alpha}) = \phi(\sin(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2})) = \psi(\cos(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2}))$, where $\phi, \psi : [0, 1] \to \mathbb{R}$. There are benefits to both alternatives, as we shall see in the following sections.

On $SO(3)$ one easily checks that the functions that only depend on the rotational angle of the argument coincide exactly with the class functions – functions that are constant on conjugacy classes, i.e. $f(\boldsymbol{x}) = f(\boldsymbol{y}\boldsymbol{x}\boldsymbol{y}^{-1})$ for all $\boldsymbol{x}, \boldsymbol{y} \in SO(3)$. In other words, for any class function f on $SO(3)$ there is a uniquely determined $\tilde{f} : [0, \pi] \to \mathbb{C}$ such that $f(\boldsymbol{x}) = \tilde{f}(\omega(\boldsymbol{x}))$ for all $\boldsymbol{x} \in SO(3)$. The Haar integral for such a class function reads as

$$\int\limits_{SO(3)} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) = \int\limits_{SO(3)} \tilde{f}(\omega(\boldsymbol{x})) \, \mathrm{d}\mu(\boldsymbol{x}) = \frac{2}{\pi} \int\limits_0^\pi \tilde{f}(t) \sin^2\left(\frac{t}{2}\right) \mathrm{d}t. \tag{7}$$

We make the convention that if f is a class function on $SO(3)$ and if no confusion occurs, we subsequently drop the tilde.

Since the characters $\mathfrak{c}_\ell = \text{Tr} \, D^\ell$, $\ell \in \mathbb{N}_0$, of $SO(3)$ form an orthonormal basis for the space of class functions in $L_2(SO(3))$, we can decompose any continuous kernel possessing the symmetry mentioned above as $k(\boldsymbol{x}, \boldsymbol{\alpha}) = \sum_{\ell=0}^\infty \tilde{k}(\ell) \mathfrak{c}_\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x})$. Using the formula $\mathfrak{c}_\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x}) = \mathcal{U}_{2\ell}(\cos(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2}))$, such kernels can be expressed as an expansion of even powered Chebyshev polynomials of the second kind,

$$k(\boldsymbol{x}, \boldsymbol{\alpha}) = \sum_{\ell=0}^\infty \tilde{k}(\ell) \mathfrak{c}_\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x}) = \sum_{\ell=0}^\infty \tilde{k}(\ell) \mathcal{U}_{2\ell}\left(\cos\left(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2}\right)\right). \tag{8}$$

Alternatively, we can write:

$$k(\boldsymbol{x}, \boldsymbol{\alpha}) = \sum_{\ell=0}^\infty \tilde{k}(\ell) \mathfrak{c}_\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x}) = \sum_{\ell=0}^\infty \tilde{k}(\ell) \text{Tr}\left(D^\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x})\right)$$

$$= \sum_{\ell=0}^\infty \tilde{k}(\ell) \text{Tr}\left(D^\ell(\boldsymbol{\alpha})^* D^\ell(\boldsymbol{x})\right) = \sum_{\ell=0}^\infty \sum_{\iota=-\ell}^\ell \sum_{\nu=-\ell}^\ell \tilde{k}(\ell) D_{\iota,\nu}^\ell(\boldsymbol{x}) \overline{D_{\iota,\nu}^\ell(\boldsymbol{\alpha})}. \tag{9}$$

This is sometimes called the addition formula for Wigner D-functions.

The expression of the kernel as a series of Chebyshev polynomials (8) as well as the addition formula (9) are of prime importance in the study of kernels on $SO(3)$. The first is used to understand the kernel as the fundamental solution of a simple differential operator, which is the goal of the next section, while the second is key to localizing the kernel, which is tackled in Section 4.

2.3. Smoothness spaces

Because we are interested in approximating in the $L_p(SO(3))$ metric, we must introduce L_p smoothness spaces on $SO(3)$. This is done in two stages. In the first stage we introduce the (simpler) Sobolev and Hölder spaces, whose definitions are accessible, if not familiar to most readers. The second stage is postponed until Section 5, where we introduce Besov spaces.

Sobolev spaces on manifolds have a variety of equivalent characterizations. Perhaps the most straightforward of these is to import the definition from Euclidean space via a partition of unity and a corresponding set of diffeomorphisms. This is the approach taken in [25, 1.11]. The Sobolev space on a region $\mathcal{O}$ of $\mathbb{R}^3$ is defined to be the space of functions $f \in L_p(\mathcal{O})$ such that the quantity

$$\|f\|_{W_p^k(\mathcal{O})}^p := \sum_{\beta \leqslant k} \int\limits_{\mathcal{O}} |D^\beta f(x)|^p \, \mathrm{d}x \tag{10}$$

is finite. To transport this definition, let $(\psi_j)_{j=1}^N$ be a partition of unity of $SO(3)$, and let $(\Omega_j, h_j)_{j=1}^N$ be a corresponding collection of charts (each Ω_j is an open set in $SO(3)$ containing the closure of $\text{supp}(\psi_j)$ and each $h_j : \Omega_j \to \mathcal{O}_j \subset \mathbb{R}^3$ is a diffeomorphism).

Definition 1. For $1 \leqslant p < \infty$ the Sobolev space $W_p^k(SO(3))$ consists of functions $f \in L_p$ such that

$$\|f\|_{W_p^k}^p := \sum_{j=1}^N \left\|\left(\psi_j \circ (h_j^{-1})\right)\left(f \circ (h_j^{-1})\right)\right\|_{W_p^k(\mathcal{O}_j)}^p$$

is finite.

In this way we transport the Sobolev space definition from Euclidean space. For smoothness measured in L_∞, although the definition makes sense, we end up using the C^k spaces, whose norm is defined according to the same procedure:

$$\left\| \Delta^j f \right\|_p \lesssim \| f \|_{W_p^k}.$$

This follows from the well-known expression of Δ in local coordinates [25, 7.2.5]: namely $(\Delta f) \circ h = L(f \circ h)$ where

$$L = \frac{1}{\sqrt{\det g}} \sum_{j,k=1}^d \frac{\partial}{\partial x_j} \left(\sqrt{\det g} g^{jk} \frac{\partial}{\partial x_k} \right)$$

is a second order differential operator having smooth coefficients (here g is an invertible matrix having smooth entries and g^{jk} is the j, k entry of its inverse).

Another useful observation concerns the density of C^∞ functions in $W_p^k(SO(3))$ (and in $C^k(SO(3))$ when $p = \infty$), which follows from the density of such functions in $\mathbb{R}^3$. Each real valued function $(\psi_j f) \circ (h_j^{-1})$ can be approximated with error ε in W_p^k by a function $g_j : \mathcal{O}_j \to \mathbb{R}$ (via mollification, for instance). The desired smooth function is

$$G := \sum_{j=1}^N \tilde{\psi}_j \times G_j := \sum_{j=1}^N \tilde{\psi}_j \times (g_j \circ h_j).$$

Here we utilize a second family of smooth functions: each $\tilde{\psi}_j : SO(3) \to \mathbb{R}$ is a C^∞ cut-off function satisfying $\mathrm{supp}(\tilde{\psi}_j) \subset \Omega_j$ and $\tilde{\psi}_j(\boldsymbol{\alpha}) = 1$ for $\boldsymbol{\alpha} \in \mathrm{supp}(\psi_j)$ (the function $\tilde{\psi}_j \times G_j$ is extended by 0 outside of Ω_j). It follows that $G - f = \sum_{j=1}^N \tilde{\psi}_j(G_j - \psi_j f)$ and

$$\| G - f \|_{W_p^k(SO(3))} \leqslant \sum_{\ell=1}^N \sum_{j=1}^N \left\| \left[\psi_\ell \times \tilde{\psi}_j \times \left\{ g_j - (\psi_j \times f) \circ \left(h_j^{-1} \right) \right\} \circ h_j \right] \circ h_\ell^{-1} \right\|_{W_p^k(\mathcal{O}_\ell)}$$

$$\leqslant \sum_{\ell,j=1}^N C_{\ell,j} \left\| (\psi_\ell \times \tilde{\psi}_j) \circ h_\ell^{-1} \right\|_{C^k(\Omega_\ell)} \left\| g_j - (\psi_j \times f) \circ \left(h_j^{-1} \right) \right\|_{W_p^k(\mathcal{O}_j)}$$

where the last line follows from the fact that each $h_j \circ h_\ell^{-1}$ is a diffeomorphism from $h_\ell(\Omega_j \cap \Omega_\ell)$ to $h_j(\Omega_j \cap \Omega_\ell)$. Therefore, $\| G - f \|_{W_p^k(SO(3))} \leqslant C\varepsilon$, where C depends only on the two families of smooth functions $(\psi_j)_{j=1}^N$ and $(\tilde{\psi}_j)_{j=1}^N$, and on the diffeomorphisms h_j.

3. Surface splines on $SO(3)$

In this article we investigate approximation properties of the *surface splines* (perhaps more accurately the 'rotational surface splines'), which are the functions derived from $\phi_s(t) := t^{2s}$ or, alternatively $\psi_s(t) = (1 - t^2)^s$ with the index, s, a positive, pure half integer (which gives the order of zero of the kernel or its smoothness) related to the 'order' of the kernel by $s = s(m) = \frac{2m-3}{2}$. We have

$$k_m(\boldsymbol{x}, \boldsymbol{\alpha}) := \left(\sin\left(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2} \right) \right)^{2m-3} = \phi_s\left(\sin\left(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2} \right) \right)$$

$$= \left(1 - \cos^2\left(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2} \right) \right)^{\frac{2m-3}{2}} = \psi_s\left(\cos\left(\frac{\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})}{2} \right) \right). \tag{11}$$

To ensure continuity of the kernel, we will assume throughout the article that $m \geqslant 2$.

An analysis of their Wigner D-function expansions will reveal two remarkable qualities. First, the kernels are conditionally positive definite; this property and its consequences are discussed in Section 3.2. Second, they satisfy integral identities, valid for $f \in W_p^{2m}(SO(3))$, of the form

$$f = \int_{SO(3)} \mathcal{L}_m f(\boldsymbol{\alpha}) k_m(\cdot, \boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha}) \tag{12}$$

where $\mathcal{L}_m$ is an elliptic differential operator of order $2m$ on $SO(3)$. To be sure, there are many kernels satisfying a reproduction formula like (12) and many of them are conditionally positive definite. However, in general, it is unclear that they will have a 'nice', closed-form expression similar to (11). Similarly, there is no guarantee that they will have the property of localizability (developed in the next section), which is key to our understanding of the approximation power of the kernel (as discussed in Section 5).

We organize the section as follows. In Section 3.1, we expand (11) in terms of Chebyshev polynomials. This expansion is used to show that the kernels are conditionally positive definite, which leads to a discussion of some practical approximation schemes. This is treated in Section 3.2. In Section 3.3 we establish the reproduction formula (12) and demonstrate that the underlying differential operator $\mathcal{L}_m$ is of the form $p(\Delta)$, a degree m polynomial in the Laplace–Beltrami operator.

3.1. Chebyshev coefficients of the surface splines

Our investigation of k_m begins with studying its decomposition in characters (equivalently, the decomposition of ψ_s in even Chebyshev polynomials).

Lemma 2. *For $m \geqslant 2$, the ℓth coefficient of k_m in its expansion in terms of characters, cf. (8), is*

$$\tilde{k}_m(\ell) = \frac{2}{\pi} \frac{(2m-2)!}{(-4)^{m-1}} \prod_{j=-(m-1)}^{m-1} \left[\ell + j + \frac{1}{2} \right]^{-1}. \tag{13}$$

Proof. We consider $k_m(\boldsymbol{x}, \boldsymbol{\alpha}) = k_m(\boldsymbol{\alpha}^{-1}\boldsymbol{x}) = \sum_{\ell=0}^{\infty} \tilde{k}_m(\ell) c_\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x})$ and we want to determine the coefficients $\tilde{k}_m(\ell)$. Utilizing the fact that the characters are orthonormal we wish to compute

$$\tilde{k}_m(\ell) = \int\limits_{SO(3)} k_m(\boldsymbol{x}) \overline{c_\ell(\boldsymbol{x})} \, d\mu(\boldsymbol{x}).$$

By a change of variable, we get

$$\tilde{k}_m(\ell) = \frac{2}{\pi} \int\limits_0^\pi \left(\sin\left(\frac{\omega}{2}\right) \right)^{2m-3} \mathcal{U}_{2\ell}\left(\cos\left(\frac{\omega}{2}\right) \right) \sin^2\left(\frac{\omega}{2}\right) d\omega.$$

Rewriting $\mathcal{U}_{2\ell}(\cos(\frac{\omega}{2}))$ in the familiar form

$$\mathcal{U}_{2\ell}\left(\cos\left(\frac{\omega}{2}\right) \right) = \frac{\sin((2\ell+1)\frac{\omega}{2})}{\sin(\frac{\omega}{2})}$$

and applying another change of variable $\theta = \omega/2$, we obtain:

$$\tilde{k}_m(\ell) = \frac{4}{\pi} \int\limits_0^{\pi/2} (\sin\theta)^{2m-2} \sin\big((2\ell+1)\theta\big) d\theta.$$

Such integrals are easily attacked using simple trigonometric identities. They can also be reduced to the observation that repeatedly multiplying a periodic function by $\sin\theta$ is equivalent to applying a difference operator to its Fourier coefficients. Doing so, we obtain

$$(\sin\theta)^{2M} \sin(L\theta) = \left(-\frac{1}{4}\right)^M \sum_{j=0}^{2M} (-1)^j \binom{2M}{j} \sin\big([L-(2M-2j)]\theta\big).$$

Applying this identity with $M = m-1$ and $L = 2\ell+1$, and integrating each term $\sin([2\ell+3-(2m-2j)]\theta)$ over the interval $[0, \frac{\pi}{2}]$, we see that

$$\tilde{k}_m(\ell) = \left(-\frac{1}{4}\right)^{m-1} \frac{4}{\pi} \sum_{j=0}^{2m-2} \frac{(-1)^j}{2\ell+3-(2m-2j)} \binom{2m-2}{j}$$

$$= \left(-\frac{1}{4}\right)^{m-1} \frac{2}{\pi} \sum_{j=0}^{2m-2} \frac{(-1)^j}{\ell-m+\frac{3}{2}+j} \binom{2m-2}{j}.$$

Hence, $\tilde{k}_m(\ell)$ is a difference operator applied to the rational function $x \mapsto x^{-1}$. We simplify this type of expression in the elementary formula:

$$\sum_{j=0}^M (-1)^j \binom{M}{j} \frac{1}{L+j} = \frac{M!}{(L)(L+1)\cdots(L+M-1)(L+M)}. \tag{14}$$

Thus, we can simplify $\tilde{k}_m(\ell)$, by using (14) with $L = \ell - m + \frac{3}{2}$ and $M = 2m - 2$:

$$\tilde{k}_m(\ell) = \frac{2}{\pi} \left(-\frac{1}{4} \right)^{m-1} \frac{(2m-2)!}{(\ell - m + 3/2) \cdots (\ell + m - 1/2)},$$

which proves the lemma. □

3.2. Treating data with surface splines

At this point we observe that the kernels k_m are each conditionally positive definite of order $\ell_0 \geqslant m - 2$, and, hence, well suited for solving interpolation problems. Being conditionally positive definite of order ℓ_0 means that for any $\varXi \subset SO(3)$, $\sum_{\zeta \in \varXi} \sum_{\xi \in \varXi} a_\xi \overline{a_\zeta} A_{\xi,\zeta} > 0$ for any $a \neq 0$ in the subspace defined by $Ba = 0$, where

$$A := [k_m(\xi, \zeta)]_{\xi,\zeta \in \varXi} \quad \text{and} \quad B := [D_{\iota,\nu}^\ell(\xi)]_{\ell=0,\dots,\ell_0,\,-\ell \leqslant \iota,\nu \leqslant \ell;\,\xi \in \varXi}. \tag{15}$$

Indeed, conditional positive definiteness follows directly from the fact that $\tilde{k}_m(\ell)$ does not change sign for $\ell \geqslant m - 2$ (see [26, Proposition 4.3]).

We now present three classical, practical approximation schemes involving the kernel k_m. Namely, we approximate by functions of the form

$$s_\varXi = \sum_{\xi \in \varXi} \alpha_\xi k_m(\cdot, \xi) + \sum_{\ell=0}^{\ell_0} \sum_{\iota,\nu=-\ell}^{\ell} \beta_{\iota,\nu}^\ell D_{\iota,\nu}^\ell, \tag{16}$$

where the kernel coefficients satisfy

$$\sum_{\xi \in \varXi} \alpha_\xi D_{\iota,\nu}^\ell(\xi) = 0, \quad 0 \leqslant \ell \leqslant \ell_0; \; -\ell \leqslant \iota, \nu \leqslant \ell. \tag{17}$$

Interpolation: We define the interpolant to a function f at the centers $\varXi$ as the function of the form (16) where the interpolation conditions

$$s_{f,\varXi}(\xi) = f(\xi), \quad \xi \in \varXi,$$

and auxiliary conditions (17) are satisfied.

The interpolant is determined by solving the system

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \tag{18}$$

where $f = (f(\xi))_{\xi \in \varXi}$. It is well known that system (18) is uniquely solvable for any set of centers $\varXi$ that is unisolvent for the polynomial space Π_{ℓ_0}, see [2, Chapter 8.5] (we provide sufficient conditions for such sets, in terms of the fill distance, in the next section). Moreover, for our kernels k_m, setting up this system is particularly easy (see Note 1 below). We refer the interested reader to [2,26] for background on interpolation by conditionally positive definite kernels.

Tikhonov regularization: For treating noisy data, a common approach is to find the kernel approximant minimizing a certain quadratic form. For noisy data $(y_\xi)_{\xi \in \varXi}$ sampled at $\varXi \subset SO(3)$, we wish to solve the problem

$$\min_{s \in W_2^m(SO(3))} \left(\sum_{\xi \in \varXi} |s(\xi) - y_\xi|^2 + \lambda |s|_\mathcal{N}^2 \right)$$

where we use the seminorm $|s|_\mathcal{N}^2 = \sum_{\ell \geqslant \ell_0} \sum_{-\ell \leqslant \iota,\nu \leqslant \ell} |\hat{s}_{\iota,\nu}^\ell|^2 |\tilde{k}_m(\ell)|^{-1}$ and a smoothing parameter $\lambda > 0$ is chosen based on the level of noise inherent in the data. The minimizer is of the form (16). Because k_m is conditionally positive definite, the minimizer is determined by solving the system of equations

$$\begin{pmatrix} A + \lambda \, \mathrm{Id} & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}. \tag{19}$$

See [27, 2.4] for a detailed discussion of this problem.

Least-squares approximation: For a unisolvent set $\varXi$, the space

$$S_\varXi := \left\{ s = \sum_{\xi \in \varXi} \alpha_\xi k_m(\cdot, \xi) + p \mid p \in \Pi_{\ell_0}, \; (\alpha_\xi)_{\xi \in \varXi} \text{ satisfying (17)} \right\}$$

has dimension $\#\Xi$, and one can construct the least-squares operator P_Ξ, which is the orthogonal projector $P_\Xi : L_2(SO(3)) \to S_\Xi$ with range S_Ξ.

One may calculate the least-squares projector by using any basis $(v_\xi)_\xi$ for S_Ξ (say the basis of Lagrange functions χ_ξ where $\chi_\xi(\zeta) = \delta_{\xi,\zeta}$). In this case, the projector is simply $P_\Xi f = \sum_{\xi \in \Xi} a_\xi v_\xi$, with coefficients $\boldsymbol{a} = (a_\xi)_{\xi \in \Xi}$ given by $\boldsymbol{a} = \boldsymbol{G}^{-1}\boldsymbol{f}$, where $\boldsymbol{G}_{\xi,\zeta} = \langle v_\xi, v_\zeta \rangle$ is the Gram matrix for the basis and $\boldsymbol{f} = (\langle v_\xi, f \rangle)_{\xi \in \Xi}$. Since the elements of S_Ξ are continuous and $SO(3)$ is compact, the projector extends as an operator to each space $L_p(SO(3))$, $1 \leqslant p \leqslant \infty$. The L_p stability of such operators on compact Riemannian manifolds has recently been investigated in [28]; an open problem, [28, Conjecture 3], is whether the stability and approximation results of that article hold for the kernels presented here.

Note 1. Suppose we are given centers $\Xi = \{\xi_j \in SO(3), \ j = 0, \ldots, M-1\}$ in terms of Euler angles (as described in Section 2.1), i.e., each ξ_j is determined by $(\varphi_{1,j}, \theta_j, \varphi_{2,j})$. It is easy to show that we can write the distance of two centers ξ_i and ξ_j in terms of Euler angles; in fact we have

$$\cos\left(\frac{\text{dist}(\xi_i, \xi_j)}{2}\right) = \left(\left|\cos\left(\frac{\varphi_{1,i} - \varphi_{1,j}}{2}\right) \cos\left(\frac{\varphi_{2,i} - \varphi_{2,j}}{2}\right) \cos\left(\frac{\theta_i - \theta_j}{2}\right)\right.\right.$$
$$\left.\left. - \sin\left(\frac{\varphi_{1,i} - \varphi_{1,j}}{2}\right) \sin\left(\frac{\varphi_{2,i} - \varphi_{2,j}}{2}\right) \cos\left(\frac{\theta_i + \theta_j}{2}\right)\right|\right).$$

Thus, we can directly write down the closed-form expression for the entry $\boldsymbol{A}_{i,j}$ of the matrix $\boldsymbol{A}$ as

$$\boldsymbol{A}_{i,j} = \left(1 - \cos^2\left(\frac{\text{dist}(\xi_i, \xi_j)}{2}\right)\right)^{\frac{2m-3}{2}}.$$

Furthermore, the nonequispaced Fourier matrix $\boldsymbol{B}$ can be computed directly from (5). Such matrices have been studied in great detail and can be handled efficiently, see [9].

On the other hand, if the centers are given via their rotation axis and rotation angle, we can immediately compute the corresponding Euler angles, using the direct relation between the Euler angle and the axis-angle parameterization of $SO(3)$, see [29, Chapter 1.4.4]. By doing so, we can again easily set up the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ as shown above.

3.3. Integral representation

Returning to our investigation of the coefficients $\tilde{k}_m(\ell)$, we observe that $\tilde{k}_m(\ell)$ can be simplified by arranging factors symmetrically around $\ell + 1/2$. Doing this, we have:

$$\tilde{k}_m(\ell) = \frac{2}{\pi} \frac{(2m-2)!}{(-4)^{m-1}} \left[\prod_{j=1}^{m-1} \frac{1}{(\ell+1/2) - j}\right] \frac{1}{\ell+1/2} \left[\prod_{j=1}^{m-1} \frac{1}{(\ell+1/2) + j}\right]$$

$$= \frac{2}{\pi} \frac{(2m-2)!}{(-4)^{m-1}} \frac{1}{(\ell+1/2)} \prod_{j=1}^{m-1} \frac{1}{(\ell+1/2)^2 - j^2}. \tag{20}$$

The last line comes from rearranging terms, and noting that $[J-k][J+k] = J^2 - k^2$. Since $(\ell+1/2)^2 = \ell(\ell+1) + 1/4$, we can rewrite the expression in terms of $\ell(\ell+1)$ and obtain (after reindexing)

$$\tilde{k}_m(\ell) = \frac{1}{\pi} \frac{(2m-2)!}{(-4)^{m-1}} (2\ell+1) \prod_{j=0}^{m-1} \frac{1}{\ell(\ell+1) - [j^2 - 1/4]}. \tag{21}$$

The coefficients $\tilde{k}_m(\ell)$ involve the *symbol* for Δ, $\sigma(\Delta) = \ell(\ell+1)$. In fact, these are the reciprocals of a polynomial in the symbol, $p(\sigma(\Delta))^{-1}$, which is nonvanishing, because the numbers $[j^2 - 1/4]$ never coincide with eigenvalues of Δ. So the operator induced by k_m inverts the elliptic operator $p(\Delta)$. This is the point of the following lemma:

Lemma 3. *The kernel* $(\boldsymbol{x}, \boldsymbol{\alpha}) \mapsto k_m(\boldsymbol{x}, \boldsymbol{\alpha})$ *satisfies the formula*

$$f = \int_{SO(3)} k_m(\cdot, \boldsymbol{\alpha}) \mathcal{L}_m f(\boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha})$$

using the operator of order $2m$, $\mathcal{L}_m := \frac{\pi(-4)^{m-1}}{(2m-2)!} \prod_{j=0}^{m-1} (\Delta - r_j)$ with $r_j := j^2 - 1/4$, and which holds for all $f \in W_p^{2m}(SO(3))$ (or $C^{2m}(SO(3))$ when $p = \infty$).

Proof. For a sufficiently smooth f (say C^∞ initially), we have the absolutely convergent series $f = \sum_\ell \sum_{\iota,\nu} c^\ell_{\iota,\nu} D^\ell_{\iota,\nu}$. Using the addition formula (9), property (6) and the orthogonality of the Wigner D-functions we get

$$
\int\limits_{SO(3)} k_m(\boldsymbol{x},\boldsymbol{\alpha}) \mathcal{L}_m f(\boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha}) = \int\limits_{SO(3)} \sum_\ell \tilde{k}_m(\ell) c_\ell(\boldsymbol{\alpha}^{-1}\boldsymbol{x}) \frac{\pi(-4)^{m-1}}{(2m-2)!} \prod_{j=0}^{m-1} (\Delta - r_j) \sum_{\ell'} \sum_{\iota',\nu'} c^{\ell'}_{\iota',\nu'} D^{\ell'}_{\iota',\nu'}(\boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha})
$$

$$
= \int\limits_{SO(3)} \sum_\ell \sum_{\iota,\nu} \tilde{k}_m(\ell) D^\ell_{\iota,\nu}(\boldsymbol{x}) \overline{D^\ell_{\iota,\nu}(\boldsymbol{\alpha})}
$$

$$
\times \sum_{\ell'} \sum_{\iota',\nu'} c^{\ell'}_{\iota',\nu'} \frac{\pi(-4)^{m-1}}{(2m-2)!} \prod_{j=0}^{m-1} \big(\ell'(\ell'+1) - r_j\big) D^{\ell'}_{\iota',\nu'}(\boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha})
$$

$$
= \sum_\ell \sum_{\iota,\nu} \tilde{k}_m(\ell) \frac{\pi(-4)^{m-1}}{(2m-2)!} \prod_{j=0}^{m-1} \big[\ell(\ell+1) - r_j\big] \frac{1}{2\ell+1} c^\ell_{\iota,\nu} D^\ell_{\iota,\nu}(\boldsymbol{x}) = f(\boldsymbol{x}).
$$

The reproducing formula follows for all f by a limiting argument, using the density of C^∞ in W^{2m}_p for $1 \leqslant p < \infty$ (or C^{2m} when $p = \infty$) and the boundedness of $\mathcal{L}_m$. $\quad\square$

4. Localizing the kernel

For a set of centers $\boldsymbol{\varXi} \subset SO(3)$ we now wish to investigate a 'coefficient kernel' $\mathbf{a} : \boldsymbol{\varXi} \times SO(3) \to \mathbb{R}$ that will allow us to easily replace $k_m(\boldsymbol{x},\boldsymbol{\alpha})$ with $\bar{k}(\boldsymbol{x},\boldsymbol{\alpha}) := \sum_{\boldsymbol{\xi} \in \boldsymbol{\varXi}} \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha}) k_m(\boldsymbol{x},\boldsymbol{\xi})$ in the representation (12). To do so, we must estimate the cost of replacing the kernel, e_{k_m}, given by the 'error kernel':

$$
e_{k_m}(\boldsymbol{x},\boldsymbol{\alpha}) := \left| k_m(\boldsymbol{x},\boldsymbol{\alpha}) - \sum_{\boldsymbol{\xi} \in \boldsymbol{\varXi}} \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha}) k_m(\boldsymbol{x},\boldsymbol{\xi}) \right|.
$$

This is a bounded, rapidly decaying function, inducing an operator on L_p with small norm.

The outline of this section is as follows. In Section 4.1 we develop conditions on the coefficient kernel $\mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha})$ necessary for proving our results, and it is shown that such conditions are satisfied for sufficiently dense centers. A practical procedure for producing the coefficients is given in Section 4.2. In Section 4.3, we estimate the decay of the error kernel $e_{k_m}(\boldsymbol{x},\boldsymbol{\alpha})$.

4.1. Coefficient Kernel Conditions

The two key quantities we need to resolve are the polynomial precision (the degree of Wigner D-functions reproduced by the coefficients $\mathbf{a}(\cdot,\boldsymbol{\alpha})$) and the rate of decay of the error kernel. As in the Euclidean and spherical setting, these are related: the greater the polynomial precision, the more local the replacement error.

Definition 4 *(Coefficient Kernel Conditions).* For a set of centers $\boldsymbol{\varXi} \subset SO(3)$ there is a number $\rho = \rho(\boldsymbol{\varXi})$, reflecting the density of $\boldsymbol{\varXi}$ in $SO(3)$, and a measurable kernel $\mathbf{a} : \boldsymbol{\varXi} \times SO(3) \to \mathbb{R}$ satisfying the following three conditions:

CKC 1 (*Support*) $\mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha}) = 0$ for $\mathrm{dist}(\boldsymbol{\alpha},\boldsymbol{\xi}) > \rho$.
CKC 2 (*Precision*) For any polynomial of degree at most L, i.e., for any $p \in \boldsymbol{\varPi}_L$, the following holds:

$$
\sum_{\boldsymbol{\xi} \in \boldsymbol{\varXi}} \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha}) p(\boldsymbol{\xi}) = p(\boldsymbol{\alpha}) \quad \text{for every } \boldsymbol{\alpha} \in SO(3).
$$

CKC 3 (*Stability*) There is a constant $K \geqslant 1$ such that for every $\boldsymbol{\alpha} \in SO(3)$, we have

$$
\sum_{\boldsymbol{\xi} \in \boldsymbol{\varXi}} \big| \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha}) \big| \leqslant K.
$$

As in the Euclidean case, such a local polynomial reproduction property holds for sufficiently dense centers. In order to show this we need the following fact about polynomials on $SO(3)$.

Proposition 5. *Let $p \in \boldsymbol{\varPi}_L$. Then p restricted to a geodesic segment on $SO(3)$ gives a trigonometric polynomial of degree at most L.*

Proof. Let $p = \sum_{\ell=0}^{L} \sum_{\iota,\nu=-\ell}^{\ell} a^\ell_{\iota,\nu} D^\ell_{\iota,\nu} \in \boldsymbol{\varPi}_L$ and an arbitrary geodesic segment $G = \{\boldsymbol{x}_0 \boldsymbol{s}_\eta(\omega) : \omega \in [0,t]\}$ of length $t \in [0,2\pi)$ on $SO(3)$ be given. Here $\boldsymbol{s}_\eta(\omega)$ denotes a rotation about the rotation axis $\eta \in \mathbb{S}^2$ with rotation angle $\omega \in [0,2\pi)$. Then we

can find $\boldsymbol{y} \in SO(3)$ such that $\boldsymbol{s}_\eta(\omega) = \boldsymbol{y}^{-1}\boldsymbol{s}_z(\omega)\boldsymbol{y}$ for all $\omega \in [0, t]$. The rotation $\boldsymbol{s}_z(\omega)$ can, for example, be represented by the Euler angles $(\omega, 0, 0)$. Thus, we get, by using the representation (5), for $\boldsymbol{x}(\omega) = \boldsymbol{x}_0 \boldsymbol{s}_\eta(\omega) \in G$

$$
\begin{aligned}
p\big(\boldsymbol{x}(\omega)\big) = p\big(\boldsymbol{x}_0 \boldsymbol{y}^{-1}\boldsymbol{s}_z(\omega)\boldsymbol{y}\big) &= \sum_{\ell=0}^{L} \sum_{\iota,\nu=-\ell}^{\ell} a_{\iota,\nu}^\ell D_{\iota,\nu}^\ell \big(\boldsymbol{x}_0 \boldsymbol{y}^{-1}\boldsymbol{s}_z(\omega)\boldsymbol{y}\big) \\
&= \sum_{\ell=0}^{L} \mathrm{Tr}\big(\hat{A}_\ell^T D^\ell\big(\boldsymbol{x}_0 \boldsymbol{y}^{-1}\boldsymbol{s}_z(\omega)\boldsymbol{y}\big)\big) \\
&= \sum_{\ell=0}^{L} \mathrm{Tr}\Big(\underbrace{D^\ell(\boldsymbol{y})\hat{A}_\ell^T D^\ell\big(\boldsymbol{x}_0 \boldsymbol{y}^{-1}\big)}_{\hat{B}_\ell^T} D^\ell\big(\boldsymbol{s}_z(\omega)\big)\Big) \\
&= \sum_{\ell=0}^{L} \sum_{\iota,\nu=-\ell}^{\ell} b_{\iota,\nu}^\ell D_{\iota,\nu}^\ell\big(\boldsymbol{s}_z(\omega)\big) = \sum_{\ell=0}^{L} \sum_{\iota=-\ell}^{\ell} b_{\iota,\iota}^\ell e^{-i\iota\omega}.
\end{aligned}
$$

We note that the operator $\hat{B}_\ell$ and its entries $b_{\iota,\nu}^\ell$ are independent of ω. The final equality is a consequence of (5). $\quad\square$

The Coefficient Kernel Conditions are easily satisfied. The following lemma shows that for sufficiently dense centers, one can find a suitable kernel.

Lemma 6. *Given L, an integer reflecting the desired precision, and given centers $\boldsymbol{\varXi}$ having fill distance $h \leqslant h_0$ (where $h_0 := \frac{\pi}{CL^2}$, a constant depending only on L), there exists a coefficient kernel $\mathbf{a} : \boldsymbol{\varXi} \times SO(3) \to \mathbb{R}$ satisfying the CKCs with radius $\rho = CL^2 h$ (with $C > 0$ a global constant) and $K = 2$.*

Before providing the proof, we remark that the restriction $\rho \leqslant \pi$ forces $h \leqslant \frac{\pi}{CL^2}$, which is the only restriction placed on h.

Proof. Proving the existence of a coefficient kernel follows a 'norming set argument', which relies on being able to capture the norm of a polynomial over a ball by its restriction to the centers inside the ball:

$$
\|P\|_{L_\infty(B(\boldsymbol{p}_0,\rho))} \leqslant 2\|P_{|\boldsymbol{\varXi}\cap B(\boldsymbol{p}_0,\rho)}\|_{\ell_\infty(\boldsymbol{\varXi}\cap B(\boldsymbol{p}_0,\rho))}. \tag{22}
$$

Utilizing this to obtain a local polynomial reproduction is a fairly standard procedure, and the basics of our argument can be found in [17, Lemma 4.2] or [2, Chapter 3]. So all we need to show is (22).

The inequality (22) relies on locating a center $\boldsymbol{\xi} \in \boldsymbol{\varXi}$ near the point where the uniform norm is attained, $\boldsymbol{p} \in B(\boldsymbol{p}_0, \rho)$, so that a geodesic segment of length ω, starting at $\boldsymbol{p}$ and passing through $\boldsymbol{\xi}$ lies entirely within $B(\boldsymbol{p}_0, \rho)$. Specifically, we require

$$
\frac{\mathrm{dist}(\boldsymbol{p}, \boldsymbol{\xi})}{\omega} \leqslant \frac{1}{16L^2}. \tag{23}
$$

The geodesic segment is determined by the curve $\gamma : [-\omega/2, \omega/2] \mapsto SO(3)$, parameterized by arclength. The main estimate we employ is Videnskiĭ's inequality [30] which reads as follows. Let T be a trigonometric polynomial of degree at most L, then

$$
\big|T'(t)\big| \leqslant 2L^2 \cot(\omega/4)\|T\|_{L_\infty(-\omega/2,\omega/2)} \quad \text{for } \omega/2 < \pi, \ |t| \leqslant \omega/2.
$$

The trigonometric polynomial $\tau = P \circ \gamma$ is the restriction of P, the polynomial in question, to the geodesic segment. We then observe that $\tau(-\omega/2) - \tau(-\omega/2 + \mathrm{dist}(\boldsymbol{p}, \boldsymbol{\xi})) \leqslant \int_{-\omega/2}^{-\omega/2+\mathrm{dist}(\boldsymbol{p},\boldsymbol{\xi})} |\tau'(t)|\,dt$. Applying Videnskiĭ's inequality

$$
P(\boldsymbol{p}) - P(\boldsymbol{\xi}) \leqslant \frac{8L^2}{\omega}\mathrm{dist}(\boldsymbol{p},\boldsymbol{\xi})\|\tau\|_{L_\infty(-\omega/2,\omega/2)} \leqslant \frac{1}{2}P(\boldsymbol{p}).
$$

This gives us the desired inequality: $P(\boldsymbol{p}) \leqslant 2P(\boldsymbol{\xi})$.

The fact that (23) holds follows by employing a cone condition. On $SO(3)$, a cone $c(\boldsymbol{q}, \theta, \nu, r)$ is the union of geodesic segments of length r emanating from $\boldsymbol{q}$ and having the initial tangent vector within θ of the tangent vector ν. We call θ the aperture and ν the direction of the cone. It is an elementary fact that every ball in $SO(3)$ satisfies a uniform cone condition. I.e., there is $\theta > 0$ so that for any $\boldsymbol{p}_0 \in SO(3)$ and $\rho > 0$ and for every point in the ball $B(\boldsymbol{p}_0, \rho)$, the cone $c(\boldsymbol{p}, \theta, \nu, \rho/2) \subset B(\boldsymbol{p}_0, \rho)$, where ν is the tangent vector at $\boldsymbol{p}$ of the geodesic segment connecting $\boldsymbol{p}$ and $\boldsymbol{p}_0$.

An appeal to geometry shows that there is a global constant C' depending only on θ and the geometry of $SO(3)$ so that a smaller ball $b \subset SO(3)$ having small radius h can be placed entirely within the cone $c(\boldsymbol{p}, \theta, \nu, \rho/2)$ with its center a distance of $C'h$ from $\boldsymbol{p}$. Thus there is a center $\boldsymbol{\xi}$ that is a distance $Ch \leqslant (C'+1)h$ from $\boldsymbol{p}$, for which the geodesic segment connecting $\boldsymbol{p}$ and $\boldsymbol{\xi}$ of length $\rho/2$ is entirely within $B(\boldsymbol{p}_0, \rho)$. Therefore, $\rho = 32L^2Ch$ suffices. $\quad\square$

4.2. Constructing coefficients in practice

Although the previous lemma guarantees the existence of a coefficient kernel sufficient for our purposes, it does not provide a viable method for generating such kernels in practice. To do so, we make the extra assumption that the centers are quasi-uniform: that h and q, as defined in Section 2, are related by $h/q \leqslant \mathfrak{m}$ for a fixed 'mesh ratio' $\mathfrak{m} < \infty$. Armed with this assumption and using (4), we can estimate the number of centers in a given ball of radius $\rho = CL^2h$ (as was used in the previous lemma):

$$\#\boldsymbol{\varXi}_{\boldsymbol{\alpha}} = \#\bigl(B(\boldsymbol{\alpha}, \rho) \cap \boldsymbol{\varXi}\bigr) \leqslant \frac{b_2(CL^2h)^3}{b_1 q^3} \leqslant \left(\frac{\pi}{2}\right)^2 C^3 L^6 \mathfrak{m}^3.$$

We may determine a modified coefficient kernel, which we call $\tilde{\mathbf{a}}$ that satisfies the Coefficient Kernel Conditions by taking the ℓ_2 minimizing solution at each $\boldsymbol{\alpha}$ of the system $\boldsymbol{B}\mathbf{a} = \mathrm{D}_{\boldsymbol{\alpha}}$, where $(\mathrm{D}_{\boldsymbol{\alpha}})_{(j,k,\ell)} = D_{jk}^{\ell}(\boldsymbol{\alpha})$ is the vector of Wigner D-functions evaluated at $\boldsymbol{\alpha}$, the matrix $\boldsymbol{B}$ is the nonequispaced Fourier matrix introduced in (15). I.e., the matrix whose rows are the D-functions evaluated at the various $\boldsymbol{\xi}$ in $\boldsymbol{\varXi}_{\boldsymbol{\alpha}}$. I.e., $(\boldsymbol{B})_{(j,k,\ell),\boldsymbol{\xi}} = D_{jk}^{\ell}(\boldsymbol{\xi})$.

It follows that the ℓ_2 minimizing solution is given by $\boldsymbol{B}^*(\boldsymbol{B}\boldsymbol{B}^*)^{-1}\mathrm{D}_{\boldsymbol{\alpha}}$, so we obtain a coefficient kernel $\tilde{\mathbf{a}}(\boldsymbol{\xi}, \boldsymbol{\alpha})$ satisfying conditions CKC 1 and CKC 2 of Definition 4 via $(\tilde{\mathbf{a}}(\boldsymbol{\xi}, \boldsymbol{\alpha}))_{\boldsymbol{\xi} \in \boldsymbol{\varXi}_{\boldsymbol{\alpha}}} = \boldsymbol{B}^*(\boldsymbol{B}\boldsymbol{B}^*)^{-1}\mathrm{D}_{\boldsymbol{\alpha}}$ and zero extension. Moreover, it is clear that on the subset

$$\bigl\{\boldsymbol{\alpha}' \in SO(3) \,\big|\, \boldsymbol{\varXi}_{\boldsymbol{\alpha}'} = \boldsymbol{\varXi}_{\boldsymbol{\alpha}}\bigr\}$$

(i.e., the set of $\boldsymbol{\alpha}'$ for which the sets $\boldsymbol{\varXi}_{\boldsymbol{\alpha}'}$ coincide) each $\tilde{\mathbf{a}}(\boldsymbol{\xi}, \cdot)$ is a Wigner D-function. Hence, $\tilde{\mathbf{a}}$ is measurable.

To show that $\tilde{\mathbf{a}}$ satisfies CKC 3 we can compare it to the kernel $\mathbf{a}$ guaranteed by Lemma 6 to observe that

$$\begin{aligned}
\bigl\|\tilde{\mathbf{a}}(\cdot, \boldsymbol{\alpha})\bigr\|_{\ell_1(\boldsymbol{\varXi})} &\leqslant (\#\boldsymbol{\varXi}_{\boldsymbol{\alpha}})^{1/2} \bigl\|\tilde{\mathbf{a}}(\cdot, \boldsymbol{\alpha})\bigr\|_{\ell_2(\boldsymbol{\varXi})} \leqslant \sqrt{(\pi/2)^2 C^3 L^6 \mathfrak{m}^3} \bigl\|\mathbf{a}(\cdot, \boldsymbol{\alpha})\bigr\|_{\ell_2(\boldsymbol{\varXi})} \\
&\leqslant \frac{\pi}{2}\sqrt{C^3 L^6 \mathfrak{m}^3} \bigl\|\mathbf{a}(\cdot, \boldsymbol{\alpha})\bigr\|_{\ell_1(\boldsymbol{\varXi})} \leqslant \pi\sqrt{C^3 L^6 \mathfrak{m}^3}.
\end{aligned}$$

So CKC 3 is satisfied with $K = \pi\sqrt{C^3 L^6 \mathfrak{m}^3}$.

4.3. Estimating the cost of replacing the kernels

A consequence of the Coefficient Kernel Conditions is that for any $\boldsymbol{x} \in SO(3)$ and a smooth class function κ, there is a convenient mechanism to estimate the error $e_\kappa(\boldsymbol{x}, \boldsymbol{\alpha})$ in terms of an interval around $\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})$. For fixed $\boldsymbol{x} \in SO(3)$ we choose to express the kernel κ in terms of $\tau_{\boldsymbol{x}}(\boldsymbol{\zeta}) = \cos(\frac{\omega(\boldsymbol{\zeta}^{-1}\boldsymbol{x})}{2})$, by way of the formula $\kappa(\boldsymbol{x}, \boldsymbol{\zeta}) = \vartheta(\tau_{\boldsymbol{x}}^2(\boldsymbol{\zeta}))$ assuming ϑ is smooth on the interval $\mathcal{I}_{\boldsymbol{x}} := [\min Q_{\boldsymbol{x}}, \max Q_{\boldsymbol{x}}]$, where the set $Q_{\boldsymbol{x}}$ depends on the rotation angles used, $Q_{\boldsymbol{x}} := \{\tau_{\boldsymbol{x}}^2(\boldsymbol{\alpha})\} \cup \{\tau_{\boldsymbol{x}}^2(\boldsymbol{\xi}): \boldsymbol{\xi} \in \boldsymbol{\varXi} \cap B(\boldsymbol{\alpha}, \rho)\}$. The cost of replacing the kernel can be estimated in terms of the length of the interval $\mathcal{I}_{\boldsymbol{x}}$ and the size of derivatives of ϑ purely on $\mathcal{I}_{\boldsymbol{x}}$. This is the gist of the following lemma:

Lemma 7. *Given a kernel* $\kappa(\boldsymbol{x}, \boldsymbol{\zeta}) = \vartheta(\tau_{\boldsymbol{x}}^2(\boldsymbol{\zeta}))$, *with* $\vartheta \in C^{L+1}(\mathcal{I}_{\boldsymbol{x}})$, *a coefficient kernel* $\mathbf{a}$ *satisfying the CKCs with precision* L, *and the replacement error* $e_\kappa(\boldsymbol{x}, \boldsymbol{\alpha}) = |\kappa(\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})) - \sum_{\boldsymbol{\xi} \in \boldsymbol{\varXi}} \mathbf{a}(\boldsymbol{\xi}, \boldsymbol{\alpha})\kappa(\omega(\boldsymbol{\xi}^{-1}\boldsymbol{x}))|$. *We have the estimate*:

$$e_\kappa(\boldsymbol{x}, \boldsymbol{\alpha}) \leqslant \frac{\|\mathbf{a}(\cdot, \boldsymbol{\alpha})\|_{\ell_1(\boldsymbol{\varXi})}}{(L+1)!} |\mathcal{I}_{\boldsymbol{x}}|^{L+1} \max_{t \in \mathcal{I}_{\boldsymbol{x}}} \bigl|\vartheta^{(L+1)}(t)\bigr|. \tag{24}$$

Proof. Let both $\boldsymbol{x}$ and $\boldsymbol{\alpha}$ be fixed, and choose the Taylor polynomial of degree L, $q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}$, of ϑ expanded about $t_{\boldsymbol{x}}(\boldsymbol{\alpha}) = \tau_{\boldsymbol{x}}^2(\boldsymbol{\alpha})$. Now $q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}$ may be rewritten as a linear combination of even degree Chebyshev polynomials,

$$q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}(t) = \sum_{\ell=0}^{L} c_\ell \mathcal{U}_{2\ell}(\sqrt{t}).$$

This follows because $\mathcal{U}_{2\ell}$ is of exact degree 2ℓ and is a linear combination of even degree monomials only, so $\mathcal{U}_{2\ell}(\sqrt{t})$ is a polynomial of exact degree ℓ, and, thus, the set $(\mathcal{U}_{2\ell}(\sqrt{t}))_{\ell=0}^{L}$ is a basis for set of polynomials up to degree L. Note, furthermore, that $q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}(t_{\boldsymbol{x}}(\boldsymbol{\alpha})) - \sum \mathbf{a}(\boldsymbol{\xi}, \boldsymbol{\alpha})q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}(t_{\boldsymbol{x}}(\boldsymbol{\xi})) = 0$ by the addition formula, since

$$q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}\big(\tau_{\boldsymbol{x}}^2(\boldsymbol{\zeta})\big) = \sum_{\ell=0}^{L} c_\ell \mathcal{U}_{2\ell}\big(\tau_{\boldsymbol{x}}(\boldsymbol{\zeta})\big) = \sum_{\ell \leqslant L} c_\ell \sum_{k,m} D_{k,m}^\ell(\boldsymbol{x})\,\overline{D_{k,m}^\ell(\boldsymbol{\zeta})}$$

and each $D_{k,m}^\ell(\boldsymbol{\zeta})$ is annihilated by $\mu = \delta_{\boldsymbol{\alpha}} - \sum \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha})\delta_{\boldsymbol{\xi}}$.

Consequently, $e_\kappa(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant \sum_{\boldsymbol{\xi}} |\mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha})||\vartheta(t_{\boldsymbol{x}}(\boldsymbol{\xi})) - q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}(t_{\boldsymbol{x}}(\boldsymbol{\xi}))|$, and the lemma follows from the remainder in Taylor's theorem:

$$\big|\vartheta\big(t_{\boldsymbol{x}}(\boldsymbol{\xi})\big) - q_{L,t_{\boldsymbol{x}}(\boldsymbol{\alpha})}\big(t_{\boldsymbol{x}}(\boldsymbol{\xi})\big)\big| \leqslant \frac{1}{(L+1)!}\big|t_{\boldsymbol{x}}(\boldsymbol{\xi}) - t_{\boldsymbol{x}}(\boldsymbol{\alpha})\big|^{L+1} \max_{t \in \mathcal{I}_{\boldsymbol{x}}}\big|\vartheta^{(L+1)}(t)\big|. \qquad \square$$

We now measure e_{k_m}, the cost of replacing $k_m(\boldsymbol{x},\boldsymbol{\alpha})$ by a linear combination of its copies, $\sum_{\boldsymbol{\xi}} \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha})k_m(\boldsymbol{x},\boldsymbol{\xi})$.

Lemma 8. *Assume* $\mathbf{a}$ *is a coefficient kernel satisfying the CKCs with radius* ρ, *precision* $L \geqslant 2m$ *and stability constant* $K = \sup_{\boldsymbol{\alpha}} \|\mathbf{a}(\cdot,\boldsymbol{\alpha})\|_{\ell_1(\boldsymbol{\Xi})}$. *Then the replacement error of the kernel* k_m *satisfies*

$$e_{k_m}(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant C(L,m)K\rho^{2m-3}\left(1 + \frac{\mathrm{dist}(\boldsymbol{x},\boldsymbol{\alpha})}{\rho}\right)^{2m-4-L},$$

where $C(L,m)$ *is a constant that only depends on* L *and* m.

Proof. We break this into two parts, the first concerns $\boldsymbol{x}$ near to $\boldsymbol{\alpha}$ (say $\mathrm{dist}(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant 2\rho$) where we can make use of the high order zero of the kernel, while the second treats $\boldsymbol{x}$ sufficiently far from $\boldsymbol{\alpha}$ so that we can utilize the smoothness of the kernel and Lemma 7.

Case 1. Assume $\mathrm{dist}(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant 2\rho$. This implies that $\mathrm{dist}(\boldsymbol{x},\boldsymbol{\xi}) \leqslant 3\rho$ for every $\boldsymbol{\xi} \in \mathrm{supp}(\mathbf{a}(\cdot,\boldsymbol{\alpha}))$. Writing the kernel as $k_m(\boldsymbol{x},\boldsymbol{\zeta}) = (\sin \frac{\mathrm{dist}(\boldsymbol{x},\boldsymbol{\zeta})}{2})^{2m-3}$ and using the stability of the coefficient kernel, we see that

$$\left|k_m(\boldsymbol{x},\boldsymbol{\alpha}) - \sum \mathbf{a}(\boldsymbol{\xi},\boldsymbol{\alpha})k_m(\boldsymbol{x},\boldsymbol{\xi})\right| \leqslant (1+K)\left(\sin\frac{3}{2}\rho\right)^{2m-3} \leqslant (1+K)\left(\frac{3}{2}\right)^{2m-3}\rho^{2m-3}.$$

Because $1 + K \leqslant 2K \leqslant (3/2)^2 K$, it follows that $e_{k_m}(\boldsymbol{x},\boldsymbol{\alpha})$ is less than $(3/2)^{2m-1}K\rho^{2m-3}$. The condition that $\mathrm{dist}(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant 2\rho$ implies that $3 \geqslant 1 + \mathrm{dist}(\boldsymbol{x},\boldsymbol{\alpha})/\rho$, which handles the first case.

Case 2. We use the fact, from the definition of the surface spline (11), that $k(\boldsymbol{x},\boldsymbol{\zeta}) = (1 - \cos^2 \frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2})^s$ with $2s = 2m - 3$. Since $0 \leqslant \mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x}) \leqslant \pi$, we can rewrite the kernel as $k(\boldsymbol{x},\boldsymbol{\zeta}) = \vartheta_s(t_{\boldsymbol{x}}(\boldsymbol{\zeta}))$, with $t_{\boldsymbol{x}}(\boldsymbol{\zeta}) := \cos^2(\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2}) = (\tau_{\boldsymbol{x}}(\boldsymbol{\zeta}))^2$ and $\vartheta_s := (1 - \cdot)^s$. We note that ϑ_s is smooth for t less than 1. Lemma 7 tells us that the error is controlled by $|\mathcal{I}_{\boldsymbol{x}}|^{L+1}$. The length of $\mathcal{I}_{\boldsymbol{x}}$ can be written as

$$\begin{aligned}
\big|t_{\boldsymbol{x}}(\boldsymbol{\xi}) - t_{\boldsymbol{x}}(\boldsymbol{\zeta})\big| &= \left|\cos^2\left(\frac{\mathrm{dist}(\boldsymbol{\xi},\boldsymbol{x})}{2}\right) - \cos^2\left(\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2}\right)\right| \\
&= \left|\sin\left(\frac{\mathrm{dist}(\boldsymbol{\xi},\boldsymbol{x})}{2}\right) - \sin\left(\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2}\right)\right|\left|\sin\left(\frac{\mathrm{dist}(\boldsymbol{\xi},\boldsymbol{x})}{2}\right) + \sin\left(\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2}\right)\right|.
\end{aligned}$$

Thus, for $\boldsymbol{x}$ in a sufficiently wide annulus about $\boldsymbol{\alpha}$, say $2^n\rho \leqslant \mathrm{dist}(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant 2^{n+1}\rho$, we have $2^{n-1}\rho \leqslant \mathrm{dist}(\boldsymbol{x},\boldsymbol{\xi}) \leqslant 2^{n+2}\rho$ for all $\boldsymbol{\xi} \in \mathrm{supp}(\mathbf{a}(\cdot,\boldsymbol{\alpha}))$, and we can estimate $|\mathcal{I}_{\boldsymbol{x}}|$ by

$$|\mathcal{I}_{\boldsymbol{x}}| \leqslant \frac{|\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{\xi})|}{2} \times \frac{|\mathrm{dist}(\boldsymbol{x},\boldsymbol{\xi}) + \mathrm{dist}(\boldsymbol{x},\boldsymbol{\zeta})|}{2} \leqslant \rho \times \big(2^{n+2}\rho\big).$$

Applying Lemma 7 gives

$$e_{k_m}(\boldsymbol{x},\boldsymbol{\alpha}) \leqslant \frac{K}{(L+1)!}\big\|\vartheta_s^{(L+1)}\big\|_{L_\infty(\mathcal{I}_{\boldsymbol{x}})}\rho^{2(L+1)}2^{(n+2)(L+1)}.$$

Differentiating gives $\|\vartheta_s^{(L+1)}\|_{L_\infty(\mathcal{I}_{\boldsymbol{x}})} \leqslant |\prod_{j=0}^{L}(s-j)|\max_{t \in \mathcal{I}_x}(1-t)^{s-L-1}$, and we note that the maximum is attained at $t = t_{\boldsymbol{x}}(\boldsymbol{\zeta}) = \cos^2\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2}$ for the rotation $\boldsymbol{\zeta} \in \{\boldsymbol{\alpha}\} \cup \mathrm{supp}(\mathbf{a}(\cdot,\boldsymbol{\alpha}))$ nearest to $\boldsymbol{x}$. Rewriting $1 - t_{\boldsymbol{x}}(\boldsymbol{\zeta})$ as $\sin^2\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2}$, and using the inequality $\sin\theta \geqslant \frac{2}{\pi}\theta$ gives $1 - t_{\boldsymbol{x}}(\boldsymbol{\zeta}) \geqslant (\frac{2}{\pi}\frac{\mathrm{dist}(\boldsymbol{\zeta},\boldsymbol{x})}{2})^2 \geqslant (\frac{2^{n-1}\rho}{\pi})^2$. Thus,

$$\begin{aligned}
\big\|\vartheta_s^{(L+1)}\big\|_{L_\infty(\mathcal{I}_x)} &\leqslant |s(s-1)\cdots(s-L)|\left(\frac{2^{2(n-1)}\rho^2}{\pi^2}\right)^{s-L-1} \\
&= |s(s-1)\cdots(s-L)|\pi^{2(L+1-s)}\rho^{2(s-L-1)}2^{2(n-1)(s-L-1)}.
\end{aligned}$$

Together with the estimate of the length of $|\mathcal{I}_{\boldsymbol{x}}|$, we obtain

$$e_{k_m}(\boldsymbol{x}, \boldsymbol{\alpha}) \leqslant K \frac{(8\pi)^{2(L+1-s)}}{(L+1)!} |s(s-1)\cdots(s-L)| \rho^{2s} 2^{(n+2)(2s-1-L)}.$$

Because $2^{n+2} \geqslant 1 + \frac{\text{dist}(\boldsymbol{x},\boldsymbol{\alpha})}{\rho}$, the result follows.   □

5. Approximation scheme and main results

We are now ready to introduce the approximation scheme central to our error estimates. Given a coefficient kernel **a** satisfying the CKCs with precision $L \geqslant 2m$, we define the operator $T_{\boldsymbol{\Xi}}$ on $W_p^{2m}(SO(3))$ by

$$T_{\boldsymbol{\Xi}} f(\boldsymbol{x}) = \sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} A_{\boldsymbol{\xi}} k_m(\boldsymbol{x}, \boldsymbol{\xi}) \tag{25}$$

with $A_{\boldsymbol{\xi}} := \int_{SO(3)} \mathcal{L}_m f(\boldsymbol{\alpha}) \mathbf{a}(\boldsymbol{\xi}, \boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha})$ (recalling that the operator $\mathcal{L}_m$ was defined in Lemma 3).

The approximant

$$T_{\boldsymbol{\Xi}} f = \int\limits_{SO(3)} \sum \mathbf{a}(\boldsymbol{\xi}, \boldsymbol{\alpha}) k_m(\cdot, \boldsymbol{\xi}) \mathcal{L}_m f(\boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha})$$

is the function obtained by replacing the kernel in the integral representation by linear combinations of its nearby shifts.

An apparent drawback of this approach is the requirement that f should be in the domain of $\mathcal{L}_m$: this is a restrictive smoothness condition on the target function. To remedy this, we must modify the approach slightly, by splitting the (less smooth) target function into two parts: a component of sufficient smoothness (in L_p) and a rough but negligible component. The approximant to a general function will be obtained by applying the operator $T_{\boldsymbol{\Xi}}$ to its smooth part only. Precisely how this split is accomplished will be set aside for now, and revisited later in this section.

5.1. Functions of full smoothness

We obtain error estimates by bounding the norm of the operator induced by the error kernel, i.e.,

$$E_{k_m} : L_p\big(SO(3)\big) \to L_p\big(SO(3)\big), \qquad E_{k_m} g(\boldsymbol{x}) = \int\limits_{SO(3)} e_{k_m}(\boldsymbol{x}, \boldsymbol{\alpha}) g(\boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha}).$$

The bound on such an operator controls the error of the global approximant, as we will see in the following theorem.

Theorem 9. *Let $m \geqslant 2$, and let $\mathbf{a} : \boldsymbol{\Xi} \times SO(3) \to \mathbb{R}$ be a coefficient kernel satisfying CKCs with radius ρ, precision $L \geqslant 2m$, and stability constant K. For $f \in W_p^{2m}(SO(3))$, $1 \leqslant p < \infty$, the approximant converges, in L_p, to f with error bounded by*

$$\|f - T_{\boldsymbol{\Xi}} f\|_p \leqslant C \rho^{2m} \|f\|_{W_p^{2m}(SO(3))}.$$

Likewise, for $f \in C^{2m}(SO(3))$, the approximant converges in L_∞ with error

$$\|f - T_{\boldsymbol{\Xi}} f\|_\infty \leqslant C \rho^{2m} \|f\|_{C^{2m}(SO(3))}.$$

Proof. For $1 \leqslant p \leqslant \infty$ we can estimate the approximation error by

$$\|f - T_{\boldsymbol{\Xi}} f\|_p = \left\| \int\limits_{SO(3)} \mathcal{L}_m f(\boldsymbol{\alpha}) \left(k_m(\cdot, \boldsymbol{\alpha}) - \sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}} a(\boldsymbol{\xi}, \boldsymbol{\alpha}) k_m(\cdot, \boldsymbol{\xi}) \right) d\mu(\boldsymbol{\alpha}) \right\|_p$$

$$\leqslant \left\| \int\limits_{SO(3)} |\mathcal{L}_m f(\boldsymbol{\alpha})| e_{k_m}(\cdot, \boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha}) \right\|_p \leqslant \|E_{k_m}\|_{p \to p} \|\mathcal{L}_m f\|_p.$$

We need to consider only two values of p, namely $p = 1$ and ∞, because the others follow by interpolation. The L_1 operator norm is controlled by $M_1 := \sup_{\boldsymbol{\alpha} \in SO(3)} \int e_{k_m}(\boldsymbol{x}, \boldsymbol{\alpha}) \, d\mu(\boldsymbol{x})$. Likewise, the L_∞ norm is controlled by $M_\infty := \sup_{\boldsymbol{x} \in SO(3)} \int e_{k_m}(\boldsymbol{x}, \boldsymbol{\alpha}) \, d\mu(\boldsymbol{\alpha})$. Since the estimate from Lemma 8 controls $e_{k,m}$ by a quantity involving $\omega(\boldsymbol{\alpha}^{-1}\boldsymbol{x})$ only, the error kernel is generated by a class function, and we can make a change of variable to see that the result is the same for L_1 and L_∞ (and, hence, all other $1 < p < \infty$ as well):

$$\max\left(M_1, M_\infty\right) \leqslant \int\limits_{SO(3)} C\rho^{2m-3}\left(1 + \frac{\mathrm{dist}(\cdot, \boldsymbol{\alpha})}{\rho}\right)^{2m-4-L} \mathrm{d}\mu(\boldsymbol{\alpha})$$

$$\leqslant C\rho^{2m-3}\int\limits_0^\pi\left(1 + \frac{\omega}{\rho}\right)^{2m-4-L}\sin^2\frac{\omega}{2}\,\mathrm{d}\omega$$

$$\leqslant C\rho^{2m}\int\limits_0^\infty (1+\theta)^{2m-4-L}\theta^2\,\mathrm{d}\theta \leqslant C\rho^{2m}.$$

Interpolating over $1 \leqslant p \leqslant \infty$ we obtain

$$\|E_{k_m}\|_{p\to p} \leqslant C\rho^{2m},$$

and, hence, $\|f - T_{\boldsymbol{\Xi}} f\|_p \leqslant \|E_{k_m}\|_{p\to p}\|\mathcal{L}_m f\|_p \leqslant C\rho^{2m}\|\mathcal{L}_m f\|_p.$ $\quad\square$

Note 2. The size of the coefficients $A_{\boldsymbol{\xi}}$ can be estimated as follows. Given a coefficient kernel **a** satisfying the CKCs with stability constant K, we have

$$\|A\|_{\ell_1(\boldsymbol{\Xi})} \leqslant K\|\mathcal{L}_m f\|_1.$$

Indeed,

$$\sum_{\boldsymbol{\xi}\in\boldsymbol{\Xi}}|A_{\boldsymbol{\xi}}| \leqslant \int\limits_{SO(3)}\sum_{\boldsymbol{\xi}\in\boldsymbol{\Xi}}\big|\mathbf{a}(\boldsymbol{\alpha},\boldsymbol{\xi})\big|\big|\mathcal{L}_m f(\boldsymbol{\alpha})\big|\,\mathrm{d}\mu(\boldsymbol{\alpha}) \leqslant K\int\limits_{SO(3)}\big|\mathcal{L}_m f(\boldsymbol{\alpha})\big|\,\mathrm{d}\mu(\boldsymbol{\alpha}).$$

Applying Lemma 6 to the previous theorem gives the following corollary.

Corollary 10. *For centers $\boldsymbol{\Xi}$ having density $h < h_0$ (with h_0 determined by L as in Lemma 6), there exists a coefficient kernel such that the approximant defined in (25) converges like*

$$\|f - T_{\boldsymbol{\Xi}} f\|_p \leqslant Ch^{2m}\|f\|_{W_p^{2m}(SO(3))}$$

with coefficients satisfying the estimate

$$\|A\|_{\ell_1(\boldsymbol{\Xi})} \leqslant 2\|\mathcal{L}_m f\|_1.$$

5.2. Functions of lower smoothness

To treat more general functions, we make use of a common K-functional argument, one practically ubiquitous in approximation theory. This will permit us to define approximants for functions in Besov spaces of lower order (and ultimately get good error bounds for such functions as well).

Definition 11. For $0 < s \leqslant 2m$ and $1 \leqslant p < \infty$, we define the Besov space $B_{p,\infty}^s(SO(3))$ as the collection of functions in $L_p(SO(3))$ for which the following expression

$$\|f\|_{B_{p,\infty}^s} := \sup_{t>0} t^{-s}K(f, t)$$

is finite, where the K-functional $K(f, \cdot):(0,\infty)\to(0,\infty)$ is defined as

$$K(f,t) := \inf\big\{\|f - g\|_p + t^{2m}\|g\|_{W_p^{2m}}\colon g\in W_p^{2m}\big\}.$$

For $p = \infty$, the definition is the same after substituting $L_p(SO(3))$ by $C(SO(3))$ and W_p^{2m} by C^{2m}.

There are abundant alternative characterizations of Besov spaces. We have chosen the preceding because it captures precisely the properties we need from a function of lower smoothness. It is possible to follow the example of Definition 1, by transporting the Besov norm from $\mathbb{R}^3$. For a general manifold this would not work, it is possible only because $SO(3)$ is compact (specifically, it works because the diffeomorphisms and partitions of unity employed have finite cardinality – see [25, 1.11.3, Remark 2 and 2.4.7] for a discussion of this). We recommend [25, 1.11, and Chapter 7] and the references therein for background on Besov spaces on Riemannian manifolds.

An immediate consequence of the definition is that, for an arbitrary function in $B_{p,\infty}^s$ and $\rho > 0$, there is $g_\rho \in W_p^{2m}$ simultaneously satisfying

$$\|g_\rho\|_{W_p^{2m}} \leqslant 2\rho^{s-2m}\|f\|_{B_{p,\infty}^s}, \tag{26}$$

$$\|f - g_\rho\|_p \leqslant 2\rho^s\|f\|_{B_{p,\infty}^s}. \tag{27}$$

In the setting of Theorem 9 (namely, with radius ρ), the approximant to f is then $s_{f,\Xi} := T_\Xi g_\rho$.

Note 3. It is important to point out that the function g_ρ can be constructed by the same method employed in Section 2.3 to create the C^∞ approximant. Namely, one may employ a C^∞ partition of unity $(\psi_j)_{j=1}^N$, a corresponding sequence of smooth diffeomorphisms, and a sequence of cut-offs $(\tilde{\psi}_j)_{j=1}^N$ to generate an approximant of the form

$$g_\rho = \sum_{j=1}^N \tilde{\psi}_j \times (g_{j,\rho} \circ h_j)$$

where $g_{j,\rho} : \mathbb{R}^3 \to \mathbb{R}$ approximates the function $f_j = (f \times \psi_j) \circ h_j^{-1} \in B_{p,\infty}^s(\mathbb{R}^3)$ in the sense of (26) and (27). The Euclidean approximants $g_{j,\rho}$ can be constructed, for example, by the method [31, Theorem 5.33] or [18, Lemma 5.1].

Theorem 12. *Let $s < 2m$ and $1 \leqslant p \leqslant \infty$. Suppose the coefficient kernel $\mathbf{a} : \Xi \times SO(3) \to \mathbb{R}$ satisfies CKCs with radius ρ, precision $L \geqslant 2m$, and stability constant K. Then for $f \in B_{p,\infty}^s$ there is $s_{f,\Xi}$ satisfying*

$$\|f - s_{f,\Xi}\|_p \leqslant C\rho^s\|f\|_{B_{p,\infty}^s}.$$

Proof. This follows by applying Theorem 9 to g_ρ and using inequalities (26) and (27). □

Note 4. Again, we can estimate the size of the coefficients. For $f \in B_{p,\infty}^s$ we have

$$\|A\|_{\ell_1(\Xi)} \leqslant C\rho^{s-2m}\|f\|_{B_{p,\infty}^s}.$$

Applying Lemma 6 once more, we obtain:

Corollary 13. *For centers Ξ having density $h < h_0$ (with h_0 determined by L as in Lemma 6), and for $f \in B_{p,\infty}^s$, there is an approximant $s_{f,\Xi}$ satisfying*

$$\|f - s_{f,\Xi}\|_p \leqslant Ch^s\|f\|_{B_{p,\infty}^s}$$

with coefficients satisfying the estimate

$$\|A\|_{\ell_1(\Xi)} \leqslant Ch^{s-2m}\|f\|_{B_{p,\infty}^s}.$$

6. Concluding remarks

The results of the last section have value beyond solving a challenging (if theoretical) problem, and even beyond providing benchmarks for other approximation schemes. When approximation schemes based on projection (e.g., interpolation and least-squares projection, discussed in Section 4.2) are *stable*, application of a *Lebesgue lemma* implies that such schemes are *near best*. Very recently, in [32], it has been demonstrated that kernel interpolation on smooth, compact Riemannian manifolds is stable for a certain family of positive definite kernels. Under similar circumstances in [28], least-squares projection has been shown to be stable in L_p for each $1 \leqslant p \leqslant \infty$. In the setting of the 2-sphere, $\mathbb{S}^2$, it is known that these kernels provide precise (if theoretical) L_p rates of approximation. Hence, interpolation on $\mathbb{S}^2$ (an easily implemented approximation scheme) and least-squares approximation inherit these precise rates for all levels of smoothness up to a saturation order of the kernel.

The kernels employed in [32,28] have variants on $SO(3)$ as well, and they are very similar to the kernels considered here. Both families are composed of fundamental solutions of elliptic differential operators which serve as reproducing kernels for the Sobolev spaces $W_2^m(SO(3))$ (albeit for different, but equivalent inner products). This raises the strong likelihood that similar stability results hold for the kernels considered here, or that, by modifying the techniques of this article, precise error estimates can be shown for the kinds of kernels considered in [32,28]. The upshot is that precise error estimates for 'best' approximation may lead directly to precise error estimates for any stable linear approximation scheme obtained by projection.

In addition to this, the approach we take has the potential to work in far greater generality: specifically on other structures like Lie groups and homogeneous spaces. It relies on two basic properties that, at first, seem to have little to do with groups, but rather with (compact) manifolds. These are:

- the representation (12) involving the fundamental solution for a differential operator: iterated, perturbed Laplace–Beltrami operators;
- the ability to efficiently replace this fundamental solution by a linear combination of nearby translates incurring little error (Lemma 8).

Finding kernels that satisfy both of these simultaneously is a challenge; the second property seems especially difficult without some extra structure. However, when the manifold is a group or a homogeneous space, it is often possible to exploit some underlying symmetry. In this article we considered kernels, $\kappa(\boldsymbol{x}, \boldsymbol{\alpha})$, derived from 'class functions', i.e. functions depending on $\boldsymbol{\alpha}^{-1}\boldsymbol{x}$ – meaning $\kappa(\boldsymbol{x}, \boldsymbol{\alpha}) = f(\boldsymbol{\alpha}^{-1}\boldsymbol{x})$ – but possessing an added symmetry; the function f depends only on the rotational angle of $\boldsymbol{\alpha}^{-1}\boldsymbol{x}$ (which on $SO(3)$ is the same as the Riemannian distance). It is precisely this symmetry – ultimately embodied in the addition formula which relates the basic elements of harmonic analysis on the space (the eigenfunctions of the Laplace–Beltrami operator) to an orthonormal basis for the space of class functions, and which, in turn, can be related to certain orthogonal polynomials – that ultimately allowed us to effectively replace the kernel.

Acknowledgment

References

[1] M.D. Buhmann, Radial Basis Functions: Theory and Implementations, Cambridge Monogr. Appl. Comput. Math., vol. 12, Cambridge University Press, Cambridge, 2003.

[2] H. Wendland, Scattered Data Approximation, Cambridge Monogr. Appl. Comput. Math., vol. 17, Cambridge University Press, Cambridge, 2005.

[3] G.S. Chirikjian, A.B. Kyatkin, Engineering Applications of Noncommutative Harmonic Analysis, CRC Press, Boca Raton, 2000.

[4] K.G. van den Boogaart, R. Hielscher, J. Prestin, H. Schaeben, Kernel-based methods for inversion of the Radon transform on $SO(3)$ and their applications to texture analysis, J. Comput. Appl. Math. 199 (1) (2007) 122–140.

[5] H. Schaeben, R. Hielscher, J. Fundenberger, D. Potts, J. Prestin, Orientation density function-controlled pole probability density function measurements: automated adaptive control of texture goniometers, J. Appl. Cryst. 40 (1) (2007) 570–579.

[6] J. Castrillon-Candas, V. Siddavanahalli, C. Bajaj, Nonequispaced Fourier transforms for protein–protein docking, ICES report 05-44, Univ. Texas.

[7] A. Yershova, S. LaValle, Deterministic sampling methods for spheres and $SO(3)$, in: ICRA 04. International Conference on Robotics and Automation, vol. 4, IEEE, New Orleans, 2004, pp. 3974–3980.

[8] M. Gräf, D. Potts, Sampling sets and quadrature formulae on the rotation group, Numer. Funct. Anal. Optim. 30 (7–8) (2009) 665–688.

[9] D. Potts, J. Prestin, A. Vollrath, A fast algorithm for nonequispaced Fourier transforms on the rotation group, Numer. Algorithms 52 (3) (2009) 355–384.

[10] M. Gräf, S. Kunis, Stability results for scattered data interpolation on the rotation group, Electron. Trans. Numer. Anal. 31 (2008) 30–39.

[11] P.J. Kostelec, D.N. Rockmore, FFTs on the rotation group, J. Fourier Anal. Appl. 14 (2) (2008) 145–179.

[12] D. Schmid, Marcinkiewicz–Zygmund inequalities and polynomial approximation from scattered data on $SO(3)$, Numer. Funct. Anal. Optim. 29 (7–8) (2008) 855–882.

[13] T. Gutzmer, Interpolation by positive definite functions on locally compact groups with application to $SO(3)$, Results Math. 29 (1–2) (1996) 69–77.

[14] F. Filbir, D. Schmid, Stability results for approximation by positive definite functions on $SO(3)$, J. Approx. Theory 153 (2) (2008) 170–183.

[15] J. Duchon, Sur l'erreur d'interpolation des fonctions de plusieurs variables par les D^m-splines, RAIRO Anal. Numér. 12 (4) (1978) 325–334.

[16] J. Duchon, Splines minimizing rotation-invariant semi-norms in Sobolev spaces, in: Constructive Theory of Functions of Several Variables, Proc. Conf., Math. Res. Inst., Oberwolfach, 1976, in: Lecture Notes in Math., vol. 571, Springer, Berlin, 1977, pp. 85–100.

[17] T. Hangelbroek, Polyharmonic approximation on the sphere, Constr. Approx., in press, doi:10.1007/s00365-010-9093-8.

[18] R. DeVore, A. Ron, Approximation using scattered shifts of a multivariate function, Trans. Amer. Math. Soc. 362 (12) (2010) 6205–6229.

[19] C. de Boor, R.A. DeVore, A. Ron, Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$, Trans. Amer. Math. Soc. 341 (2) (1994) 787–806.

[20] J. Lei, R.-Q. Jia, E.W. Cheney, Approximation from shift-invariant spaces by integral operators, SIAM J. Math. Anal. 28 (2) (1997) 481–498.

[21] G.C. Kyriazis, Approximation of distribution spaces by means of kernel operators, J. Fourier Anal. Appl. 2 (3) (1996) 261–286.

[22] R.-Q. Jia, Approximation by quasi-projection operators in Besov spaces, J. Approx. Theory 162 (1) (2010) 186–200.

[23] I.M. Gelfand, R.A. Minlos, Z.Y. Shapiro, Representations of the Rotation and Lorentz Groups and Their Applications, Pergamon Press, Oxford, 1963.

[24] J.C. Mitchell, Sampling rotation groups by successive orthogonal images, SIAM J. Sci. Comput. 30 (1) (2007/08) 525–547.

[25] H. Triebel, Theory of Function Spaces. II, Monogr. Math., vol. 84, Birkhäuser Verlag, Basel, 1992.

[26] N. Dyn, F.J. Narcowich, J.D. Ward, Variational principles and Sobolev-type estimates for generalized interpolation on a Riemannian manifold, Constr. Approx. 15 (2) (1999) 175–208.

[27] G. Wahba, Spline Models for Observational Data, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 59, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[28] T. Hangelbroek, F.J. Narcowich, X. Sun, J.D. Ward, Kernel approximation on manifolds II: The L_∞-norm of the L_2-projector, in press, arXiv:1005.2424.

[29] D.A. Varshalovich, A.N. Moskalev, V.K. Khersonskiĭ, Quantum Theory of Angular Momentum, World Scientific Publishing Co. Inc., Teaneck, NJ, 1988, irreducible tensors, spherical harmonics, vector coupling coefficients, $3nj$ symbols, translated from the Russian.

[30] V.S. Videnskiĭ, Extremal estimates for the derivative of a trigonometric polynomial on an interval shorter than its period, Soviet Math. Dokl. 1 (1960) 5–8.

[31] R.A. Adams, J.J.F. Fournier, Sobolev Spaces, 2nd ed., Pure Appl. Math. (Amst.), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[32] T. Hangelbroek, F.J. Narcowich, J.D. Ward, Kernel approximation on manifolds I: Bounding the Lebesgue constant, SIAM J. Math. Anal. 42 (4) (2010) 1732–1760.