Approximation by positive definite functions on compact groups

Wolfgang Erb,* Frank Filbir†

We consider approximation methods defined by translates of a positive definite function on a compact group. A characterization of the native space generated by a positive definite function on a compact group is presented. Starting from Bochner's Theorem, we construct examples of well localized positive definite central functions on the rotation group $\mathsf{SO}(3)$. Finally, the stability of the interpolation problem and the error analysis for the given examples are studied in detail.

AMS Subject Classification(2000): 41A05, 43A35, 41A30, 43A75

Keywords: rotation group, positive definite functions, scattered data approximation.

1 Introduction

Approximating a function f by linear combinations of translates of a given single basis function ϕ is a widely used method. The case where the underlying manifold is $\mathbb{R}^d$ and the basis function is radial symmetric (radial basis function) has been studied in great detail during the last decade (see [18] and references therein). Usually, the setting is as follows. Given a data set

*Centre for Mathematical Sciences, Munich University of Technology, Boltzmannstrasse 3, 85747 Garching, Germany. The research was supported by the DFG through the Graduate Programme Applied Algorithmic Mathematics. erb@ma.tum.de

†Institute for Biomathematics and Biometry, GSF Research Center, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany. filbir@gsf.de

$X \times \mathbb{R} = \{(\boldsymbol{x}_1, f_1), \ldots, (\boldsymbol{x}_n, f_n)\} \subset \mathbb{R}^d \times \mathbb{R}$ and a basis function $\phi : \mathbb{R}^d \to \mathbb{R}$, we suppose the f_i's to be point evaluations of an unknown function f. Now one tries to recover f by a linear combination of translates of a basis function ϕ, i.e., we have an ansatz of the following type:

$$s_f(\boldsymbol{x}) = \sum_{k=1}^{n} a_k \phi(\boldsymbol{x} - \boldsymbol{x}_k), \quad \boldsymbol{x} \in \mathbb{R}^d.$$

Assuming that s_f interpolates the data, leads to a system of linear equations for the coefficients a_k, i.e

$$\mathbf{A}_\phi \boldsymbol{a} = \boldsymbol{f}, \tag{1}$$

where $\mathbf{A}_\phi = (\phi(\boldsymbol{x}_i - \boldsymbol{x}_k))_{i,k=1}^n, \boldsymbol{a} = (a_k)_{k=1}^n$ and $\boldsymbol{f} = (f_k)_{k=1}^n$. It turned out that, in this setting, a positive definite function ϕ is a good choice as a basis function. A function $\phi : \mathbb{R}^d \to \mathbb{R}$ is called positive definite if

$$\sum_{j=1}^{n} \sum_{k=1}^{n} c_j \overline{c_k}\, \phi(\boldsymbol{x}_j - \boldsymbol{x}_k) \geq 0 \tag{2}$$

for all finite sets of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$ and arbitrary coefficients $c_1, \ldots, c_n \in \mathbb{C}$. If the inequality (2) is strict for pairwise distinct $\boldsymbol{x}_j$'s, the function ϕ is called a strictly positive definite function.

Applying such methods, several problems arise naturally.

A. Firstly, one has to clarify which functions f can be approximated by this approach. Every positive definite function ϕ is related to a reproducing kernel Hilbert space and this is the space where the approximation takes place. For this reproducing kernel Hilbert space the name native space was invented by several authors [14], [18]. The problem is that this space is initially defined as the closure of all translates of the given function. Obviously, this characterization of the space is not suitable for approximation purposes. Therefore, one has to identify the space as a subspace of a well known function space.

B. We have to make sure that the linear system (1) can be solved in a stable manner. Even when the system (1) is uniquely solvable, the condition number of the matrix $\boldsymbol{A}_\psi$ can be arbitrarily bad. A careful analysis shows how the condition number of the matrix depends on the one hand on the so-called separation distance of the points defined by

$$q_X = \min_{i \neq j} |\boldsymbol{x}_i - \boldsymbol{x}_j|,$$

and, on the other hand, on the specific properties of the basis function ϕ. In many cases a better stability can be obtained by applying a suitable preconditioning technique [3].

C. The third important question is concerned with the approximation error

$|f(\boldsymbol{x}) - s_f(\boldsymbol{x})|,\ \boldsymbol{x} \in \mathbb{R}^d$. An essential tool for the error analysis is the so-called power function. This function will provide a first estimate, and a closer inspection of this function will finally lead to an estimate in terms of the so-called fill distance

$$h_X = \sup_{\boldsymbol{x} \in \mathbb{R}^d} \inf_{\boldsymbol{x}_j \in X} |\boldsymbol{x} - \boldsymbol{x}_j|.$$

While all these questions are well studied in case where the underlying manifold is the Euclidean space $\mathbb{R}^d$ (or certain subsets of it), the analysis on other manifolds is by no means as well understood. For a comprehensive treatment of the $\mathbb{R}^d$-case we refer to [18].

In various applications we are confronted with the situation where the underlying set is a compact or locally compact group G, i.e. the data set $X \times \mathbb{R}$ is now a subset of $G \times \mathbb{R}$. Problems of such type arise in biochemistry (e.g. protein-protein-docking, force field calculations of macromolecules) in engineering (e.g. robotics) or in physics (e.g. crystallography). The monograph [5] provides a lot more examples. In many situations special matrix groups are involved. Especially the rotation group $\mathsf{SO}(3)$ is one of the most important examples, see [2], [5].

In the case where a locally compact group different from $\mathbb{R}^d$ is involved, a main problem is to come up with suitable positive definite functions. As long as the group can be embedded in the Euclidean space $\mathbb{R}^d$, as in the case of matrix groups, one might try to restrict positive definite functions on $\mathbb{R}^d$ to the manifold defined by the group. Some work in this direction has been done by J. Levesley, D.L .Ragozin [11] and by F.J. Narcowich [12]. In [4] the authors studied properties of positive definite functions on the sphere which are constructed in this way. The drawback of this approach is that it ignores the underlying algebraic structure completely. Especially the powerful harmonic analysis on groups can not be used in order to study the above mentioned problems. Thus, it is more appropriate to work directly on the group. This is the approach we will follow in this paper. A first attempt of applying tools from harmonic analysis on noncommutative locally compact groups to interpolation problems on such structures was made by T. Gutzmer [8]. Recently, a more detailed analysis of the stability problem was made by D. Schmid and the second named author in [6]. The interplay between the approximation error and the stability of the linear system (1) was first observed by R. Schaback [15] and lately extended to a far more general setting by D. Schmid [16].

The paper is organized as follows. After collecting the basics on positive definite functions in the next section, we provide a characterization of the native space in section 3. Section 4 is devoted to the application of the results from section 3 to the rotation group. In section 5 we discuss the stability problem and the error analysis for the approximation problem on $\mathsf{SO}(3)$.

2 Positive definite functions on compact groups

In this section we collect the fundamentals on harmonic analysis and positive definite functions on compact groups as far as they are necessary to understand the remaining part of the paper. We start with the definition of positive definite functions on a general topological group G.

Definition 2.1
A complex valued function ϕ on a topological group G is called positive definite if

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i\bar{c}_j\,\phi(x_i^{-1}x_j) \geq 0 \tag{3}$$

for all $c_1,\dots,c_n \in \mathbb{C},\ x_1,\dots,x_n \in G$ and $n \in \mathbb{N}$. The function ϕ is called strictly positive definite if the inequality (3) *is strict for all possible choices of $c_1,\dots,c_n \in \mathbb{C},\ x_1,\dots,x_n \in G$ and $n \in \mathbb{N}$.*

Note that for the definition of positive definiteness of a function on G it is not necessary to assume that G bears a topology. Since in this paper we are dealing with topological groups exclusively, we included this assumption in the definition above.

We denote by $P(G), CP(G)$ the set of positive definite and continuous positive definite functions respectively. It is easy to see that if $\phi,\psi \in P(G)$ so are $\phi+\psi$, $c\phi$, $\phi\cdot\psi$ for $c \geq 0$. Every positive definite function is hermitean, i.e. $\phi = \widetilde{\phi}$, where $\widetilde{\phi}(x) = \overline{\phi(x^{-1})})$, and, moreover, we have $\phi(e) \geq 0$ and $|\phi(x)| \leq \phi(e)$ for all $x \in G$. For a short proof of these facts see [13].

Let $(\pi, \mathcal{H}_\pi)$ be a unitary representation of G, i.e., π is a homomorphism from G into the group of unitary operators on the Hilbert space $\mathcal{H}_\pi$ which is continuous with respect to the strong operator topology. It is very easy to see that for every $v \in \mathcal{H}_\pi$ the function

$$\phi : G \to \mathbb{C}, \quad \phi(x) = \langle v, \pi(x)v\rangle, \tag{4}$$

is continuous and positive definite on G. Conversely, starting with a continuous positive definite function $\phi : G \to \mathbb{C}$, there is always a unitary representation $(\pi, \mathcal{H}_\pi)$ such that (4) holds. We now describe briefly this construction. Let L_x denote the left translation operator defined by $L_x f(y) = f(x^{-1}y)$. For $\phi \in CP(G)$, we define the linear vector space

$$\mathcal{I}(\phi) = \mathsf{span}\Big\{\sum_{i=1}^{n} a_i L_{x_i}\phi;\ a_1,\dots,a_n \in \mathbb{C},\ x_i \in G\Big\}.$$

For $f = \sum_{i=1}^{n} a_i L_{x_i}\phi$ and $g = \sum_{j=1}^{m} b_i L_{y_j}\phi$, we define the sesquilinear form

$$\langle f, g\rangle_\phi = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\bar{b}_j L_{x_i}\phi(y_j)$$

on $\mathcal{I}(\phi)$. It is easy to see that $\langle \cdot, \cdot \rangle_\phi$ is independent of the representation of the elements f and g and therefore well defined. It can be readily seen that $(\mathcal{I}(\phi)), \langle \cdot, \cdot \rangle_\phi)$ is an inner product space with property

$$\langle g, L_x \phi \rangle_\phi = g(x)$$

for $x \in G$ and $g \in \mathcal{I}(\phi)$. If we take the completion of $\mathcal{I}(\phi)$ with respect to the norm $\| \cdot \|_\phi = \sqrt{\langle \cdot, \cdot \rangle_\phi}$, we get a Hilbert space

$$\mathcal{H}(\phi) := cl_{\| \cdot \|_\phi} \mathcal{I}(\phi).$$

This space has the following properties.

Lemma 2.2

(i) *The translation operator L_x can be extended to $\mathcal{H}(\phi)$. Moreover, $\mathcal{H}(\phi)$ is translation invariant.*

(ii) $\langle f, L_x \phi \rangle_\phi = f(x), \quad x \in G, \ f \in \mathcal{H}(\phi)$.

(iii) *Functions in $\mathcal{H}(\phi)$ are bounded,*

(iv) *For $f_n \in \mathcal{I}(\phi)$ for all $n \in \mathbb{N}$ and $\|f_n - f\|_\phi \to 0$ for $n \to \infty$, then $f_n \to f$ uniformly on G.*

(v) *If $\langle f_n - f, g \rangle_\phi \to 0$ for $n \to \infty$ for all $g \in \mathcal{H}(\phi)$, then $f_n \to f$ pointwise.*

For a proof of this lemma see [13]. The vector space $\mathcal{H}(\phi)$ is thus a reproducing kernel Hilbert space with kernel $K(x, y) = L_x \phi(y)$. This space is called the native space of the positive definite function ϕ by some authors.

Positive definite functions on a commutative locally compact group G can be characterized in terms of the inverse Fourier-Stieltjes transforms of certain measures on the dual of G. This fundamental result is known as Bochner's Theorem. A result of this type is also available for noncommutative, compact groups. We now state the Bochner Theorem for compact groups. In order to do so, we need to summarize some basic facts from harmonic analysis on compact groups. A standard reference for all this material including Bochner's Theorem is [9].

Let G be a compact group with normalized left Haar measure μ. By $\widehat{G}$ we denote the dual object of G, i.e. the set of equivalence classes of irreducible unitary representations of G. For every equivalence class $\sigma \in \widehat{G}$ let $(\pi_\sigma, \mathcal{H}_\sigma)$ be a representative with $\dim \mathcal{H}_\sigma = d_\sigma$. The Fourier transform of a function $f \in L^1(G)$ at a point $\sigma \in \widehat{G}$ is an operator on $\mathcal{H}_\sigma$ defined by

$$\hat{f}_\sigma = \int_G f(x) \pi^*_\sigma(x) d\mu(x).$$

The set $(\hat{f}_\sigma)_{\sigma \in \widehat{G}}$ constitutes an operator-valued sequence with index set $\widehat{G}$.

Due to the Peter-Weyl Theorem, we have

$$f = \sum_{\sigma \in \widehat{G}} d_\sigma \mathsf{tr}(\hat{f}_\sigma \pi_\sigma)$$

for $f \in L^2(G)$, where the sum converges in the topology of $L^2(G)$. Furthermore, the Parseval equality

$$\|f\|_2^2 = \sum_{\sigma \in \widehat{G}} d_\sigma \mathsf{tr}\big(\hat{f}_\sigma^* \hat{f}_\sigma\big)$$

holds.

Following [9], we define for operators $A_\sigma \in B(\mathcal{H}_\sigma)$ the expression

$$\|A_\sigma\|_{\varphi_p} := \begin{cases} \Big(\sum_{i=1}^{d_\sigma} \lambda_i^p\Big)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \max\big\{\lambda_1, \lambda_2, \dots, \lambda_{d_\sigma}\big\}, & p = \infty, \end{cases}$$

where $\lambda_1, \lambda_2, \dots, \lambda_{d_\sigma}$ are the eigenvalues of $|A_\sigma| = \sqrt{A_\sigma^* A_\sigma}$. We now define the space of operator-valued sequences on the dual object by

$$c(\widehat{G}) := \prod_{\sigma \in \widehat{G}} B(\mathcal{H}_\sigma).$$

Obviously, $c(\widehat{G})$ is a $*$-algebra with pointwise defined addition, scalar multiplication, multiplication and the adjoint of operators as involution.

The spaces

$$l^p(\widehat{G}) := \Big\{A \in c(\widehat{G}) : \sum_{\sigma \in \widehat{G}} d_\sigma \|A_\sigma\|_{\varphi_p}^p < \infty\Big\}, \qquad 1 \le p \le \infty, \tag{5}$$

are Banach spaces with respect to the norm

$$\|A\|_p := \begin{cases} \Big(\sum_{\sigma \in \widehat{G}} d_\sigma \|A_\sigma\|_{\varphi_p}^p\Big)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \sup \big\|A_\sigma\|_\sigma, & p = \infty. \end{cases}$$

Moreover, the space $l^2(\widehat{G})$ is a Hilbert space with respect to

$$\langle A, B\rangle = \sum_{\sigma \in \widehat{G}} d_\sigma \mathsf{tr}\big(A_\sigma B_\sigma^*\big).$$

We are now able to state the Theorem of Bochner which provides a characterization of positive definite functions on a compact group. For the proof of this theorem we refer to [9, p.334].

Theorem 2.3
A function ϕ on the compact group G is continuous and positive definite if and only if there is a unique sequence $(A_\sigma)_{\sigma\in\widehat{G}} \in l^1(\widehat{G})$ of hermitian positive semidefinite operators $A_\sigma \in B(\mathcal{H}_\sigma)$ such that

$$\phi(x) = \sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}\big(A_\sigma \pi_\sigma(x)\big).$$

Note that an operator $A_\sigma \in B(\mathcal{H}_\sigma)$ is called positive semidefinite or positive definite if, accordingly, $\langle A_\sigma v, v\rangle_{\mathcal{H}_\sigma} \geq 0$ or $\langle A_\sigma v, v\rangle_{\mathcal{H}_\sigma} > 0$ for all $v \in \mathcal{H}_\sigma \setminus \{0\}$.

Since we always assume the positive definite function ϕ to be continuous, the Peter-Weyl Theorem implies $A_\sigma = \hat{\phi}_\sigma$.

Sometimes it is necessary to work with strictly positive definite functions. The characterization of those functions is a very difficult problem. We refer to [1] for a sufficient condition.

3 Characterization of the native space

Let G be a compact group and let $X = \{x_1, \ldots, x_n\}$ be a set of distinct points on G. In this section we are going to apply Bochner's Theorem in order to obtain a characterization of the native space $\mathcal{H}(\phi)$ for a given function $\phi \in CP(G)$. The following proposition shows that the native space $\mathcal{H}(\phi)$ is uniquely defined.

Proposition 3.1
Let $\phi \in CP(G)$ and H be a Hilbert space of functions $f : G \to \mathbb{C}$ with reproducing kernel $K(x,y) = L_x\phi(y)$. Then H coincides with $\mathcal{H}(\phi)$ and the inner products are the same.

Proof. The line of argumentation is the same as in the proof of Theorem 10.11 in [18]. For $f = \sum_{i=1}^n a_i L_{x_i}\phi \in \mathcal{I}(\phi)$, we have

$$\|f\|_H^2 = \sum_{i,j=1}^n a_i \bar{a}_j \langle L_{x_i}\phi, L_{x_j}\phi\rangle = \sum_{i,j=1}^n a_i \bar{a}_j L_{x_i}\phi(x_j) = \|f\|_\phi^2. \tag{6}$$

This shows $\mathcal{I}(\phi) \subset H$. Now let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{I}(\phi)$ converging to $f \in \mathcal{H}(\phi)$. By Lemma 2.2 (iv) $(f_n)_{n\in\mathbb{N}}$ converges pointwise, too. Due to (6), the sequence $(f_n)_{n\in\mathbb{N}}$ converges to an element $\widetilde{f} \in H$ in norm as well as pointwise. This gives us $f = \widetilde{f}$ and, consequently, $\mathcal{H}(\phi) \subseteq H$. Assume now that H does not coincide with $\mathcal{H}(\phi)$. Then we can find a nonzero element $g \in H$

orthogonal to the closed subspace $\mathcal{H}(\phi)$. But this implies $g(x) = \langle g, L_x\phi\rangle = 0$ for all $x \in G$, which is a contradiction. The equality of the inner products follows easily from the polarization identity. ■

For $\phi \in CP(G)$, let $\hat{\phi} = (\hat{\phi}_\sigma)_{\sigma\in\widehat{G}} \in l^1(\widehat{G})$ be the sequence of Fourier coefficients. By Theorem 2.3, all operators $\hat{\phi}_\sigma$ are positive definite and hermitian. This means that the square root $\hat{\phi}_\sigma^{\frac{1}{2}}$ is well defined. By $\hat{\phi}_\sigma^\dagger$ we denote the Moore-Penrose pseudo inverse of $\hat{\phi}_\sigma$. Since this operator is hermitian and positive definite as well, $\hat{\phi}^{\dagger\frac{1}{2}}$ is also well defined. Moreover, we have

$$\mathsf{Ker}(\hat{\phi}_\sigma^{\dagger\frac{1}{2}}) = \mathsf{Ker}(\hat{\phi}_\sigma), \quad \hat{\phi}_\sigma\, \hat{\phi}_\sigma^\dagger\, \hat{\phi}_\sigma = \hat{\phi}_\sigma, \quad \hat{\phi}_\sigma^\dagger\, \hat{\phi}_\sigma\, \hat{\phi}_\sigma^\dagger = \hat{\phi}_\sigma^\dagger.$$

Now we introduce the function space

$$H(\phi) = \Big\{ f \in C(G):\ \mathsf{Im}(\hat{f}_\sigma) \subseteq \mathsf{Im}(\hat{\phi}_\sigma),\ \big(\hat{\phi}_\sigma^{\dagger\frac{1}{2}}\, \hat{f}_\sigma\big)_{\sigma\in\widehat{G}} \in l^2(\widehat{G}) \Big\},$$

where Im denotes the image of the operator. As we will prove now, $H(\phi)$ characterizes the native space $\mathcal{H}(\phi)$.

Theorem 3.2
The function space $H(\phi)$ equipped with the bilinear form

$$\langle f, g\rangle_H = \sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}\big(\hat{\phi}_\sigma^\dagger\, \hat{f}_\sigma\, \hat{g}_\sigma^*\big) \tag{7}$$

is a reproducing kernel Hilbert space with reproducing kernel $K(x, y) = L_x\phi(y)$. Moreover, $H(\phi) = \mathcal{H}(\phi)$, and both inner products coincide.

Proof. The bilinear form (7) is obviously linear with respect to the first argument. Moreover, the properties of the trace operator and the fact that $\phi_\sigma^\dagger$ is hermitian gives us $\langle f, g\rangle_H = \overline{\langle g, f\rangle_H}$. The positive definiteness of $\langle\cdot,\cdot\rangle_H$ follows from the fact that $\mathsf{Im}(\hat{f}_\sigma) \subseteq \mathsf{Im}(\hat{\phi}_\sigma) = \mathsf{Ker}(\hat{\phi}_\sigma)^\perp$ and $\mathsf{Ker}(\hat{\phi}_\sigma) = \mathsf{Ker}(\hat{\phi}_\sigma^{\dagger\frac{1}{2}})$ for all $\sigma \in \widehat{G}$.

Now, let $(g_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $H(\phi)$. By definition of $\langle\cdot,\cdot\rangle_H$, this means that $(\hat{\phi}^{\dagger\frac{1}{2}}\hat{g}_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the space $l^2(\widehat{G})$. Consequently, it converges to a sequence $A \in l^2(\widehat{G})$. Moreover, the sequence $\hat{\phi}^{\frac{1}{2}}A$ lies in $l^1(\widehat{G})$ because

$$\sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}\big(|\hat{\phi}_\sigma^{\frac{1}{2}} A_\sigma|\big) \le \Big(\sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}(\hat{\phi}_\sigma)\Big)^{\frac{1}{2}} \Big(\sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}(A_\sigma A_\sigma^*)\Big)^{\frac{1}{2}} = \|\hat{\phi}^{\frac{1}{2}}\|_2^2 \|A\|_2^2 < \infty$$

by the Hölder's inequality for $l^p(\widehat{G})$ (see [9]). Hence, the function

$$g(x) := \sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}\big(\hat{\phi}_\sigma^{\frac{1}{2}} A_\sigma \pi_\sigma(x)\big)$$

is well defined, continuous and we have $\hat{g}_\sigma = \hat{\phi}_\sigma^{\frac{1}{2}} A_\sigma,\ \sigma \in \widehat{G}$. Furthermore, we have $\mathsf{Im}(A_\sigma) \subseteq \mathsf{Im}(\hat{\phi}_\sigma)$ for all $\sigma \in \widehat{G}$. Consequently, $\hat{\phi}_\sigma^{\dagger\frac{1}{2}}\, \hat{g}_\sigma\, v = \hat{\phi}_\sigma^{\dagger\frac{1}{2}}\, \hat{\phi}_\sigma^{\frac{1}{2}}\, A_\sigma\, v = A_\sigma\, v$ for all $v \in \mathcal{H}_\sigma,\ \sigma \in \widehat{G}$. Thus, we obtain $\hat{\phi}^{\dagger\frac{1}{2}}\, \hat{g} = A$. Moreover, $\mathsf{Im}(\hat{g}_\sigma) = \mathsf{Im}(\hat{\phi}_\sigma^{\frac{1}{2}} A_\sigma) \subseteq \mathsf{Im}(\hat{\phi}_\sigma)$ which finally gives $g \in H(\phi)$. Now we obtain

$$\|g - g_n\|_H = \|\hat{\phi}^{\dagger\frac{1}{2}}(\hat{g} - \hat{g}_n)\|_2 = \|A - \hat{\phi}^{\dagger\frac{1}{2}}\hat{g}_n\|_2 \overset{n\to\infty}{\longrightarrow} 0.$$

This yields the completeness of $H(\phi)$.
It remains to show that $K(x,y) = L_x\phi(y)$ is the reproducing kernel of $H(\phi)$. The assertion then follows from Proposition 3.1. At first, we see that $H(\phi)$ contains ϕ. Indeed, since

$$\|\hat{\phi}^{\dagger\frac{1}{2}}\hat{\phi}\|_2 = \Big(\sum_{\sigma\in\widehat{G}} d_\sigma \|\hat{\phi}_\sigma^{\dagger\frac{1}{2}}\, \hat{\phi}_\sigma\|_{\varphi_2}^2\Big)^{\frac{1}{2}} = \Big(\sum_{\sigma\in\widehat{G}} d_\sigma \|\hat{\phi}_\sigma^{\dagger\frac{1}{2}}\, \hat{\phi}_\sigma^{\frac{1}{2}}\, \hat{\phi}_\sigma^{\frac{1}{2}}\|_{\varphi_2}^2\Big)^{\frac{1}{2}}$$

and $\mathsf{Im}(\hat{\phi}_\sigma^{\frac{1}{2}}) \subseteq \mathsf{Im}(\hat{\phi}_\sigma)$ for all $\sigma \in \widehat{G}$ we obtain $\|\hat{\phi}^{\dagger\frac{1}{2}}\hat{\phi}\|_2 = \|\hat{\phi}\|_1$ from which the claim follows. Moreover, from $\widehat{(L_x\phi)}_\sigma = \hat{\phi}_\sigma\, \pi_\sigma^*(x)$ for all $\sigma \in \widehat{G}$ and the invariance of $\|\cdot\|_2$ with respect to a sequence of unitary operators we see that $L_x\phi \in H(\phi)$ for every $x \in G$. Finally, the reproduction property follows from

$$\begin{aligned}\langle g, L_x\phi\rangle_H &= \sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}\big(\hat{\phi}_\sigma^\dagger\, \hat{g}_\sigma \big(\hat{\phi}_\sigma \pi_\sigma^*(x)\big)^*\big) \\ &= \sum_{\sigma\in\widehat{G}} d_\sigma \mathsf{tr}\big(\hat{g}_\sigma \pi_\sigma(x)\big) = g(x).\end{aligned}$$

■

The elements of the space $\mathcal{H}(\phi)$ can therefore be regarded as continuous functions whose Fourier transform lies in a weighted $l^2(\widehat{G})$ space where the weights are given by the sequence of the pseudo inverse $\hat{\phi}^\dagger$.

4 Application to the rotation group

We now are going to apply the results of the previous section to the group of rotations in the Euclidean space $\mathbb{R}^3$, i.e.

$$\mathsf{SO}(3) = \{x \in \mathsf{GL}(3,\mathbb{R}) :\ x^T x = e,\ \det x = 1\}.$$

In order to do so, we collect briefly some fundamental facts about the group $\mathsf{SO}(3)$. The rotation group is a compact semisimple Lie group and it can be parameterized in different ways. Most suitable for us is the parameterization on the projective space. This is given as follows: Let K_π be the closed ball with radius π in $\mathbb{R}^3$ and identify antipodal points on the surface. This is the three dimensional real projective space. An element $x \in \mathsf{SO}(3)$ is determined by its rotation axis given by $x \cdot \mathbf{r} = \mathbf{r}, \|\mathbf{r}\| = 1$, and its rotation angle $\alpha(x) \in [0, \pi]$ defined by $\langle \mathbf{v}, x \cdot \mathbf{v} \rangle = \cos \alpha(x)$, where $\mathbf{v} \in \{\mathbf{r}\}^{\perp}, \ \|\mathbf{v}\| = 1$. The element $x \in \mathsf{SO}(3)$ is now identified with a point in the projective space K_π by $x \to \alpha(x)\mathbf{r}$.

A translation invariant metric on $\mathsf{SO}(3)$ is given by

$$d(x, y) := \alpha(y^{-1}x).$$

The easiest way to prove this statement is to look at the double cover of K_π which can be identified with the three dimensional unit sphere $S^3 \subset \mathbb{R}^4$. There, the rotation angle corresponds exactly with the geodesic distance.

For $l \in \mathbb{N}$ and $-l \leq n \leq l$, let Y_n^l denote the canonical orthonormal basis of spherical harmonics on the space of square integrable functions on the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ and $\mathcal{H}_l := \mathsf{span}\{Y_n^l : -l \leq n \leq l\}$. Now define

$$D^l_{m,n}(x) = \int_{\mathbb{S}^2} Y^l_m(xs)\overline{Y^l_n(s)}ds,$$

and let

$$D^l(x) = \Big(D^l_{m,n}(x)\Big)^l_{m,n=-l}, \quad l \in \mathbb{N}.$$

It is well known that the D^l, $l \in \mathbb{N}$, form a complete set of unitary irreducible representations of the rotation group. Thus, the dual object $\mathsf{SO}(3)\hat{}$ can be identified with $\mathbb{N}$ and, due to the Peter-Weyl theorem, the matrix elements $D^l_{m,n}$ constitute an orthonormal basis of $L^2(\mathsf{SO}(3))$. Bochner's theorem now reads as follows.

Theorem 4.1
A function ϕ on $\mathsf{SO}(3)$ is continuous and positive definite if and only if there is an unique sequence $(A_l)_{l\in\mathbb{N}} \in l^1(\mathsf{SO}(3)\hat{})$ of hermitian positive semidefinite operators $A_l \in B(\mathcal{H}_l)$ such that

$$\phi(x) = \sum_{l\in\mathbb{N}} (2l+1)\mathsf{tr}\big(A_l D_l(x)\big).$$

By the Peter-Weyl theorem, we have

$$A_l = \hat{\phi}_l = \int_{\mathsf{SO}(3)} \phi(x) D^l(x^{-1}) d\mu(x),$$

where μ denotes the normalized translation invariant Haar measure on $\mathsf{SO}(3)$.

Bochner's theorem states that in order to construct positive definite functions on $\mathsf{SO}(3)$ we need appropriate Fourier coefficients $(A_l)_{l\in\mathbb{N}}$. To simplify the selection, it is reasonable to look at coefficients of the form $A_l = a_l \mathsf{id}_{2l+1}$ where the a_l are complex numbers.

A function on a group is called central function if it is constant on conjugacy classes. It turns out that functions with Fourier coefficients of the form $A_l = a_l \mathsf{id}_{2l+1}$ are central functions on $\mathsf{SO}(3)$. It can be shown that central functions on $\mathsf{SO}(3)$ depend on the rotation angle only, see [7]. Due to this symmetry, central functions can be seen as an analog of the radial functions in $\mathbb{R}^d$. Fundamental central functions on $\mathsf{SO}(3)$ are the characters $\mathsf{tr}(D^l(x))$. We have

$$\mathsf{tr}\big(D^l(x)\big) = \frac{\sin\big((2l+1)\frac{\alpha(x)}{2}\big)}{\sin\big(\frac{\alpha(x)}{2}\big)} = U_{2l}\big(\cos\big(\frac{\alpha(x)}{2}\big)\big),$$

where U_l is the Chebyshev polynomial of second kind of degree l. For more details we refer to the monographs [7] and [17]. As an immediate consequence of Theorem 4.1, we now obtain

Corollary 4.2
A central function ϕ on $\mathsf{SO}(3)$ is continuous and positive definite if and only if there is an unique sequence $(a_l)_{l\in\mathbb{N}}$ of nonnegative numbers with the property $\sum_{l\in\mathbb{N}}(2l+1)^2 a_l < \infty$ such that

$$\phi(x) = \sum_{l\in\mathbb{N}} (2l+1) a_l U_{2l}\big(\cos\big(\frac{\alpha(x)}{2}\big)\big).$$

We are now prepared to come up with some examples of positive definite central functions on the rotation group. For the construction we will apply Corollary 4.2.

Examples

(a) We start from the well known formula

$$\frac{1}{1-2r\cos\alpha + r^2} = \sum_{l=0}^{\infty} r^l U_l(\cos\alpha),$$

where $0 < r < 1$ and $\alpha \in [0,\pi]$. This equation implies

$$\sum_{l=0}^{\infty} r^l U_{2l}\big(\cos\frac{\alpha}{2}\big) = \sum_{l=1}^{\infty} r^l \big(U_l(\cos\alpha) + U_{l-1}(\cos\alpha)\big) + 1$$
$$= (1+r)\sum_{l=0}^{\infty} r^l U_l(\cos\alpha) = \frac{1+r}{1-2r\cos\alpha + r^2}.$$

Thus, by Corollary 4.2, the functions

$$P_r(x) = \frac{(1-r)^2}{1-2r\cos\alpha(x) + r^2}$$

are positive definite and continuous on $\mathsf{SO}(3)$. Note that the function is scaled such that $P_r(e) = 1$. The Fourier coefficients are then given by

$$\widehat{P_r}(l) = \frac{(1-r)^2}{1+r}\frac{r^l}{2l+1}\mathsf{id}_{2l+1}, \quad l \in \mathbb{N}.$$

By Theorem 3.2, the native space of the functions P_r is given by

$$\mathcal{H}(P_r) = \left\{ g \in C(\mathsf{SO}(3));\ \sum_{l=0}^{\infty} \frac{(2l+1)^2}{r^l}\mathsf{tr}(\hat{g}_l \hat{g}_l^*) < \infty \right\}$$

with the inner product

$$\langle g, h\rangle_{P_r} = \frac{1+r}{(1-r)^2}\sum_{l=0}^{\infty}\frac{(2l+1)^2}{r^l}\mathsf{tr}(\hat{g}_l \hat{h}_l^*).$$

The scaling parameter r determines the localization of P_r around the identity element. In fact, as r tends to one the tighter the peak of P_r at the point e gets.

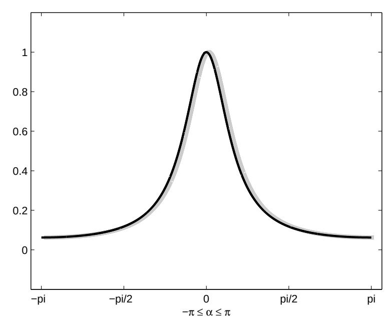

Fig.1 $P_r(\alpha)$, $r = 0.6$

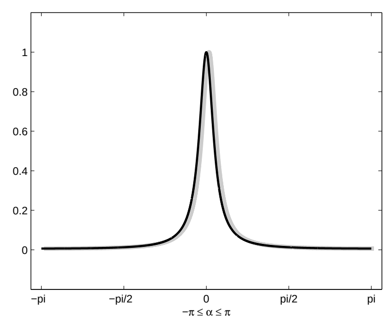

Fig.2 $P_r(\alpha)$, $r = 0.85$

(b) Let $m \in \mathbb{N}$ and define

$$S_m(x) = \left(\frac{1 + 2\cos\alpha(x)}{3}\right)^m, \quad m \in \mathbb{N}.$$

Since $S_1(x) = \frac{1}{3}U_2(\cos(\alpha(x)/2))$, the function S_1 is positive definite on $\mathsf{SO}(3)$. Therefore, S_m is positive definite, too. Moreover, thanks to $S_1(x) < S_1(e)$ for all $x \neq e$, we can localize the function around e as much as we like by taking suitable powers. Using the product formula for the Chebyshev polynomials

$$U_l(x)U_k(x) = \sum_{j=|k-l|}^{k+l} U_j(x), \quad k, l \in \mathbb{N}_0, \tag{8}$$

we get for $k=2$ and $l \geq 2$ the formula

$$U_l(x)U_2(x) = U_{l-1} + U_l + U_{l+2}.$$

Thus, by a simple induction over m one can show that

$$S_m(x) = \sum_{l=0}^{m} \frac{n_l}{3^m} U_{2l}\big(\cos\frac{\alpha(x)}{2}\big),$$

where n_l are strictly positive integers which can be computed recursively.

The native space in this case is finite dimensional and can be written as

$$\mathcal{H}(S_m) = \mathsf{span}\{D^l_{\mu,\nu};\ 0 \leq l \leq m,\ -l \leq \mu, \nu \leq l\}$$

with the inner product

$$\langle g, h\rangle_{S_m} = \sum_{l=0}^{m} \frac{3^m}{n_l} \mathsf{tr}(\hat{g}_l\, \hat{h}_l^*).$$

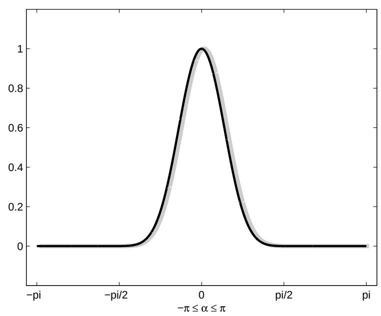

Fig.3 $S_m(\alpha),\ m = 8$

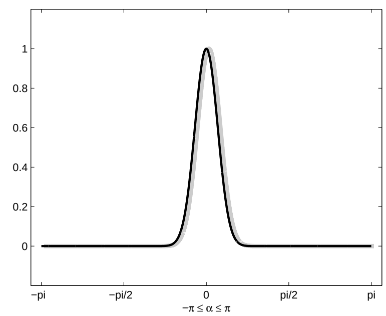

Fig.4 $S_m(\alpha),\ m = 32$

(c) The next example is given by the fundamental solution of the diffusion equation on $\mathsf{SO}(3)$. The eigenvalues of the Laplacian on $\mathsf{SO}(3)$ are $-l(l+1), l \in \mathbb{N}$. The diffusion kernel is then given by

$$G_t(x) = \sum_{l=0}^{\infty} (2l+1) e^{-l(l+1)t} U_{2l}\big(\cos\frac{\alpha(x)}{2}\big), \quad t \geq 0. \tag{9}$$

The kernel can not be expressed in closed form, but, due to the fast decay of the coefficients, it can be well approximated by truncation of the series. The native space for G_t is given by

$$\mathcal{H}(G_t) = \Big\{g \in \mathsf{SO}(3) : \sum_{l=0}^{\infty} (2l+1) e^{l(l+1)t} \mathsf{tr}(\hat{g}_l\, \hat{g}_l^*)\Big\}.$$

The scalar product reads as

$$\langle g, h\rangle_{G_t} = \sum_{l=0}^{\infty} (2l+1) e^{l(l+1)t} \mathsf{tr}(\hat{g}_l\, \hat{h}_l^*).$$

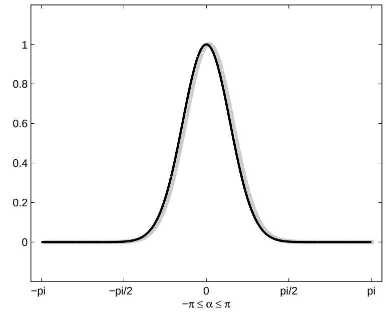

Fig.5 $G_t(\alpha)/G_t(0)$, $t = 0.1$

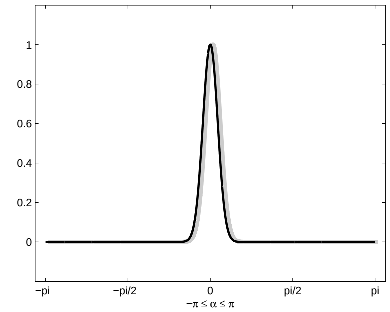

Fig.6 $G_t(\alpha)/G_t(0)$, $t = 0.01$

(d) Another way of getting a family of positive definite functions is to start from a known convergent sine series. As an example, we take the absolutely convergent series

$$\sum_{l=1}^{\infty} \frac{p^l}{l!} \sin(l\alpha) = e^{p\cos(\alpha)} \sin(p\sin(\alpha)), \quad \alpha \in [0, \pi],$$

where $p > 0$. To come up with a series of Chebyshev polynomials of even degree, we transform the latter formula appropriately. We have

$$\begin{aligned} e^{p\cos(\alpha)} \sin(p\sin(\alpha)) \cos(\frac{\alpha}{2}) &= \sum_{l=1}^{\infty} \frac{p^l}{l!} \sin(l\alpha) \cos(\frac{\alpha}{2}) \\ &= \frac{1}{2} \sum_{l=1}^{\infty} \frac{p^l}{l!} \Big(\sin\big((2l+1)\frac{\alpha}{2}\big) + \sin\big((2l-1)\frac{\alpha}{2}\big) \Big) \\ &= \frac{1}{2} \sum_{l=1}^{\infty} \Big(\frac{p^l}{l!} + \frac{p^{l+1}}{(l+1)!} \Big) \sin\big((2l+1)\frac{\alpha}{2}\big) + \frac{p}{2} \sin\big(\frac{\alpha}{2}\big) \\ &= \frac{1}{2} \sum_{l=1}^{\infty} \frac{p^l}{(l+1)!} (p+l+1) \sin\big((2l+1)\frac{\alpha}{2}\big) + \frac{p}{2} \sin\big(\frac{\alpha}{2}\big). \end{aligned}$$

Dividing now by $\sin(\frac{\alpha}{2})$, we get

$$\begin{aligned} e^{p\cos(\alpha)} \frac{\sin(p\sin(\alpha))}{\tan(\frac{\alpha}{2})} &= \frac{1}{2} \sum_{l=1}^{\infty} \frac{p^l}{(l+1)!} (p+l+1) \frac{\sin\big((2l+1)\frac{\alpha}{2}\big)}{\sin\big(\frac{\alpha}{2}\big)} + \frac{p}{2} \\ &= \frac{1}{2} \sum_{l=1}^{\infty} \frac{p^l}{(l+1)!} (p+l+1) U_{2l}\Big(\cos\big(\frac{\alpha}{2}\big)\Big) + \frac{p}{2}. \end{aligned}$$

Finally, scaling by $\frac{1}{2e^p p}$, we have

$$E_p(x) = e^{-p(1-\cos\alpha(x))} \frac{\sin(p\sin\alpha(x))}{2p\tan(\frac{\alpha(x)}{2})}, \quad p > 0.$$

Thus, from the computations above we obtain

$$\mathcal{H}(E_p) = \left\{ g \in C(\mathsf{SO}(3));\ \sum_{l=1}^{\infty} \frac{(2l+1)^2 (l+1)!}{p^{l-1}(p+l+1)} \mathsf{tr}(\hat{g}_l \hat{g}_l^*) < \infty \right\}$$

with the inner product

$$\langle g, h \rangle_{E_p} = 4e^p \hat{g}(0)\overline{\hat{h}(0)} + \sum_{l=1}^{\infty} \frac{(2l+1)^2 4e^p (l+1)!}{p^{l-1}(p+l+1)} \mathsf{tr}(\hat{g}_l \hat{h}_l^*).$$

Similarly to S_m, the larger the value p is chosen the better the function E_p localizes at the identity e.

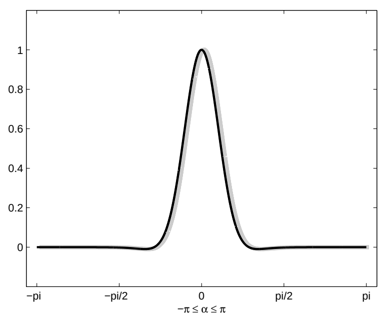

Fig.7 $E_p(\alpha)$, $p = 4$

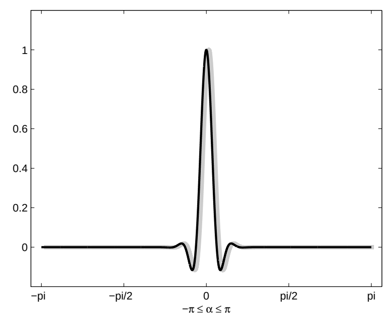

Fig.8 $E_p(\alpha)$, $p = 16$

5 Approximation problems

We now turn to the problem of approximating a function from the native space by a linear combination of the corresponding positive definite function. Let us briefly describe the problem. Assume $\phi \in CP(G)$ and $X \times \mathbb{C} = \{(x_1, \xi), \ldots, (x_n, \xi_n)\} \subset G \times \mathbb{C}$ to be a given data set. Let us further assume that $\xi_j = g(x_j)$, where $g \in \mathcal{H}(\phi)$. Now we would like to approximate g by

$$s_g(x) = \sum_{j=1}^{n} a_j\, \phi(x_j^{-1} x), \tag{10}$$

where the coefficients a_j are determined according to $s_g(x_j) = \xi_j$. This leads to the linear system

$$\boldsymbol{A}_\phi\, \boldsymbol{a} = \boldsymbol{\xi}, \tag{11}$$

where $\boldsymbol{A}_\phi = (\phi(x_i^{-1} x_j))_{i,j=1}^n, \boldsymbol{a} = (a_1, \ldots, a_n)^t,\ \boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)^t$. The matrix $\boldsymbol{A}_\phi$ is hermitian and positive semidefinite. Now, one has to address at least two problems. Firstly, we have to make sure that (10) can be solved in a stable manner. Therefore, it is important to have good estimates for the condition number

$$\kappa(\boldsymbol{A}_\phi) = \frac{\lambda_{max}}{\lambda_{min}}. \tag{12}$$

Here $\lambda_{min}, \lambda_{max}$ denotes the smallest and the greatest eigenvalue of $\boldsymbol{A}_\phi$ respectively. Secondly, we need estimates for the approximation error $|g(x) - s_g(x)|$. In order to come up with error estimates, it is useful to represent the approximating function s_g in a different way. This representation is given by

$$s_g(x) = \sum_{i=1}^{n} u_i(x) g(x_i). \tag{13}$$

The functions u_i are defined pointwise by

$$\boldsymbol{A}_\phi \, \boldsymbol{u}(x) = \boldsymbol{\phi}(x), \tag{14}$$

where $\boldsymbol{u}(x) = (u_1(x), \ldots, u_n(x))^T$ and $\boldsymbol{\phi}(x) = (\phi(x_1^{-1}x), \ldots, \phi(x_n^{-1}x))^T$. Using the fact that $\mathcal{H}(\phi)$ is a reproducing kernel Hilbert space with kernel $L_x\phi$, we obtain

$$\begin{aligned} |g(x) - s_g(x)| &= |g(x) - \sum_{i=1}^{n} u_i(x) g(x_i)| = |\langle g, L_x\phi - \sum_{i=1}^{n} u_i(x) L_{x_i}\phi \rangle_\phi| \\ &\leq \|g\|_\phi \|L_x\phi - \sum_{i=1}^{n} u_i(x) L_{x_i}\phi\|_\phi. \end{aligned} \tag{15}$$

Usually, the function $P_{X,\phi}(x) = \|L_x\phi - \sum_{i=1}^{n} u_i(x) L_{x_i}\phi\|_\phi$ is referred to as power function, see [18]. The power function depends on the point x, on the set X and on ϕ, but does not depend on g. Using the reproducing kernel property, we obtain

$$P_{X,\phi}^2(x) = \phi(e) - 2\sum_{j=1}^{n} u_j(x)\phi(x_j^{-1}x) + \sum_{j=1}^{n}\sum_{k=1}^{n} u_j(x) u_k(x) \phi(x_j^{-1}x_k). \tag{16}$$

It can be shown that the quadratic form (16) is minimized pointwise by the vector $\boldsymbol{u}(x)$.

5.1 Stability

The estimate of the condition number $\kappa(\boldsymbol{A}_\phi)$ is strongly related to the separation distance of the set X. This parameter is defined as

$$q_X := \min\{d(x_i, x_j) : \, 1 \leq i, j \leq n, \, i \neq j\}, \tag{17}$$

where $d(x,y) = \alpha(y^{-1}x)$. Let $\phi \in CP(\mathsf{SO}(3))$ be a central positive definite function with Fourier coefficients $\hat{\phi}_l = a_l \mathsf{id}_{2l+1}$. Now, let

$$a_{\min}(M) := \min_{l \le 2M} \{a_l\}, \quad M \in \mathbb{N}.$$

Then we have the following result.

Theorem 5.1
Let $\phi \in CP(\mathsf{SO}(3))$ be a central positive definite function on $\mathsf{SO}(3)$. Then we have

$$\lambda_{\min}(\boldsymbol{A}_\phi) \ge \frac{1}{2} a_{\min}(M)(M+1)^3,$$

for any M satisfying

$$M + 1 \ge \frac{14.65}{q_X}.$$

Proof. We consider the polynomial kernel

$$F_M(x) = \sum_{l=0}^{M} U_{2l}\Big(\cos\frac{\alpha(x)}{2}\Big) = \frac{\sin\big((M+1)\frac{\alpha}{2}\big)^2}{\sin(\frac{\alpha}{2})^2}. \tag{18}$$

and define the auxiliary function $G_M := \frac{a_{\min}(M)}{(M+1)} F_M^2$. We have

$$G_M(e) = \frac{a_{\min}(M)}{(M+1)} F_M^2(e) = a_{\min}(M)(M+1)^3. \tag{19}$$

For hermitian positive definite operators $A_l, B_l \in \mathcal{H}_l$ we write $A_l \le B_l$ if $B_l - A_l$ is a positive definite operator. We are now showing that the Fourier coefficients of G_M are dominated by the Fourier coefficients of ϕ in the sense that $(\widehat{G_M})_l \le \hat{\phi}_l$ for each $l \in \mathbb{N}$. We apply the product formula (8) of the Chebyshev polynomials to get

$$\begin{aligned}
F_M^2 &= \Big(\sum_{k=0}^{M-1} U_{2k} + U_{2M}\Big)^2 = \Big(\sum_{k=0}^{M-1} U_{2k}\Big)^2 + U_{2M}^2 + 2\sum_{k=0}^{M} U_{2M}U_{2k} \\
&= \Big(\sum_{k=0}^{M-1} U_{2k}\Big)^2 + \sum_{k=0}^{2M} U_{2k} + 2\sum_{k=0}^{M-1}\sum_{j=M-k}^{M+k} U_{2j} \\
&= \Big(\sum_{k=0}^{M-1} U_{2k}\Big)^2 + \sum_{k=0}^{2M} (1 + 2(M - |M-k|)) U_{2k}.
\end{aligned}$$

Taking the Fourier transform on both sides, we get

$$\begin{aligned}
(\widehat{F_M^2})_l &= \Big(\big(\sum_{k=0}^{M-1} U_{2k}\big)^2\Big)^{\widehat{}}_l + \frac{1 + 2(M - |M-l|)}{2l+1} \chi_{[0,2M]}(l) \mathsf{id}_{2l+1} \\
&\le \Big(\big(\sum_{k=0}^{M-1} U_{2k}\big)^2\Big)^{\widehat{}}_l + \chi_{[0,2M]}(l) \mathsf{id}_{2l+1}.
\end{aligned}$$

Induction over M, yields

$$(\widehat{F_M^2})_l \leq (M+1)\mathsf{id}_{2l+1}, \quad 0 \leq l \leq 2M,$$

$$(\widehat{F_M^2})_l = 0_{2l+1}, \quad l > 2M.$$

Thus, we get the desired estimate

$$(\widehat{G_M})_l = \frac{a_{\min}(M)}{(M+1)}(\widehat{F_M^2})_l \leq a_{\min}(M)\chi_{[0,2M]}(l)\mathsf{id}_{2l+1} \leq \hat{\phi}_l, \quad l \in \mathbb{N}. \tag{20}$$

The function $\phi - G_M$ is therefore positive definite by Theorem 4.1 and Gerschgorin's theorem implies

$$\begin{aligned} \sum_{i,j=1}^{n} c_i\overline{c_j}\phi(x_i^{-1}x_j) &\geq \sum_{i,j=1}^{n} c_i\overline{c_j}G_M(x_i^{-1}x_j) \\ &\geq \|c\|_2^2\left(G_M(e) - \max_{1\leq j\leq n}\sum_{i=1,i\neq j}^{n}|G_M(x_i^{-1}x_j)|\right). \end{aligned} \tag{21}$$

In order to obtain our desired estimate, we show

$$\max_{1\leq j\leq n}\sum_{i=1,i\neq j}^{n}|G_M(x_i^{-1}x_j)| \leq \frac{G_M(e)}{2}, \tag{22}$$

for an appropriate $M > 0$. Now, we assume w.l.o.g that the maximum in (22) is attained at $x_1 = e$. Then we get

$$\max_{1\leq j\leq n}\sum_{i=1,i\neq j}^{n}|G_M(x_i^{-1}x_j)| = \sum_{j=2}^{n}|G_M(x_j)|.$$

We now decompose the rotation group into subsets of the form

$$R_{q_X,m} := \{x \in \mathsf{SO}(3) : \; mq \leq d(x,e) < (m+1)q\}.$$

From [6], we have the following upper bound for the number of points from X which can be contained in a set $R_{q_x,m}$:

$$|R_{q_X,m} \cap X| \leq 141m^2.$$

For the function G_M, we have

$$|G_M(x)| = \frac{a_{\min}(M)}{M+1}\frac{\sin\left((M+1)\frac{\alpha}{2}\right)^4}{\sin(\frac{\alpha}{2})^4} \leq \frac{a_{\min}(M)\pi^4}{(M+1)\alpha(x)^4} = \frac{\pi^4 G_M(e)}{((M+1)\alpha(x))^4}.$$

Thus, an upper bound for G_M on $R_{q_X,m}$ is given by

$$|G_M(x)| \leq G_M(e)\frac{\pi^4}{((M+1)mq_X)^4}, \quad x \in R_{q_X,m}.$$

Summarizing all statements, we finally get the following estimate

$$\sum_{j=2}^{n} |G_M(x_j)| \leq \sum_{m=1}^{\infty} |R_{q_X,m} \cap X| \sup_{x \in R_{q_X,m}} |G_M(x)|$$
$$\leq G_M(e) \frac{141\pi^4}{((M+1)q_X)^4} \sum_{m=1}^{\infty} \frac{1}{m^2} \leq G_M(e) \frac{24\pi^6}{((M+1)q_X)^4}.$$

In order to ensure that the last expression is less then $\frac{1}{2} G_M(e)$, we have to chose

$$M + 1 \geq \frac{(48\pi^6)^{\frac{1}{4}}}{q_X} \geq \frac{14.65}{q_X}.$$

■

Example

The following table contains the lower bounds for $a_{min}(M)$ and $\lambda_{min}(\boldsymbol{A}_\phi)$ for the functions introduced in section 4. To simplify things, we apply Theorem 5.1 with $M + 1 = 16/q_X$.

	$a_{min}(M)$	λ_{min}
P_r	$\frac{(1-r)^2}{1+r} \frac{r^{2M}}{4(M+1)}$	$32 \frac{(1-r)^2}{(1+r)r^2} r^{\frac{32}{q_X}} q_X^{-2}$
G_t	$e^{-4(M+1)^2 t}$	$2^{11} e^{-\frac{32^2}{q_X^2} t} q_X^{-3}$
E_p	$\frac{p^{2M-1}}{8e^p (2M+1)!}$	$\frac{2^8}{p^3 e^p} \frac{p^{32/q_X}}{\Gamma(32/q_X)} q_X^{-3}$

Remark that this method doesn't work with the function S_m since in that case only finite many Fourier coefficients are different from zero.

5.2 Error estimates

To get an error estimate for $|g(x) - s_g(x)|, g \in \mathcal{H}(f)$, we have to find a bound for the power function $P_{X,\phi}(x)$. This will be done in terms of the fill distance

$$h_X = \sup_{x \in \mathsf{SO}(3)} \inf_{x_j \in X} d(x, x_j). \tag{23}$$

The fill distance specifies the maximal distance between a point $x \in \mathsf{SO}(3)$ and the set X and thus describes how uniformly the points of the set X are distributed over the manifold $\mathsf{SO}(3)$.

Let

$$\mathcal{P}_N = \mathsf{span}\{D^l_{m,n} : \ 1 \leq l \leq N, \ -l \leq m, n \leq l\}. \tag{24}$$

The next lemma is a consequence of the Markov inequality for functions from $\mathcal{P}_N$. The proof is the same as in the case where the underlying manifold is $\mathbb{S}^d$ (see [18] or [10]) and is therefore omitted.

Lemma 5.2
Let $X = \{x_1, \dots, x_n\}$ be a subset of $\mathsf{SO}(3)$ with fill distance $h_X \leq \frac{1}{2N}$. Then there exist functions $\tilde{u}_i : \mathsf{SO}(3) \to \mathbb{R}$ such that

(i) $\sum_{i=1}^{n} \tilde{u}_i(x) p(x_i) = p(x)$ *for all* $p \in \mathcal{P}_N(\mathsf{SO}(3))$ *and* $x \in \mathsf{SO}(3)$.

(ii) $\sum_{i=1}^{n} |\tilde{u}_i(x)| \leq 2$ *for all* $x \in \mathsf{SO}(3)$.

With this preliminary we are able to prove the following error estimate.

Theorem 5.3
Let $\phi \in CP(\mathsf{SO}(3))$ be a central function with Fourier coefficients $a_l \mathsf{id}_{2l+1}, l \in \mathbb{N}$. Let $g \in \mathcal{H}(\phi)$, $X = \{x_1, \dots, x_n\}$, and assume $\frac{1}{2N+1} \leq h_X \leq \frac{1}{2N}$. Then, for $x \in \mathsf{SO}(3)$, we have

$$|g(x) - s_g(x)| \leq 3 \left(\sum_{l=N+1}^{\infty} (2l+1)^2 a_l \right)^{\frac{1}{2}} \|g\|_\phi.$$

Proof. We have
$$|g(x) - s_g(x)| \leq P_{X,\phi}(x) \|g\|_\phi.$$

Let $x_0 = x$ and $u_0(x) = -1$. Then, by using Corollary 4.2, we obtain the following expression for the power function

$$\begin{aligned} P^2_{X,\phi}(x) &= \|L_x\phi - \sum_{i=1}^{n} u_i(x) L_{x_i}\phi\|^2_\phi = \|\sum_{i=0}^{n} u_i(x) L_{x_i}\phi\|^2_\phi \\ &= \sum_{l=0}^{\infty} (2l+1) a_l \sum_{i=0}^{n}\sum_{j=1}^{n} u_i(x)\overline{u_j(x)} U_{2l}\Big(\cos\Big(\frac{\alpha(x_i^{-1}x_j)}{2}\Big)\Big). \end{aligned}$$

Since the quadratic form (16) is minimized pointwise by the functions u_i, we obtain an upper bound by replacing the u_i's by the functions $\tilde{u}_i$ from the previous lemma. Setting again $\tilde{u}_0(x) = -1$, we get

$$\begin{aligned} P^2_{X,f}(x) &\leq \sum_{l=N+1}^{\infty} (2l+1) a_l \sum_{i,j=0}^{n} \tilde{u}_i(x)\overline{\tilde{u}_j(x)} U_{2l}\Big(\cos\Big(\frac{\alpha(x_i^{-1}x_j)}{2}\Big)\Big) \\ &\leq \sum_{l=N+1}^{\infty} (2l+1)^2 a_l \sum_{i,j=0}^{n} |\tilde{u}_i(x)\overline{\tilde{u}_j(x)}| \\ &\leq \sum_{l=N+1}^{\infty} (2l+1)^2 a_l \left(1 + \sum_{i=1}^{n} |\tilde{u}_i(x)|\right)^2 \leq 9 \sum_{l=N+1}^{\infty} (2l+1)^2 a_l. \end{aligned}$$

Note that, since the function ϕ is central and in $CP(\mathsf{SO}(3))$, Bochner's theorem 4.2 ensures the existence of $\sum_{l=N+1}^{\infty}(2l+1)^2 a_l$. ■

We can express the error estimates in terms of the fill distance h_X if some information on the decay of the Fourier coefficients is known.

Theorem 5.4
Under the same assumptions as in Theorem 5.3, we have

(1) If $(2l+1)^2 a_l \leq c(2l+1)^{-s}$ *with* $s > 1$*, then*

$$\|g - s_g\|_\infty \leq C\, h_X^{\frac{s-1}{2}}\, \|g\|_\phi.$$

(2) If $(2l+1)^2 a_l \leq c e^{-s(2l+1)}$ *with* $s > 0$*, then*

$$\|g - s_g\|_\infty \leq C\, e^{-\frac{s}{2(h_X)}} \|g\|_\phi.$$

Proof. Inserting the assumptions on the Fourier coefficients in Theorem 5.3, we get

$$(1) \quad \sum_{l=N+1}^{\infty} (2l+1)^2 a_l \leq c \int_N^\infty (1+2l)^{-s} dl = \frac{c}{2(s-1)}(2N+1)^{-s+1} \leq \frac{c}{2(s-1)}(h_X)^{s-1},$$

$$(2) \quad \sum_{l=N+1}^{\infty} (2l+1)^2 a_l \leq c \int_N^\infty e^{-s(2l+1)} dl = \frac{c}{2s} e^{-s(2N+1)} \leq \frac{c}{2s} e^{-s/h_X}.$$

This proves the statements. ■

References

[1] M. Allali, T. Przebinda. *Strictly positive definite functions on a compact group*. Proc. Am. Math. Soc., **129** (5), 1459–1462, 2000.

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

[3] R.K. Beatson, J.B. Cherrie, C.T. Mouat, *Fast fitting of radial basis functions: Methods based on preconditioned GMRES iteration*, Adv. Comput. Math., **11**, 253-270, 1999.

[4] W. zu Castell, F. Filbir. *Radial basis functions and corresponding zonal series expansions on the sphere*. J. Approx. Theory, **134**, 65–79, 2005.

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

[6] F. Filbir, D. Schmid. *Stability results for approximation by positive definite functions on compact groups*, GSF-preprint 07-06 (http://ibb.gsf.de/preprints.php), to appear in J. Approx. Theory.

[7] I.M. Gelfand, R.A. Minlos, Z.Ya. Shapiro. *Representations of the rotation and Lorentz groups and their application*, Pergamon Press, Oxford, 1963.

[8] T. Gutzmer. *Interpolation by positive definite functions on locally compact groups with application to* $\mathsf{SO}(3)$. Res. in Math., **29**, 69-77, 1996.

[9] E. Hewitt, K. A. Ross. *Abstract Harmonic Analysis II.* Springer, Berlin, 1970.

[10] K. Jetter, J. Stöckler, J.D. Ward. *Error estimates for scattered data interpolation on spheres*. Mathematics of Computation, **68** (226), 733–747, 1999.

[11] J. Levesley, D. L. Ragozin, *The Density of Translates of Zonal Kernels on Compact Homogeneous Spaces*. J. Approx. Theory, **103**, 252–268, 2000.

[12] F.J. Narcowich. *Generalized Hermite Interpolation and positive definite kernels on a Riemannian manifold*. J. Math. Anal. Appl., **199**, 165–194, 1995.

[13] Z. Sasvári. *Positive Definite and Definitizable Functions.* Akademie Verlag, Berlin, 1994.

[14] R. Schaback, *Native Hilbert spaces for radial basis functions I.* in M. W. Müller et. al., eds., New Developments in Approximation Theory. 2nd International Dortmund Meeting (IDoMAT '98), vol. 132, Int. Ser. Numer. Math., pp. 255-282. Basel, Birhäuser Verlag, 1999.

[15] R. Schaback, *Error estimates and condition numbers for radial basis function interpolation*, Adv. Comput. Math., **3**, 251-264, 1995.

[16] D. Schmid, *A trade-off principle in connection with the approximation by positive definite kernels.* in M. Neamtu, L.L. Schumaker, eds., Approximation Theory 12: San Antonio 2007, pp. 348-359. Brentwood, Nashboro Press, 2008.

[17] N.J. Vilenkin. *Special Functions and the Theory of Group Representations.* American Mathematical Society, Rhode Island, 1968.

[18] H. Wendland. *Scattered Data Approximation.* Cambridge University Press, Cambridge, 2005.