TECHNISCHE UNIVERSITÄT
CHEMNITZ

Diplomarbeit

Inversion of the Circular Average Transform on the Sphere

Michael Quellmalz

September 2014

Betreuer: Dr. Ralf Hielscher

Advisor / primary reviewer: Dr. Ralf Hielscher
Secondary reviewer: Prof. Dr. Daniel Potts

Abstract

The reconstruction of a function defined on the two-sphere from its mean values along all circles of the sphere is an overdetermined problem. Every circle of the sphere is the intersection of the sphere with a plane. We say that two circles are perpendicular if the respective planes are. This thesis revolves around the circular average transform $\mathcal{T}$, which computes the mean values along all those circles that are perpendicular the equator of the sphere. We describe a singular value decomposition of the circular average transform based on the Funk-Hecke formula for spherical harmonics.

Any function f that is symmetrical with respect to the equator can be reconstructed from its circular average transform $\mathcal{T}f$, but the reconstruction is an ill-posed problem. We describe a fast algorithm for the reconstruction based on the discrete spherical Fourier transform. Therefore, we sample the function $\mathcal{T}f$ at discrete data points and assume that the sampled values are disturbed by white noise. As a regularization scheme we use the mollifier method, which essentially means that the smoothing in the reconstruction process is done by a convolution with the mollifier. Finally, we construct a family of mollifiers that minimizes the maximum risk of the reconstruction error assuming that the original function f has a certain degree of smoothness with respect to Sobolev spaces on the sphere.

Contents

List of Figures

Danksagungen (Acknowledgments)

Zuerst möchte ich auf diesem Wege Dr. Ralf Hielscher meinen Dank aussprechen. Er ist der Betreuer und erste Gutachter dieser Diplomarbeit. Er lieferte die Idee zu dieser Arbeit und stand mir während der Anfertigung stets mit Rat und Tat zur Seite. Darüber hinaus danke ich Prof. Dr. Daniel Potts für die fortwährende Unterstützung. Er ist außerdem der zweite Gutachter dieser Arbeit.

Besonderer Dank gilt auch meinen Eltern, die mich mein gesamtes Studium hindurch unterstützten.

1 Introduction

The problem of reconstructing a function from its integrals along certain lower-dimensional submanifolds has been studied since the early twentieth century. It is associated with the terms reconstructive integral geometry [42] and geometric tomography [17]. In a paper from 1917, Radon [44] introduced what later became known as the Radon transform, which maps a function defined on the plane to its mean values along all lines lying in the plane. The Radon transform provides the mathematical background of the computerized tomography, whose development started in the 1960s and which has become a crucial part of medical diagnostics since, cf. [49]. An even earlier work was published in 1913 by Funk [16]. He described the problem of reconstructing a function on the two-sphere knowing its mean values along all great circles of the sphere. The computation of these mean values is known as the Funk transform or spherical Radon transform. This is not to be confused with the similarly named spherical mean Radon transform, which computes the mean values over spheres of functions defined on the three-dimensional space, cf. [22, 36].

What Funk did for great circles can be generalized to other classes of circles on the unit sphere $\mathbb{S}^2$. One can consider the integration along all circles with a fixed radius. This is known as the translation operator, cf. [48, 7, 24]. Another example, introduced by Helgason [24, II.1.C], is the spherical slice transform, which computes the means along all circles that contain a fixed point of the sphere.

Any circle of the sphere is the intersection of the sphere with a plane. We call two circles perpendicular if the respective planes are. In this thesis, we consider the circular average transform which computes the averages of a function along all circles perpendicular to the equator of the sphere. Let us denote with $\boldsymbol{e}_\sigma = (\cos\sigma, \sin\sigma, 0)^\top$, $\sigma \in \mathbb{T} = [0, 2\pi)$, the point on the equator of the sphere with the latitude σ. The circular average transform of a continuous function $f : \mathbb{S}^2 \to \mathbb{C}$ is defined by

$$\mathcal{T} f(\sigma, t) = \frac{1}{2\pi\sqrt{1-t^2}} \int_{\langle \boldsymbol{y}, \boldsymbol{e}_\sigma \rangle = t} f(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}, \quad (\sigma, t) \in \mathbb{T} \times (-1, 1),$$

where the integration is carried out with respect to the arc length on the circle. For $t = \pm 1$, we set $\mathcal{T} f(\sigma, \pm 1) = f(\pm \boldsymbol{e}_\sigma)$.

The circular average transform first arose in 2010 in a problem related to photoacoustic tomography. Zangerl and Scherzer [59] described an algorithm for inverting the circular average transform using a connection to the circular Radon transform. We will describe this method in Subsection 3.3.1.

The circular average transform is a compact operator on $L^2(\mathbb{S}^2)$, thus it has a singular value decomposition. In Theorem 3.9, we describe a singular system $(Y_{n,k}, B_{n,k}, \lambda(n,k))_{n,k}$ such that $\mathcal{T} Y_{n,k} = \lambda(n,k) B_{n,k}$ for all $n \in \mathbb{N}_0$ and $k = -n, \ldots, n$. The spherical harmonics $Y_{n,k}$ are the basis functions of the singular value decomposition on the sphere. The basis functions $B_{n,k}(\sigma, t)$ in the range of $\mathcal{T}$ are the products of the exponential functions $\sigma \mapsto \mathrm{e}^{\mathrm{i}k\sigma}$ in the σ variable and the Legendre polynomials P_n in the t variable, normalized with respect to the L^2-norm, see (2.29). The singular values $\lambda(n,k) \in \mathbb{C}$ are known explicitly, see (3.15).

Reconstruction. We want to reconstruct a function f given its circular average transform $g = \mathcal{T} f$. Like it is the case for many other problems of this type, the inversion of the circular average transform is an ill-posed problem. This means that the reconstruction of f is very sensitive to small errors in the data g. Any practical observation incorporates some noise. So, instead of g, we only have the noisy data $g + \varepsilon$.

There are two main kinds of error assumptions in inverse problems. In the deterministic framework, the noise is assumed to be bounded in some norm, i.e. $\|\varepsilon\| \leq r$, cf. [29]. On the other hand, there is the statistical framework we chose in this thesis, where the noise is described by a random process, cf. [5]. Here, the number of observations of g goes to infinity, while the norm $\|\varepsilon\|$ converges to zero in the deterministic setting.

In most practical applications, the number of observations is finite, so it makes sense to sample the function g on discrete data points $(\sigma_l, t_m) \in \mathbb{T} \times [-1,1]$ for $l = 1, \ldots, L,\ m = 1, \ldots, M$. We consider the discrete noisy data

$$g^{\varepsilon}(\sigma_l, t_m) = g(\sigma_l, t_m) + \varepsilon(l,m), \quad l = 1, \ldots, L,\ m = 1, \ldots, M,$$

where $[\varepsilon(l,m)]_{l,m}$ is an uncorrelated, zero-mean random variable, see Definition 4.1.

To overcome the ill-posedness of the reconstruction problem, we use the mollifier method (cf. [38]), i.e. we aim at recovering a smoothened version of f, namely

$$\psi \star f(\boldsymbol{x}) = \int_{\mathbb{S}^2} \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) f(\boldsymbol{y}) \,\mathrm{d}\boldsymbol{y}, \quad \boldsymbol{x} \in \mathbb{S}^2,$$

where $\psi : [-1,1] \to \mathbb{R}$ is some mollifier, see Section 4.1. We are going to use the estimator

$$f_{\psi}^{\varepsilon} = \psi \star \left(\mathcal{T}^{\dagger} \mathcal{L}_N g^{\varepsilon} \right)$$

where $\mathcal{T}^{\dagger}$ denotes the pseudo-inverse of $\mathcal{T}$ and $\mathcal{L}_N$ is a (hyper-)interpolation of degree N of the sampled data. The $L \times M$-point hyperinterpolation

$$\mathcal{L}_N g = \sum_{n=0}^{N} \sum_{k=-n}^{n} \left(\sum_{l=1}^{L} \sum_{m=1}^{M} w_{l,m} g(\sigma_l, t_m) \overline{B_{n,k}(\sigma_l, t_m)} \right) B_{n,k}.$$

is a Fourier series with respect to the basis functions $B_{n,k}$, truncated at degree N, where the Fourier coefficients of g are computed using a quadrature rule. One should be aware that $\mathcal{L}_N$ is not necessarily an interpolation. We say that $\mathcal{L}_N$ is exact for a function $g \in C(\mathbb{T} \times [-1, 1])$ if $\mathcal{L}_N g = g$.

We assume that the function f belongs to the class of Sobolev-smooth functions

$$\mathscr{F}(s, S) = \left\{ f \in H^s\left(\mathbb{S}^2\right) \cap \mathrm{N}(\mathcal{T})^\perp : \|f\|_{H^s(\mathbb{S}^2)} \leq S \right\}$$

where $s, S > 0$ and $\mathrm{N}(\mathcal{T})$ denotes the nullspace of the operator $\mathcal{T}$. The Sobolev space $H^s(\mathbb{S}^2)$ is equal to the set of functions whose weak derivatives up to the order s exist and are square-integrable, provided s is an integer, cf. Subsection 2.2.3.

The minimax risk of the estimator f_ψ^ε is defined as

$$\inf_{\psi} \sup_{f \in \mathscr{F}(s,S)} \mathbb{E} \left\| f_\psi^\varepsilon - f \right\|_{L^2(\mathbb{S}^2)}^2,$$

cf. [53, 5]. Our aim is to find optimal mollifiers ψ^* for which the above infimum is attained, i.e.

$$\sup_{f \in \mathscr{F}(s,S)} \mathbb{E} \left\| f_{\psi^*}^\varepsilon - f \right\|_{L^2(\mathbb{S}^2)}^2 = \inf_{\psi} \sup_{f \in \mathscr{F}(s,S)} \mathbb{E} \left\| f_\psi^\varepsilon - f \right\|_{L^2(\mathbb{S}^2)}^2.$$

Since this is a rather tough problem, we look for asymptotically optimal mollifiers instead, as the number of sampling points goes to infinity.

In the first stage, we prove an asymptotic upper bound for the minimax risk in (4.36). Therefore, we need to make two assumptions on the $L \times M$-point hyperinterpolation $\mathcal{L}_N$ that we apply to the sampled data on the set $\mathbb{T} \times [-1, 1]$, cf. Definition 4.7. The first one is pretty natural, it says that $\mathcal{L}_N$ should be exact for the basis functions $B_{n,k}$ if $n \leq N$. Secondly, we assume that there exist some constants $\underline{\gamma}, \overline{\gamma} > 0$ such that

$$\underline{\gamma} \frac{4\pi}{LM} \leq \sum_{l=1}^{L} \sum_{m=1}^{M} \left| w_{l,m} B_{n,k}(\sigma_l, t_m) \right|^2 \leq \overline{\gamma} \frac{4\pi}{LM}, \quad n = 0, \ldots, N, \; k = -n, \ldots, n,$$

if L and M are sufficiently large. This condition is satisfied with $\underline{\gamma} = \overline{\gamma} = 1$ if the weights $w_{l,m}$ are constant, i.e. independent of l and m. Since constant-weight quadrature rules with a certain exactness on the unit interval and thus on $\mathbb{T} \times [-1, 1]$ are difficult to achieve, we also prove that the above condition holds for the hyperinterpolation based on the Fejér quadrature on the unit interval, cf. Corollary 4.11.

In the second stage, in Lemma 4.15, we prove that the mollifier can be chosen out of a class of mollifiers, namely

$$\psi_{\tilde{N}} = \sum_{n=0}^{\tilde{N}} \frac{2n+1}{4\pi} \left(1 - \left(\frac{n + \frac{1}{2}}{\tilde{N} + \frac{1}{2}} \right)^s \right) P_n, \quad \tilde{N} \geq 0,$$

then the error is still below the upper bound we derived for the minimax risk. This also gives us an asymptotic lower bound for the minimax risk.

Theorem 4.17 serves as the main result of this thesis. It shows the asymptotic optimality of the so-defined family of mollifiers $\psi_{\tilde{N}}$ for the class $\mathscr{F}(s,S)$ provided s is sufficiently large. In particular, we prove that there exists a sequence $\tilde{N}(L,M)$ such that the minimax risk is asymptotically achieved by using the mollifier $\psi_{\tilde{N}(L,M)}$ where the number of sampling points LM goes to infinity, i.e.

$$\sup_{f\in\mathscr{F}(s,S)} \mathbb{E}\left\| f - f^{\varepsilon}_{\psi_{\tilde{N}(L,M)}} \right\|^2_{L^2(\mathbb{S}^2)} \simeq \inf_{\psi\in L^2([-1,1])} \sup_{f\in\mathscr{F}(s,S)} \mathbb{E}\left\| f - f^{\varepsilon}_{\psi}\right\|^2_{L^2(\mathbb{S}^2)}.$$

Furthermore, the minimax risk decreases of the order $(LM)^{-\frac{2s}{2s+3}}$.

In [5], the Dirichlet kernel $\psi^{\mathrm{Dir}}_{\tilde{N}} = \sum_{n=1}^{\tilde{N}} (2n+1)/(4\pi)\, P_n$ was considered as mollifier in a more general setting. It was shown that the error decreases of the same asymptotic rate as for the optimal mollifiers. However, they assumed that the Fourier coefficients of g were given, compared to the sampled values of g in our setting.

Structure of this thesis. The rest of this thesis consists of three chapters. The next chapter describes some basics that will be useful later on. In Section 2.1, we will present some facts about quadrature and approximation on the relevant manifolds for the circular average transform, particularly the unit sphere $\mathbb{S}^2$ and the side of the cylinder $\mathbb{T}\times[-1,1]$. We generalize a result from Sloan [50] for the interpolation error in Theorem 2.2. An introduction to the theory of spherical harmonics, which form a complete system of orthogonal polynomials on the sphere, follows in Section 2.2.

Chapter 3 is about the definition and properties of the circular average transform. At first, we define the circular average transform in terms of the mean operator on the sphere, which maps a function defined on the sphere to its mean values along all circles. Then we construct some maps that connect the circular average transform with the circular Radon transform and the Radon transform whose inversions have already been subject to research, see e.g. [58, 46, 2, 41, 44, 24] and [44, 24], respectively. Using stereographic projection, we can relate the circular average transform to the circular Radon transform in the plane, which calculates the mean values of a function defined on the plane along all circles centered on a line, see Theorem 3.7. Furthermore, we can use orthogonal projection to establish a connection with the Radon transform, see Theorem 3.8. In Section 3.4, we compute the singular value decomposition of the circular average transform. Based on this and the fast spherical Fourier transform [31], we present a fast algorithm for the computation of the circular average transform in Section 3.5.

Chapter 4 is about the inversion of the circular average transform. At first, we present the setting we use for the regularization of the problem. Then, in Section 4.2, we show that the expected reconstruction error can be decomposed in two parts and examine those parts separately yielding to upper bounds of the error. In Section 4.3, we compute lower bounds of the error and show that the family of mollifiers (4.37) is asymptotically optimal. In Section 4.4, we describe an application of the inverse circular average transform in the context of photoacoustic tomography. An algorithm that computes the estimator f^{ε}_{ψ} numerically is presented in Section 4.5. A final conclusion follows in Chapter 5.

2 Preliminaries

We introduce some common symbols describing the asymptotic growth of sequences. For simplicity, let $(\alpha_j)_j$ and $(\beta_j)_j$ be two non-zero sequences of real numbers. The sign "$\lesssim$" means "asymptotically less than or equal to". We say that $\alpha_j \lesssim \beta_j$ for $j \to \infty$ if $\limsup_{j\to\infty} \alpha_j/\beta_j \leq 1$. The symbol "$\gtrsim$" stands for "asymptotically less than or equal to" and is defined analogously. Furthermore, $\alpha_j \simeq \beta_j$ if $\alpha_j \lesssim \beta_j$ and $\alpha_j \gtrsim \beta_j$. For the well-known Big Oh notation, only the order of growth matters. We say that $\alpha_j \in \mathcal{O}(\beta_j)$ if there exists a $c \in \mathbb{R}$ such that $\alpha_j \lesssim c\beta_j$. Moreover, we introduce the Big Theta notation, $\alpha_j \in \Theta(\beta_j)$ if $\alpha_j \in \mathcal{O}(\beta_j)$ and $b(j) \in \mathcal{O}(\alpha_j)$.

2.1 Quadrature

In this section, we present some basic facts about quadrature, a.k.a. numerical integration. We first describe the general case and we will come to quadrature on specific manifolds later. The first part of our considerations is based on [50].

Let $\Omega \subset \mathbb{R}^d$ be either the closure of an open, connected set or a closed smooth manifold of lower dimension. Let μ be a finite measure on Ω. We usually write $\mathrm{d}x$ instead of $\mathrm{d}\mu(x)$ if it is clear which measure is meant. We denote by $|\Omega| = \int_\Omega 1 \,\mathrm{d}\mu(x)$ the volume of Ω. The space $C(\Omega)$ is the vector space of all continuous functions defined on Ω with values in the complex numbers $\mathbb{C}$. Equipped with the supremum norm $\|f\|_{C(\Omega)} = \|f\|_\infty = \sup_{x\in\Omega} |f(x)|$, the space of continuous functions is a Banach space. The Lebesgue space $L^2(\Omega)$ consists of all measurable functions $f : \Omega \to \mathbb{C}$ that satisfy $\int_\Omega |f(x)|^2 \,\mathrm{d}\mu(x) < \infty$. It is well-known that $L^2(\Omega)$ is a separable Hilbert space with the inner product

$$\langle f, g\rangle := \int_\Omega f(x)\overline{g(x)} \,\mathrm{d}x, \quad f, g \in L^2(\Omega) \tag{2.1}$$

and the norm

$$\|f\|_{L^2(\Omega)} = \sqrt{\int_\Omega |f(x)|^2 \,\mathrm{d}x}, \quad f \in L^2(\Omega).$$

Let e_n, $n \in \mathbb{N}$, be an orthonormal basis of $L^2(\Omega)$, that is:

- $\langle e_n, e_m \rangle = \delta_{nm}$, where δ denotes the Kronecker delta, and
- every function $f \in L^2(\Omega)$ can be written as Fourier series
$$f = \sum_{n=1}^{\infty} \hat{f}(n)\, e_n$$
with the Fourier coefficients
$$\hat{f}(n) = \int_{\Omega} f(x) e_n(x)\, \mathrm{d}x, \quad n \in \mathbb{N}.$$

In $L^2(\Omega)$, like in every other Hilbert space, Parseval's identity

$$\|f\|^2_{L^2(\Omega)} = \sum_{n=1}^{\infty} \left|\hat{f}(n)\right|^2, \quad f \in L^2(\Omega), \tag{2.2}$$

holds, cf. [57, V.4.9,(b),(vi)]. Let

$$\mathcal{P}_N f = \sum_{n=1}^{N} \langle f, e_n \rangle\, e_n. \tag{2.3}$$

denote the orthogonal projection of the function $f \in L^2(\Omega)$ onto

$$\Pi_N(\Omega) = \operatorname{span}\{e_n : n = 1, \ldots, N\},$$

where span denotes the set of all finite linear combinations. Let a quadrature formula Q on Ω be given by

$$Qf = \sum_{j=1}^{J} w_j f(x_j), \quad f \in C(\Omega) \tag{2.4}$$

with the quadrature nodes $x_j \in \Omega$ and the corresponding weights $w_j \in \mathbb{R}$, $j = 1, \ldots, J$. The quadrature formula Q is called exact for a set $A \subset L^2(\Omega) \cap C(\Omega)$ if

$$Qf = \int_{\Omega} f(x)\, \mathrm{d}x \text{ for all } f \in A.$$

In case that Q is exact for $A = \Pi_N(\Omega)$, we say that Q has the order of exactness N.

Applying the quadrature rule Q to the computation of the inner product (2.1) leads to the *discretized inner product*

$$\langle f, g \rangle_Q = \sum_{j=1}^{J} f(x_j)\, \overline{g(x_j)} w_j, \quad f, g \in C(\Omega).$$

It is easy to check that $\langle \cdot, \cdot \rangle_Q$ is a non-negative sesquilinear form. But, in general, $\langle \cdot, \cdot \rangle_Q$ is not an inner product because it lacks the positive definiteness. If the quadrature Q is exact for $\{e_n e_m : n, m = 1, \ldots, N\}$, then the discrete orthogonality relation

$$\langle e_n, e_m \rangle_Q = \delta_{nm}, \quad n, m = 1, \ldots, N, \tag{2.5}$$

holds. Now we can define a discretized version of the orthogonal projection $\mathcal{P}_N$ from (2.3),

$$\mathcal{L}_N f = \sum_{n=1}^{N} \langle f, e_n \rangle_Q \, e_n. \tag{2.6}$$

The operator $\mathcal{L}_N$ is referred to as hyperinterpolation. If the equations

$$\mathcal{L}_N f\,(x_j) = f\,(x_j)\,, \quad j = 1, \ldots, J, \tag{2.7}$$

hold for all continuous functions f, then $\mathcal{L}_N$ is called interpolatory. The following two lemmas with properties of the discretized inner product and the hyperinterpolation have been shown in [50, Lemma 5]. There, it was assumed that the functions e_n are polynomials, but the proof is basically the same.

Lemma 2.1. *Let $f \in C(\Omega)$, $N \in \mathbb{N}$, and the quadrature Q be exact for the set*

$$\operatorname{span}\left\{e_n \overline{e_m} : n, m = 1, \ldots, N\right\}.$$

Then the following statements hold:

1. $\langle f - \mathcal{L}_N f, g \rangle_Q = 0$ *for all* $g \in \Pi_N(\Omega)$
2. $\langle \mathcal{L}_N f, \mathcal{L}_N f \rangle_Q \leq \langle f, f \rangle_Q$

Proof. Let $g \in \Pi_N$, then, by the assumed exactness of Q and the discrete orthogonality relation (2.5),

$$\begin{aligned}
\langle f - \mathcal{L}_N f, g \rangle_Q &= \left\langle f - \sum_{n=1}^{N} \langle f, e_n \rangle_Q \, e_n, g \right\rangle_Q \\
&= \left\langle f - \sum_{n=1}^{N} \langle f, e_n \rangle_Q \, e_n, \sum_{n'=1}^{N} \langle g, e_{n'} \rangle_Q \, e_{n'} \right\rangle_Q \\
&= \left\langle f, \sum_{n=1}^{N} \langle g, e_n \rangle_Q \, e_n \right\rangle_Q - \sum_{n=1}^{N} \langle f, e_n \rangle_Q \langle g, e_n \rangle_Q \\
&= 0
\end{aligned}$$

which shows the first equation. This also implies that

$$\begin{aligned}
\langle \mathcal{L}_N f, \mathcal{L}_N f \rangle_Q &= \langle f, \mathcal{L}_N f \rangle_Q \\
&= \langle f, f \rangle_Q - \langle f, f - \mathcal{L}_N f \rangle_Q \\
&= \langle f, f \rangle_Q - \langle f - \mathcal{L}_N f, f - \mathcal{L}_N f \rangle_Q
\end{aligned}$$

The second equation follows because $\langle g, g \rangle_Q \geq 0$ for all continuous functions g. □

Theorem 2.2. *Let $f \in C(\Omega)$, $N \in \mathbb{N}$, and the quadrature Q be exact for*

$$\operatorname{span}\{e_n e_m : n, m = 0, \ldots, N\}$$

and for constant functions. Then

$$\|\mathcal{L}_N f - f\|_{L^2(\Omega)} \leq 2\sqrt{|\Omega|} \inf_{g \in \Pi_N} \|f - g\|_{C(\Omega)}. \tag{2.8}$$

Proof. From Lemma 2.1, we obtain the estimate

$$\begin{aligned}
\|\mathcal{L}_N f\|_{L^2(\Omega)}^2 &= \langle \mathcal{L}_N f, \mathcal{L}_N f \rangle_Q \\
&\leq \langle f, f \rangle_Q \\
&= \sum_{j=1}^{J} w_j |f(x_j)|^2 \\
&\leq \sum_{j=1}^{J} w_j \|f\|_{C(\Omega)}^2 \\
&= |\Omega| \|f\|_{C(\Omega)}^2 .
\end{aligned}$$

In the last line, we have used the exactness of Q for constant functions. For any $p \in \Pi_N(\Omega)$, we have $\mathcal{L}_N p = p$ and, by the above,

$$\begin{aligned}
\|\mathcal{L}_N f - f\|_{L^2(\Omega)} &\leq \|\mathcal{L}_N (f - p)\|_{L^2(\Omega)} + \|f - p\|_{L^2(\Omega)} \\
&\leq \sqrt{|\Omega|} \|f - p\|_{C(\Omega)} + \sqrt{|\Omega|} \|f - p\|_{C(\Omega)} .
\end{aligned}$$

Since this estimate holds for all $p \in \Pi_N(\Omega)$, it implies (2.8). □

2.1.1 The unit interval

Let $\Omega = [-1, 1]$ and let μ be the Lebesgue measure on the interval $[-1, 1]$. We denote with $\Pi_n(\mathbb{R})$ the set of polynomials in one variable that have a degree not greater than n. The orthogonal polynomials in this space are the well-known Legendre polynomials P_n, $n \in \mathbb{N}_0$. The following formulas can be found in [1, Section 22]. An explicit expression is given through Rodrigues' formula

$$P_n(t) = \frac{1}{2^n n!} \cdot \frac{\mathrm{d}^n}{\mathrm{d}t^n} \left(\left(t^2 - 1 \right)^n \right), \quad t \in [-1, 1]. \tag{2.9}$$

The Legendre polynomials P_n are polynomials of degree n with $P_n(1) = 1$. They satisfy the three-term recurrence relation

$$P_n(t) = \frac{2n-1}{n} t P_{n-1}(t) - \frac{n-1}{n} P_{n-2}(t), \quad n \geq 1, \tag{2.10}$$

which is initialized by the equations

$$P_0(t) \equiv 1, \quad P_{-1}(t) \equiv 0.$$

The orthogonal relation for the Legendre polynomials is

$$\int_{-1}^{1} P_n(t)\, P_m(t)\, \mathrm{d}t = \frac{2}{2n+1}\delta_{nm}, \quad n, m \in \mathbb{N}_0, \tag{2.11}$$

where δ denotes the Kronecker delta.
Any function $f \in L^2([-1,1])$ can be expanded into a Legendre series

$$f = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \hat{f}(n) P_n \tag{2.12}$$

with the Legendre coefficients

$$\hat{f}(n) = 2\pi \int_{-1}^{1} f(t) P_n(t)\, \mathrm{d}t, \quad n \in \mathbb{N}_0. \tag{2.13}$$

The following inequality holds for the Legendre polynomials, cf. [1, 22.14.9]

$$|P_n(\cos\theta)| \leq \min\left(\sqrt{\frac{2}{\pi n}} \frac{1}{\sqrt{\sin\theta}}, 1\right), \quad 0 < \theta < \pi,\ n \in \mathbb{N}_0. \tag{2.14}$$

An asymptotic approximation of the Legendre polynomials is given by Stieltjes' generalization of the Laplace-Heine formula, cf. [51, 8.21], for $n \to \infty$

$$P_n(\cos\theta) = \sqrt{\frac{2}{\pi n}} \frac{\cos\left(\left(n + \frac{1}{2}\right)\theta - \frac{\pi}{4}\right)}{\sqrt{\sin\theta}} + \mathcal{O}\left((n \sin\theta)^{-3/2}\right), \quad 0 < \theta < \pi. \tag{2.15}$$

2.1.1.1 Quadrature on the unit interval

There are several different choices of quadrature formulas

$$Q_M f = \sum_{m=1}^{M} \omega_m f(t_m) = \int_{-1}^{1} f(t)\, \mathrm{d}\mu(t) + R_M f \tag{2.16}$$

for some measure μ on the unit interval $[-1,1]$. We call Q_M exact of degree $n \in \mathbb{N}_0$ if $R_M f = 0$ for all polynomials f having degree up to n. The m-th Lagrange basis polynomial l_m is a polynomial of degree $M-1$ which is uniquely determined by $l_m(t_k) = \delta_{mk}$, $m, k =$

$1, \dots, M$. For any pairwise disjoint set of nodes $\{t_m : m = 1, \dots, M\}$, we can use the weights

$$\omega_m = \int_{-1}^{1} l_m(t) w(t) \, \mathrm{d}\mu(t), \quad m = 1, \dots, M, \tag{2.17}$$

which are also known as the Cotes or Christoffel numbers. Using the so-defined weights, the quadrature Q_M is exact of degree $M-1$.

The *Gaussian quadrature* uses as nodes t_m the zeros of the orthogonal polynomials with respect to μ. The Gaussian quadrature is exact of degree $2M-1$, see [52, Section 19]. When μ is the Lebesgue measure, i.e. $\mathrm{d}\mu(t) = \mathrm{d}t$, the orthogonal polynomials are the Legendre polynomials (2.9) and the corresponding Gaussian rule is called the *Gauss-Legendre quadrature.*

The Gauss-Chebyshev quadrature of the first kind [13] is the Gaussian quadrature for the Chebyshev weight of the first kind $\mathrm{d}\mu(t) = \frac{1}{\sqrt{1-t^2}} \, \mathrm{d}t$. This quadrature uses the Chebyshev nodes of the first kind

$$t_m^{\mathrm{C1}} = \cos\left(\theta_m^{\mathrm{C1}}\right), \quad \theta_m^{\mathrm{C1}} = \frac{(2m-1)\pi}{2M}, \quad m = 1, \dots, M. \tag{2.18}$$

The Gauss-Chebyshev quadrature of the first kind is denoted by

$$Q_M^{\mathrm{GC1}} f = \frac{\pi}{M} \sum_{m=1}^{M} f\left(\cos\left(\theta_m^{\mathrm{C1}}\right)\right), \quad f \in C([-1,1]). \tag{2.19}$$

The Gauss-Chebyshev quadrature of the second kind [11] is the Gaussian quadrature for the Chebyshev weight of the second kind $\mathrm{d}\mu(t) = \sqrt{1-t^2} \, \mathrm{d}t$. This quadrature uses the Chebyshev nodes of the second kind

$$t_m^{\mathrm{C2}} = \cos\left(\theta_m^{\mathrm{C2}}\right), \quad \theta_m^{\mathrm{C2}} = \frac{m\pi}{M+1}, \quad m = 1, \dots, M. \tag{2.20}$$

The Gauss-Chebyshev quadrature of the second kind is denoted by

$$Q_M^{\mathrm{GC2}} f = \frac{\pi}{M+1} \sum_{m=1}^{M} \left(\sin\left(\theta_m^{\mathrm{C2}}\right)\right)^2 f\left(\cos\left(\theta_m^{\mathrm{C2}}\right)\right), \quad f \in C([-1,1]). \tag{2.21}$$

Using the Chebyshev nodes of the second kind for a quadrature rule for the Lebesque measure leads to the *Fejér quadrature* [18, 55]. The weights (2.17) of the Fejér quadrature can be expressed through the explicit formula

$$\omega_m^{\mathrm{F}} = \frac{4}{M+1} \sin\left(\theta_m^{\mathrm{C2}}\right) \sum_{j=1}^{\lfloor M/2 \rfloor} \frac{\sin\left((2j-1)\,\theta_m^{\mathrm{C2}}\right)}{2j-1}, \quad m = 1, \dots, M. \tag{2.22}$$

The Fejér quadrature is exact for polynomials of degree $M-1$ with respect to the Lebesgue measure $\mathrm{d}\mu(t) = \mathrm{d}t$. The Clenshaw-Curtis quadrature is very similar to the Fejér quadrature: in (2.20), $m = 1, \dots, M$ is replaced by $m = 0, \dots, M+1$ and the weights are chosen like in (2.17), cf. [52]. There are also explicit formulas for the Clenshaw-Curtis weights.

The asymptotic behavior of the nodes and weights of the quadrature formulas are of interest. The asymptotic distribution of the nodes is the same for Legendre, Clenshaw-Curtis and Fejér, so the nodes θ_m are almost uniformly distributed on $[0, \pi]$ for M large, cf. [33, 12]. Equivalently, the asymptotic distribution of t_m is given by $(1-t^2)^{-1/2}$.
For the Cotes numbers ω_m, the circle theorem of Davis and Rabinowitz [9] states that

$$\frac{M}{\pi}\omega_m^{(M)} \simeq \sqrt{1-\left(t_m^{(M)}\right)^2} \tag{2.23}$$

as $M \to \infty$ for the nodes and weights of the Gauss-Legendre quadrature. This theorem was extended to the Clenshaw-Curtis and Fejér quadrature in [37]. Figure 2.1 on page 21 gives a visualization of the circle theorem. A further estimate presented in [14] states that for every quadrature with non-negative weights w_m, $m = 1, \ldots, M$, which is exact for polynomials with degree $\leq 2k+1$, its weights satisfy

$$\omega_m < \frac{\pi}{k}\left(\sqrt{1-x_m^2} + \frac{\pi}{k}\left|t_m\right|\right). \tag{2.24}$$

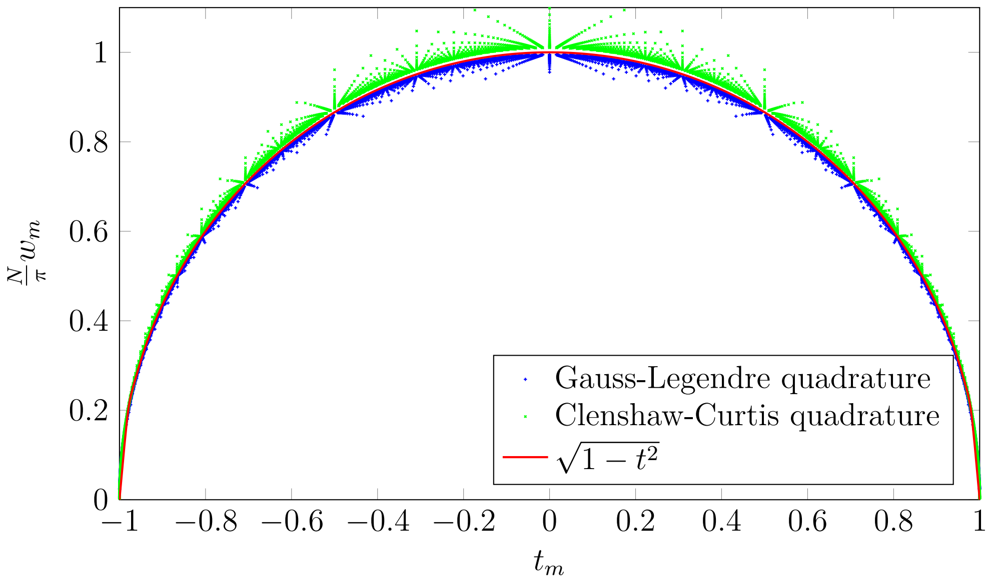

Figure 2.1: Plot of $M\omega_m/\pi$ for $M = 10, \ldots, 100$ where ω_m are the weights for the M-point Gauss-Legendre and Clenshaw-Curtis quadrature. The red curve is the graph of the function $\sqrt{1-t^2}$ to illustrate the equation (2.23) from the circle theorem.

Up to this point, we chose the nodes and computed the corresponding weights to achieve exactness of the quadrature. Alternatively, we can fix the quadrature weights to be constant and try to choose appropriate nodes t_m. This is called *Chebyshev-type quadrature*. When the quadrature is exact for constant functions, we have $\sum_{m=1}^{M} \omega_m = \int_{-1}^{1} 1 \,\mathrm{d}t = 2$ which shows that ω_m must be $2/M$. So we define the Chebyshev-type quadrature rule by

$$Q_M^{\mathrm{C}} f = \frac{2}{M}\sum_{m=1}^{M} f\left(t_m\right). \tag{2.25}$$

Chebyshev-type quadrature rules with degree of exactness M are known to exist, but, unfortunately, for $M > 10$ the nodes t_m are not all in the unit interval, instead $t_m \in \mathbb{C} \setminus [-1, 1]$ for at least one $m \in \{1, \ldots, M\}$, cf. [19]. When we require that $t_m \in [-1, 1]$ for all m, such quadratures having degree of exactness n exist with the number of nodes M_n satisfying

$$0.269n^2 < M_{2n-1} < 5.657n^2,$$

which was shown in [32]. One should note that the estimate $M_{2n-1} = \mathcal{O}(n^2)$ for the Chebyshev-type quadrature is very bad compared to $M_{2n-1} = \mathcal{O}(n)$ for the quadrature formulas from above like Gaussian and Clenshaw-Curtis. This is the reason why Chebyshev-type rules are rarely used in practice for $N > 10$.

2.1.2 The one-dimensional torus

Another well-described case is $\Omega = \mathbb{T}$ where $\mathbb{T}$ is the one-dimensional torus which is the same as the one-dimensional sphere. Then, the orthonormal polynomials can be written as

$$[0, 2\pi] \to \mathbb{C}, \, \sigma \mapsto \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}k\sigma}, \quad k \in \mathbb{Z},$$

where $\mathbb{Z}$ denotes the set of all integers. The degree of these polynomials is defined by $\deg \mathrm{e}^{\mathrm{i}k(\cdot)} = |k|$. The Fourier expansion of a function $f \in L^2(\mathbb{T})$ is given by

$$f(\sigma) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \hat{f}(k) \, \mathrm{e}^{\mathrm{i}k\sigma}, \quad \sigma \in \mathbb{T},$$

with the Fourier coefficients

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(\sigma) \mathrm{e}^{-\mathrm{i}k\sigma} \, \mathrm{d}\sigma, \quad k \in \mathbb{Z}. \tag{2.26}$$

On the torus, quadrature rules with equidistant nodes $2\pi l/L$, $l = 0, \ldots, L-1$, and equal weights $2\pi/L$ are a convenient choice. The quadrature is exact of degree L and the hyperinterpolation of degree $L - 1$ with this quadrature is interpolatory. Applying such quadrature to the computation of the Fourier coefficients in (2.26) yields to the discrete Fourier transform. The discrete Fourier transform $\mathrm{DFT}(\boldsymbol{f})$ of a vector $\boldsymbol{f} \in \mathbb{C}^L$ is defined by

$$[\mathrm{DFT}(\boldsymbol{f})](k) = \sum_{l=0}^{L-1} f(l) \mathrm{e}^{-2\pi \mathrm{i} kl/L}, \quad k = 0, \ldots, L-1. \tag{2.27}$$

The discrete Fourier transform can be computed efficiently in $\mathcal{O}(L \log L)$ steps using the famous Fast Fourier Transform (FFT) which was introduced in [6]. The inverse discrete Fourier transform (IDFT) is defined by

$$f(l) = \frac{1}{L} \sum_{k=0}^{L-1} \hat{f}(k) \mathrm{e}^{2\pi \mathrm{i} kl/L}. \tag{2.28}$$

2.1.3 Tensor product of Lebesgue spaces

We consider the space $L^2(\mathbb{T}\times[-1,1])$ with the Lebesgue measure. This is the tensor product space of $L^2(\mathbb{T})$ and $L^2([-1,1])$. The set $\mathbb{T}\times[-1,1]$ can be imagined as the side of a cylinder. We define the functions

$$B_{n,k}:\mathbb{T}\times[-1,1]\to\mathbb{C},\ (\sigma,t)\mapsto\frac{\sqrt{2n+1}}{\sqrt{4\pi}}\mathrm{e}^{\mathrm{i}k\sigma}P_n(t),\quad n\in\mathbb{N}_0,\ k\in\mathbb{Z},\tag{2.29}$$

which are just the products of the Legendre polynomials P_n and the exponential function $\mathrm{e}^{\mathrm{i}k(\cdot)}$ normalized with respect to the L^2 norm. The set $\{B_{n,k}:n\in\mathbb{N}_0,k\in\mathbb{Z}\}$ forms an orthonormal basis of polynomials in $L^2(\mathbb{T}\times[-1,1])$.

For quadrature rules on $\mathbb{T}\times[-1,1]$, we also use the ansatz of a tensor product. Let

$$Q_L^{\mathrm{S}}g=\sum_{l=1}^{L}\omega_l^{\mathrm{S}}g(\sigma_l),\quad g\in C(\mathbb{T})$$

and

$$Q_M^{\mathrm{I}}h=\sum_{m=1}^{M}\omega_m^{\mathrm{I}}h(t_m),\quad h\in C([-1,1])$$

be quadrature rules on the torus and the unit interval, respectively. Then, we define the quadrature

$$Qf=\sum_{m=1}^{M}\sum_{l=1}^{L}w_{l,m}f(\sigma_l,t_m),\quad f\in C(\mathbb{T}\times[-1,1])$$

where $w_{l,m}=\omega_l^{\mathrm{S}}\omega_m^{\mathrm{I}}$.

Proposition 2.3. *Let $N\in\mathbb{N}$ and Q be the quadrature on $\mathbb{T}\times[-1,1]$ as above, where Q_L^{S} is exact of degree $L=2N$, Q_M^{I} is exact of degree $M-1=2N$, and the respective hyperinterpolations are interpolatory. Then the hyperinterpolation*

$$\mathcal{L}_{N,Q}f(\sigma_l,t_m)=\sum_{n=0}^{N}\sum_{k=-N}^{N}\langle f,B_{n,k}\rangle_Q B_{n,k}(\sigma_l,t_m)$$

is an interpolation.

Proof. Let $N\in\mathbb{N}$. We have

$$\begin{aligned}\mathcal{L}_{N,Q}f(\sigma_l,t_m)&=\sum_{n=0}^{N}\sum_{k=-N}^{N}\langle f,B_{n,k}\rangle_Q B_{n,k}(\sigma_l,t_m)\\&=\sum_{n=0}^{N}\sum_{k=-N}^{N}\sum_{m'=1}^{M}\sum_{l'=1}^{L}w_{l',m'}f(\sigma_{l'},t_{m'})B_{n,k}(\sigma_l,t_{m'})B_{n,k}(\sigma_l,t_m)\\&=\sum_{n=0}^{N}\frac{2n+1}{4\pi}\sum_{k=-N}^{N}\sum_{m'=1}^{M}\sum_{l'=1}^{L}\omega_{l'}^{\mathrm{S}}\omega_{m'}^{\mathrm{I}}f(\sigma_{l'},t_{m'})\mathrm{e}^{\mathrm{i}k\sigma_{l'}}P_n(t_{m'})\mathrm{e}^{\mathrm{i}k\sigma_l}P_n(t_m)\\&=\sum_{n=0}^{N}\frac{2n+1}{4\pi}\sum_{m'=1}^{M}\omega_{m'}^{\mathrm{I}}P_n(t_{m'})P_n(t_m)\left(\sum_{k=-N}^{N}\sum_{l'=1}^{L}\omega_{l'}^{\mathrm{S}}f(\sigma_{l'},t_{m'})\mathrm{e}^{\mathrm{i}k\sigma_{l'}}\mathrm{e}^{\mathrm{i}k\sigma_l}\right).\end{aligned}$$

From the assumed exactness, we observe that

$$\begin{aligned}\mathcal{L}_{N,Q} f\left(\sigma_l, t_m\right) &= \sum_{n=0}^{N} \frac{2n+1}{2} \sum_{m'=1}^{M} \omega_{m'}^{\mathrm{I}} P_n\left(t_{m'}\right) P_n\left(t_m\right) f\left(\sigma_l, t_{m'}\right) \\ &= f\left(\sigma_l, t_m\right),\end{aligned}$$

which shows that $\mathcal{L}_{N,Q}$ is an interpolation. □

2.2 Approximation theory on the sphere

In this section, we present some basic results of the approximation theory on the two-dimensional sphere $\mathbb{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3 : |\boldsymbol{x}| = 1\}$ which we are going to use later. Here, $|\boldsymbol{x}|$ denotes the Euclidean norm of the vector $\boldsymbol{x}$. Further information about these topics can be found in books such as [15] or [7].

A vector $\boldsymbol{x} \in \mathbb{S}^2$ can be written in the spherical coordinates φ and ϑ as

$$\boldsymbol{x}\left(\varphi, \vartheta\right) = \begin{pmatrix} \cos\left(\varphi\right)\sin\left(\vartheta\right) \\ \sin\left(\varphi\right)\sin\left(\vartheta\right) \\ \cos\left(\vartheta\right) \end{pmatrix},$$

where $\varphi \in [0, 2\pi)$ is the azimuth and $\vartheta \in [0, \pi]$ is the polar angle. We denote with $\boldsymbol{e}_i$ the i-th unit vector and write $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3$. In the spherical coordinate system, the integral of a function $f : \mathbb{S}^2 \to \mathbb{C}$ on the sphere is parametrized by

$$\int_{\mathbb{S}^2} f\left(\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} = \int_0^{\pi} \int_0^{2\pi} f\left(\cos\varphi \sin\vartheta, \sin\varphi \sin\vartheta, \cos\vartheta\right) \mathrm{d}\varphi \sin\vartheta \mathrm{d}\vartheta. \tag{2.30}$$

The measure of the whole surface of the sphere is 4π.

A function $f : \mathbb{S}^2 \to \mathbb{C}$ is called a radial function with respect to some vector $\boldsymbol{y} \in \mathbb{S}^2$ if there exists a function $g : [-1, 1] \to \mathbb{C}$ such that $f\left(\boldsymbol{x}\right) = g\left(\langle \boldsymbol{x}, \boldsymbol{y}\rangle\right)$ for all $\boldsymbol{x} \in \mathbb{S}^2$. Since the surface measure is invariant under rotations, we are free in the choice a coordinate system and we choose one where $\boldsymbol{y}$ is the north pole $(0, 0, 1)$ of the sphere. Then the integral over f equals

$$\int_{\mathbb{S}^2} f\left(\boldsymbol{x}\right)\, \mathrm{d}\boldsymbol{x} = \int_0^{\pi} \int_0^{2\pi} g\left(\cos\vartheta\right)\, \mathrm{d}\varphi \sin\vartheta\, \mathrm{d}\vartheta = \int_{-1}^{1} g(t)\, \mathrm{d}t. \tag{2.31}$$

2.2.1 Harmonic polynomials on the sphere

In order to define an orthonormal basis of polynomials in the space $L^2\left(\mathbb{S}^2\right)$, we follow the approach from [7, Charpter 1] throughout this section.

We denote the set of all polynomials in three variables having degree up to n with $\Pi_n(\mathbb{R}^3)$. A polynomial $p : \mathbb{R}^3 \to \mathbb{R}$ is called homogeneous of degree n if p is a linear combination of monomials of degree n. For a multi-index $\alpha \in \mathbb{N}_0^3$, a monomial is $\boldsymbol{x} \mapsto \boldsymbol{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$ and has degree $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$. An equivalent definition is the following: we call a polynomial p homogeneous of degree n if $p(r\boldsymbol{x}) = r^n p(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^3$ and $r \in \mathbb{R}$. Let $\mathscr{P}_n(\mathbb{R}^3)$ denote all polynomials homogeneous of degree n. A polynomial $p \in \mathscr{P}_n(\mathbb{R}^3)$ is called harmonic if $\Delta p = 0$ where Δ denotes the Laplacian

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

The Laplacian restricted to the two-sphere is often called the Laplace-Beltrami operator. Let $\mathscr{H}_n(\mathbb{S}^2)$ denote the set of all harmonic polynomials, homogenous of degree n on $\mathbb{R}^3$, restricted to the sphere. A function $Y_n \in \mathscr{H}_n(\mathbb{S}^2)$ is called a spherical harmonic.

Proposition 2.4. *The harmonic spaces $\mathscr{H}_n(\mathbb{S}^2)$ are orthogonal with respect to n, i.e. if $Y_n \in \mathscr{H}_n(\mathbb{S}^2)$ and $Y_m \in \mathscr{H}_m(\mathbb{S}^2)$ with $m \neq n$, then $\langle Y_n, Y_m \rangle = 0$ where $\langle \cdot, \cdot \rangle$ denotes the L^2 inner product.*

Proof. Since Y_n is homogeneous, $Y_n(r\boldsymbol{x}) = r^{n-1} Y_n(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{S}^2$ and consequently

$$\frac{\partial Y_n}{\partial r}(r\boldsymbol{x}) = n r^{n-1} Y_n(\boldsymbol{x}),$$

where ∂ denotes the partial derivative. Using Green's second identity, we can conclude

$$\begin{aligned} (m-n)\langle Y_n, Y_m \rangle &= (m-n) \int_{\mathbb{S}^2} Y_n Y_m \,\mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{S}^2} \left(Y_n \frac{\partial Y_m}{\partial r} - Y_m \frac{\partial Y_n}{\partial r} \right) \mathrm{d}\boldsymbol{x} \\ &= \int_{B_1(0)} (Y_n \Delta Y_m - Y_m \Delta Y_n) \,\mathrm{d}\boldsymbol{x} = 0. \end{aligned}$$

□

Proposition 2.5. *The homogeneous polynomials can be decomposed into harmonic polynomials through*

$$\mathscr{P}_n\left(\mathbb{S}^2\right) = \bigoplus_{j=0}^{n/2} \|\boldsymbol{x}\|^{2j} \mathscr{H}_{n-2j}\left(\mathbb{S}^2\right).$$

Proof. We show the result using induction over $n \in \mathbb{N}_0$. The basis is clear since $\mathscr{H}_0 = \mathscr{P}_0$ and $\mathscr{H}_1 = \mathscr{P}_1$. For $n \geq 2$ we note that $\Delta \mathscr{P}_n \subset \mathscr{P}_{n-2}$ and thus

$$\dim \mathscr{P}_n \leq \dim \mathscr{H}_n + \dim \mathscr{P}_{n-2}.$$

From the induction hypothesis, it follows that

$$\|\boldsymbol{x}\|^2 \mathscr{P}_{n-2}\left(\mathbb{S}^2\right) = \bigoplus_{j=0}^{n/2} \|\boldsymbol{x}\|^{2j+2} \mathscr{H}_{n-2j-2}\left(\mathbb{S}^2\right).$$

Because of the orthogonality of the harmonic polynomials shown in Proposition 2.4,

$$\dim \mathscr{P}_n \geq \dim \mathscr{H}_n + \dim \mathscr{P}_{n-2}.$$

Hence, we have

$$\mathscr{P}_n = \mathscr{H}_n \oplus \|\boldsymbol{x}\|^2 \mathscr{P}_{n-2}.$$

□

With the help of Proposition 2.5, we can calculate the dimension of the harmonic polynomial spaces $\mathscr{H}_n$: The dimension of the space of homogeneous polynomials $\mathscr{P}_n$ is

$$\dim \mathscr{P}_n = \begin{pmatrix} n+2 \\ n \end{pmatrix} = \frac{(n+2)(n+1)}{2},$$

which can be seen by counting the number of three-dimensional multi-indices α with $|\alpha| = n$. Then

$$\dim \mathscr{H}_n = \dim \mathscr{P}_n - \dim \mathscr{P}_{n-2} = \frac{(n+2)(n+1) - n(n-1)}{2} = 2n+1.$$

Let $Y_{n,k}$, $k = -n, \dots, n$ be an orthonormal basis of $\mathscr{H}_n$. Proposition 2.5 implies that every polynomial on the sphere can be written as a linear combination of spherical harmonics $Y_{n,k}$. Since the sphere is compact, the Stone-Weierstrass theorem states that every continuous function can be approximated uniformly by spherical harmonics.

Zonal harmonics. For $n \in \mathbb{N}_0$, the reproducing kernel $Z_n : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{C}$ of $\mathscr{H}_n$ is solely defined by the reproducing property

$$Y_n(\boldsymbol{x}) = \int_{\mathbb{S}^2} Y_n(\boldsymbol{y}) Z_n(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \quad \text{for all } Y_n \in \mathscr{H}_n,\ \boldsymbol{x} \in \mathbb{S}^2. \tag{2.32}$$

The existence and uniqueness of the reproducing kernel follows from Riesz's representation theorem applied to the linear functional $\mathscr{H}_n \to \mathbb{C}$, $Y_n \mapsto Y_n(\boldsymbol{y})$ for every $\boldsymbol{y} \in \mathbb{S}^2$. Let $Y_{n,k}$, $k = -n, \dots, n$ be an orthonormal basis of $\mathscr{H}_n$, then it can be easily verified by inserting into (2.32) that

$$Z_n(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=-n}^{n} Y_{n,k}(\boldsymbol{x}) \overline{Y_{n,k}(\boldsymbol{y})}, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2. \tag{2.33}$$

This shows that the zonal harmonics Z_n are polynomials of degree n in both arguments. Note that (2.33) is independent from the particular choice of the basis $Y_{n,k}$ since the reproducing kernel is unique.

Proposition 2.6. *For $n \in \mathbb{N}_0$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2$, the reproducing kernel $Z_n(\boldsymbol{x}, \boldsymbol{y})$ depends only on the geodesic distance of the arguments $\boldsymbol{x}$ and $\boldsymbol{y}$. It can be written as a multiple of the Legendre polynomials by*

$$Z_n(\boldsymbol{x}, \boldsymbol{y}) = \frac{2n+1}{4\pi} P_n(\langle \boldsymbol{x}, \boldsymbol{y} \rangle), \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2. \tag{2.34}$$

Proof. At first, we show that $Z_n(\boldsymbol{x}, \boldsymbol{y})$ depends only on the geodesic distance of $\boldsymbol{x}$ and $\boldsymbol{y}$. We observe that the space $\mathscr{H}_n$ is invariant under rotations, because the Laplacian Δ is invariant under rotations. Let $SO(3)$ denote the three-dimensional rotation group which is the set of all 3×3-dimensional orthogonal matrices with determinant 1. For a rotation matrix $Q \in SO(3)$, the functions $Y_{n,k}(Q\circ)$, $k = -n, \ldots, n$ form an orthonormal basis of $\mathscr{H}_n$, so we have $Z_n(Q\circ, Q\circ) = Z_n$. For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2$ there exists a rotation matrix Q such that

$$Q\boldsymbol{x} = (0,0,1) \text{ and } Q\boldsymbol{y} = \left(0, \sqrt{1-t^2}, t\right)$$

for some $t \in [-1,1]$. Because the reproducing kernel Z_n is unique, this shows that Z_n only depends on $t = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ and we write

$$F_n(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) = Z_n(\boldsymbol{x}, \boldsymbol{y}).$$

The zonal harmonic Z_n is a polynomial of degree n in both arguments. So the function

$$Z_n(\boldsymbol{x}, \boldsymbol{e}_3) = F_n(\langle \boldsymbol{x}, \boldsymbol{e}_3 \rangle) = F_n(x_3)$$

is a polynomial of degree n in $\boldsymbol{x}$ and since it only depends on x_3, we obtain that F_n is a polynomial of degree n.

Now we show the orthogonality of F_n on the interval $[-1,1]$. For $n, n' \in \mathbb{N}_0$ with $n \neq n'$, we have for some $\boldsymbol{y} \in \mathbb{S}^2$ by (2.31) and (2.33)

$$\begin{aligned}
\int_{-1}^{1} F_n(t)\overline{F_{n'}(t)}\,\mathrm{d}t &= \int_{\mathbb{S}^2} Z_n(\boldsymbol{x}, \boldsymbol{y})\,\overline{Z_{n'}(\boldsymbol{x}, \boldsymbol{y})}\,\mathrm{d}\boldsymbol{x} \\
&= \sum_{k=-n}^{n} \sum_{k'=-n'}^{n'} \overline{Y_{n,k}(\boldsymbol{y})} Y_{n',k'}(\boldsymbol{y}) \int_{\mathbb{S}^2} Y_{n,k}(\boldsymbol{x})\,\overline{Y_{n',k'}(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} \\
&= 0
\end{aligned}$$

where we have used the orthogonality property of the spherical harmonics $Y_{n,k}$ from Proposition 2.4.

Next, we show the normalization of F_n. For any $\boldsymbol{y} \in \mathbb{S}^2$, we have

$$\begin{aligned}
\int_{-1}^{1} |F_n(t)|^2\,\mathrm{d}t &= \sum_{k,k'=-n}^{n} \overline{Y_{n,k}(\boldsymbol{y})} Y_{n,k'}(\boldsymbol{y}) \int_{\mathbb{S}^2} Y_{n,k}(\boldsymbol{x})\,\overline{Y_{n,k'}(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} \\
&= \sum_{k=-n}^{n} \overline{Y_{n,k}(\boldsymbol{y})} Y_{n,k}(\boldsymbol{y}) \\
&= \sum_{k=-n}^{n} |Y_{n,k}(\boldsymbol{y})|^2 = Z_n(\boldsymbol{y}, \boldsymbol{y}) = F_n(1).
\end{aligned}$$

Thus, the function $\sum_{k=-n}^{n} |Y_{n,k}|^2$ must be constant. Again by of the assumed orthonormality of the spherical harmonics $Y_{n,k}$, we have

$$\begin{aligned}\sum_{k=-n}^{n} |Y_{n,k}(\boldsymbol{y})|^2 &= \sum_{k=-n}^{n} \frac{1}{4\pi} \int_{\mathbb{S}^2} |Y_{n,k}(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} \\ &= \frac{2n+1}{4\pi}.\end{aligned}$$

Putting the things together, this shows that the polynomials F_n are orthogonal on the interval $[-1,1]$ with respect to the Lebesgue measure. Because the orthogonal polynomials are unique, the functions F_n must be equal to the Legendre polynomials P_n multiplied by a constant factor depending only on n. The factor can be determined by using $F_n(1) = (2n+1)/(4\pi)$ and $P_n(1) = 1$. □

Combining formulas (2.33) and (2.34), we obtain the so-called addition theorem for spherical harmonics:

$$\sum_{k=-n}^{n} Y_{n,k}(\boldsymbol{x}) Y_{n,k}(\boldsymbol{y}) = \frac{2n+1}{4\pi} P_n(\langle \boldsymbol{x}, \boldsymbol{y} \rangle), \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2, \ n \in \mathbb{N}_0. \tag{2.35}$$

2.2.2 Spherical harmonics in spherical coordinates

In the following, we give an explicit representation of the spherical harmonics in spherical coordinates shown in [40]. Therefore, we use the associated Legendre polynomials

$$P_{n,k} : [-1,1] \to \mathbb{R}, \quad n \in \mathbb{N}_0, \ k = -n, \dots, n,$$

which are given in terms of the Legendre polynomials for $k \geq 0$ by

$$P_{n,k}(t) = (-1)^k \left(1 - t^2\right)^{k/2} \frac{\mathrm{d}^k}{\mathrm{d}t^k} P_n(t), \quad t \in [-1,1], \tag{2.36}$$

and

$$P_{n,-k}(t) = (-1)^k \frac{(n-k)!}{(n+k)!} P_{n,k}(t), \quad t \in [-1,1]. \tag{2.37}$$

They satisfy the three-term recurrence relation (see [20, 8.735.2])

$$\sqrt{1-t^2} P_{n,k+1}(t) = (n-k)\, t P_{n,k}(t) - (n+k) P_{n-1,k}(t), \tag{2.38}$$

which is initialized for $k = 0$ by the Legendre polynomials $P_n = P_{n,0}$. The spherical harmonics $Y_{n,k} \in \mathscr{H}_n(\mathbb{S}^2)$ are defined in spherical coordinates by

$$Y_{n,k}(\boldsymbol{x}(\varphi, \vartheta)) = \sqrt{\frac{2n+1}{4\pi}} \sqrt{\frac{(n-k)!}{(n+k)!}} P_{n,k}(\cos\vartheta) \, \mathrm{e}^{\mathrm{i}k\varphi}, \quad n \in \mathbb{N}_0, \ k = -n, \dots, n. \tag{2.39}$$

The so-defined spherical harmonics $Y_{n,k}$, $k = -n, \dots, n$ form an orthonormal basis of $\mathscr{H}_n(\mathbb{S}^2)$. Therefore, every function $f \in L^2(\mathbb{S}^2)$ can be represented by a Fourier series with respect to the spherical harmonics as it was discussed in Subsection 2.2.1 for the general case.

2.2.3 Sobolev spaces on the sphere

We introduce function spaces containing functions of a certain smoothness, so-called Sobolev spaces. More background information about Sobolev spaces on the sphere can be found in [15, section 5.1].

Definition 2.7. For $s \geq 0$, the Sobolev space $H^s(\mathbb{S}^2)$ is defined by

$$H^s\left(\mathbb{S}^2\right) := \left\{ f \in L^2\left(\mathbb{S}^2\right) : \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \left(n + \frac{1}{2}\right)^{2s} \left|\hat{f}(n,k)\right|^2 < \infty \right\}. \tag{2.40}$$

The Sobolev space is equipped with the inner product

$$\langle f, g \rangle_s := \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \left(n + \frac{1}{2}\right)^{2s} \hat{f}(n,k)\, \overline{\hat{g}(n,k)}$$

and the induced norm $\|f\|_{H^s(\mathbb{S}^2)} = \sqrt{\langle f, f \rangle_s}$.

The Sobolev space $\left(H^s(\mathbb{S}^2), \langle \cdot, \cdot \rangle_s\right)$ is a Hilbert space. The Sobolev spaces are nested, $H^s(\mathbb{S}^2) \subset H^t(\mathbb{S}^2)$ for $s > t$. The equality $H^0(\mathbb{S}^2) = L^2(\mathbb{S}^2)$ follows directly from the definition. If s is an integer, then $H^s(\mathbb{S}^2)$ can be thought of as the space containing all functions that are square-integrable together with their weak derivatives of the order s.

Proposition 2.8 (Sobolev embedding theorem, cf. [25])**.** *If $s > 1$, the Sobolev space $H^s(\mathbb{S}^2)$ can be embedded continuously into the space of continuous functions $C(\mathbb{S}^2)$. Furthermore, there exists a constant $c > 0$ independent of f such that*

$$\|f\|_{C(\mathbb{S}^2)} \leq c\, \|f\|_{H^s(\mathbb{S}^2)}, \quad f \in C\left(\mathbb{S}^2\right).$$

Proof. Let $f \in H^s(\mathbb{S}^2)$ with $s > 1$ and $\boldsymbol{x} \in \mathbb{S}^2$. Then, by the Cauchy-Schwartz inequality and the addition theorem (2.35), we have for $N \in \mathbb{N}_0$

$$\begin{aligned}
\left| \sum_{n=0}^{N} \sum_{k=-n}^{n} \hat{f}(n,k)\, Y_{n,k}(\boldsymbol{x}) \right| &\leq \sum_{n=0}^{N} \sum_{k=-n}^{n} \left| \hat{f}(n,k) \left(n + \frac{1}{2}\right)^{s} \right| \left|Y_{n,k}(\boldsymbol{x})\right| \left(n + \frac{1}{2}\right)^{-s} \\
&\leq \|f\|_{H^s(\mathbb{S}^2)} \sqrt{\sum_{n=0}^{N} \sum_{k=-n}^{n} \left|Y_{n,k}(\boldsymbol{x})\right|^2 \left(n + \frac{1}{2}\right)^{-2s}} \\
&\leq \|f\|_{H^s(\mathbb{S}^2)} \sqrt{\sum_{n=0}^{N} \frac{2n+1}{4\pi} \left(n + \frac{1}{2}\right)^{-2s}} \\
&= \|f\|_{H^s(\mathbb{S}^2)} \sqrt{\sum_{n=0}^{N} \frac{1}{2\pi} \left(n + \frac{1}{2}\right)^{1-2s}}.
\end{aligned}$$

The sum converges for $N \to \infty$ if and only if $1 - 2s < -1$, or equivalently $s > 1$. This shows that the Fourier series of f converges uniformly to a continuous function provided $s > 1$. Since the Fourier series converges to f in the L^2-norm, we obtain that

$$\begin{aligned} |f(\boldsymbol{x})| &= \left| \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \hat{f}(n,k)\, Y_{n,k}(\boldsymbol{x}) \right| \\ &\leq \|f\|_{H^s(\mathbb{S}^2)} \sqrt{\sum_{n=0}^{\infty} \frac{1}{2\pi} \left(n + \frac{1}{2}\right)^{1-2s}}. \end{aligned}$$

□

2.2.4 Spherical convolution

The spherical convolution $\psi \star f$ of $\psi \in L^2([-1,1])$ with a spherical function $f \in L^2(\mathbb{S}^2)$ is defined as

$$\psi \star f(\boldsymbol{x}) = \int_{\mathbb{S}^2} \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) f(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}, \quad \boldsymbol{x} \in \mathbb{S}^2,$$

cf. [7, Section 2.1]. In this definition, $\boldsymbol{y} \mapsto \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)$ can be taken as a radial function on the sphere. The spherical convolution satisfies the following Young inequality:

Proposition 2.9. *The spherical convolution* $\star : L^2([-1,1]) \times L^2(\mathbb{S}^2) \to L^2(\mathbb{S}^2)$, $(\psi, f) \mapsto \psi \star f$ *is bilinear and bounded operator satisfying*

$$\|\psi \star f\|_{L^2(\mathbb{S}^2)} \leq \|\psi\|_{L^2([-1,1])} \|f\|_{L^2(\mathbb{S}^2)}.$$

Proof. Let $\psi \in L^2([-1,1])$ and $f \in L^2(\mathbb{S}^2)$, then, by Fubini's theorem,

$$\begin{aligned} \|\psi \star f\|_{L^2(\mathbb{S}^2)}^2 &= \int_{\mathbb{S}^2} \left| \int_{\mathbb{S}^2} \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) f(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \right|^2 \mathrm{d}\boldsymbol{x} \\ &\leq \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} |\psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) f(\boldsymbol{y})|^2 \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{S}^2} \left(\int_{\mathbb{S}^2} |\psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)|^2 \, \mathrm{d}\boldsymbol{x} \right) |f(\boldsymbol{y})|^2 \, \mathrm{d}\boldsymbol{y}. \end{aligned}$$

Since $\boldsymbol{x} \mapsto |\psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)|^2$ is a radial function, we can use (2.31) to evaluate the inner integral

$$\int_{\mathbb{S}^2} |\psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)|^2 \, \mathrm{d}\boldsymbol{x} = \int_{-1}^{1} |\psi(t)|^2 \, \mathrm{d}t = \|\psi\|_{L^2([-1,1])}^2.$$

Hence,

$$\|\psi \star f\|_{L^2(\mathbb{S}^2)}^2 \leq \|\psi\|_{L^2([-1,1])}^2 \int_{\mathbb{S}^2} |f(\boldsymbol{y})|^2 \, \mathrm{d}\boldsymbol{y}.$$

□

The following theorem is known as the Funk-Hecke formula, c.f. [7, Theorem 1.2.9].

Proposition 2.10. *Let* $\psi \in L^2([-1,1])$ *and* $Y_n \in \mathscr{H}_n(\mathbb{S}^2)$*, then*

$$\psi \star Y_n(\boldsymbol{x}) = \int_{\mathbb{S}^2} \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) Y_n(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \hat{\psi}(n) Y_n(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{S}^2 \tag{2.41}$$

with

$$\hat{\psi}(n) = 2\pi \int_{-1}^{1} \psi(t) P_n(t) \, \mathrm{d}t = 2\pi \langle \psi, P_n \rangle_{L^2[-1,1]}, \quad n \in \mathbb{N}_0. \tag{2.42}$$

Proof. At first, we assume that ψ is a polynomial of degree $N \in \mathbb{N}$, so we expand ψ in terms of the Legendre polynomials,

$$\psi(t) = \sum_{m=0}^{N} \hat{\psi}(m) \frac{2m+1}{4\pi} P_m(t) = \sum_{m=0}^{N} \hat{\psi}(m) F_m(t).$$

From the reproducing property (2.32) of Z_n, it follows that

$$\begin{aligned} \psi \star Y_n(\boldsymbol{x}) &= \int_{\mathbb{S}^2} Y_n(\boldsymbol{y}) \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) \, \mathrm{d}\boldsymbol{y} \\ &= \sum_{m=0}^{N} \hat{\psi}(m) \int_{\mathbb{S}^2} Y_n(\boldsymbol{y}) F_m(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) \, \mathrm{d}\boldsymbol{y} \\ &= \sum_{m=0}^{N} \hat{\psi}(m) \int_{\mathbb{S}^2} Y_n(\boldsymbol{y}) Z_m(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \\ &= \hat{\psi}(n) Y_n(\boldsymbol{x}). \end{aligned}$$

Now, let ψ be an arbitrary function in $L^2([-1,1])$, then ψ can be approximated with respect to the L^2 norm by polynomials, which follows from the Weierstrass approximation theorem and the fact that the set of continuous function is dense in $L^2([-1,1])$. Together with the boundedness of the convolution shown in Proposition 2.9, this implies (2.41). □

A simple corollary from Proposition 2.10 is the equation

$$\psi \star f = \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \hat{\psi}(n) \hat{f}(n,k) Y_{n,k} \tag{2.43}$$

where $f \in L^2(\mathbb{S}^2)$, $\psi \in L^2[-1,1]$, and $\hat{\psi}(n)$ is defined in (2.42). So the spherical convolution of two functions is just the multiplication of their Fourier coefficients. The identity (2.43) is an analogue to the convolution theorem for functions on the torus.

Convolutions can also be used for approximating functions. Therefore, we convolute a spherical function $f \in L^2(\mathbb{S}^2)$ with a mollifier $\psi \in L^2([-1,1])$, such that $\psi \star f$ is "near" to f. A sequence of such such mollifiers $(\psi_j)_{j\in\mathbb{N}}$ is called an approximate identity if for all $f \in L^2(\mathbb{S}^2)$,

$$\lim_{j\to\infty} \|f - \psi_j \star f\|_{L^2(\mathbb{S}^2)} = 0.$$

An equivalent characterization of approximate identities is that

$$\lim_{j\to\infty} \hat{\psi}_j(n) = 1 \text{ for all } n \in \mathbb{N}_0,$$

see [15, Section 8.1]. The following theorem gives an upper bound for the approximation.

Proposition 2.11. *Let $f \in H^s(\mathbb{S}^2)$ for $s \geq 0$ and $\psi \in L^2([-1,1])$, then the error of the approximation of f by $\psi \star f$ is bounded by*

$$\|f - \psi \star f\|_{L^2(\mathbb{S}^2)} \leq \sup_{n\in\mathbb{N}_0} \left(\frac{\left|1 - \hat{\psi}(n)\right|}{\left(n + \frac{1}{2}\right)^s} \right) \|f\|_{H^s(\mathbb{S}^2)}$$

where $\hat{\psi}(n)$ is defined in (2.42)*.*

Proof. From Parseval's identity (2.2) and the Funk-Hecke formula (2.41), we obtain

$$\begin{aligned} \|f - \psi \star f\|^2_{L^2(\mathbb{S}^2)} &= \sum_{n=0}^{\infty} \sum_{k=-n}^{n} |\langle f - \psi \star f, Y_{n,k}\rangle|^2 \\ &= \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \left|\hat{f}(n,k) - \hat{\psi}(n)\hat{f}(n,k)\right|^2 \\ &= \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \frac{\left|1 - \hat{\psi}(n)\right|^2}{\left(n + \frac{1}{2}\right)^{2s}} \left(n + \frac{1}{2}\right)^{2s} \left|\hat{f}(n,k)\right|^2. \end{aligned} \tag{2.44}$$

Now we can write the supremum of the first factor in (2.44) left of the sums, insert the H^s-norm of f from Definition 2.7, and we see that

$$\|f - \psi \star f\|^2_{L^2(\mathbb{S}^2)} \leq \sup_{n\in\mathbb{N}_0} \left(\frac{\left|1 - \hat{\psi}(n)\right|^2}{\left(n + \frac{1}{2}\right)^{2s}} \right) \|f\|^2_{H^s(\mathbb{S}^2)}.$$

□

3 Circular Averages on the Sphere

3.1 The mean operator on the sphere

In this section, we describe properties of the mean operator on the sphere $\mathcal{M}$ that computes the mean values along circles of functions defined on the two-dimensional unit sphere $\mathbb{S}^2$. Every circle on the sphere can be described as the intersection of $\mathbb{S}^2$ with a plane $\{\boldsymbol{y} \in \mathbb{R}^3 : \langle \boldsymbol{x}, \boldsymbol{y} \rangle = t\}$ where $\boldsymbol{x} \in \mathbb{S}^2$ and $t \in [-1, 1]$. The diameter of the sphere that is perpendicular to the circle is called the axis of the circle. The endpoints of this axis, i.e. $\boldsymbol{x}$ and $-\boldsymbol{x}$, are called the poles of the circle, cf. [27].

The mean operator $\mathcal{M} : C(\mathbb{S}^2) \to C(\mathbb{S}^2 \times [-1, 1])$ is defined for every function $f \in C(\mathbb{S}^2)$ by

$$\mathcal{M}f(\boldsymbol{x}, t) = \begin{cases} \frac{1}{2\pi\sqrt{1-t^2}} \int_{\langle \boldsymbol{x}, \boldsymbol{y} \rangle = t} f(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}), & \boldsymbol{x} \in \mathbb{S}^2, t \in (-1, 1) \\ f(\boldsymbol{x}, \pm t), & \boldsymbol{x} \in \mathbb{S}^2,\ t = \pm 1 \end{cases} \tag{3.1}$$

(cf. [7, Section 2.1], [48]) where $\mathrm{d}s$ denotes the arc length along the circle $\{\boldsymbol{y} \in \mathbb{S}^2 : \langle \boldsymbol{x}, \boldsymbol{y} \rangle = t\}$, which has the poles $\boldsymbol{x}$ and $-\boldsymbol{x}$ the radius $\sqrt{1-t^2}$. So the spherical mean $\mathcal{M}f(\boldsymbol{x}, t)$ is just the mean value of f on that circle. For $t = \pm 1$, the set $\{\boldsymbol{y} \in \mathbb{S}^2 : \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \pm 1\}$ consists only of the single point $\boldsymbol{x}$ on the sphere, so $\mathcal{M}f(\boldsymbol{x}, \pm 1) = f(\pm \boldsymbol{x})$ can be seen as the “mean value” of f on the set $\{\pm \boldsymbol{x}\}$.

The factor $2\pi\sqrt{1-t^2}$ in (3.1) is equal to the circumference of a circle $\{\boldsymbol{y} \in \mathbb{S}^2 : \langle \boldsymbol{x}, \boldsymbol{y} \rangle = t\}$. Hence, the mean operator applied to a constant function $f \equiv c \in \mathbb{C}$ gives a function that is constant in both $\boldsymbol{x}$ and t and has the same value as f. Moreover, the following symmetry property of the mean operator on the sphere is easy to check,

$$\mathcal{M}f(\boldsymbol{x}, t) = \mathcal{M}f(-\boldsymbol{x}, -t), \quad (\boldsymbol{x}, t) \in \mathbb{S}^2 \times [-1, 1].$$

As a consequence of (2.31), the integration over the whole sphere can be written with the help of the mean operator on the sphere as

$$\int_{\mathbb{S}^2} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 2\pi \int_{-1}^{1} \mathcal{M}f(\boldsymbol{y},t)\,\mathrm{d}t \tag{3.2}$$

for any $\boldsymbol{y} \in \mathbb{S}^2$.

Theorem 3.1 ([4, Section 4.2], [7, Section 2.1])**.** *Let $f \in L^2(\mathbb{S}^2)$, $Y_n \in \mathscr{H}_n(\mathbb{S}^2)$ and $t \in [-1,1]$, then*

$$\langle \mathcal{M}f(\cdot,t), Y_n\rangle = P_n(t)\langle f, Y_n\rangle, \quad n \in \mathbb{N}_0. \tag{3.3}$$

Proof. Let $g \in L^2[-1,1]$ and $n \in \mathbb{N}_0$. For a spherical harmonic $Y_n \in \mathscr{H}_n$, we have

$$\begin{aligned} g \star Y_n(\boldsymbol{x}) &= \int_{\mathbb{S}^2} g(\langle \boldsymbol{x},\boldsymbol{y}\rangle)\, Y_n(\boldsymbol{y})\ \mathrm{d}\boldsymbol{y} \\ &= \int_{-1}^{1} \int_{\langle \boldsymbol{x},\boldsymbol{y}\rangle = t} g(t)\, Y_n(\boldsymbol{y})\ \mathrm{d}s(\boldsymbol{y})\ \mathrm{d}t \\ &= 2\pi \int_{-1}^{1} g(t)\,\mathcal{M}Y_n(\boldsymbol{x},t)\ \mathrm{d}t, \quad \boldsymbol{x} \in \mathbb{S}^2. \end{aligned} \tag{3.4}$$

Using the Funk-Hecke formula (2.41), we have on the other side

$$g \star Y_n = 2\pi Y_n \int_{-1}^{1} g(t) P_n(t)\,\mathrm{d}t. \tag{3.5}$$

Combining the equations (3.4) and (3.5) yields

$$\int_{-1}^{1} g(t)\,\mathcal{M}Y_n(\cdot,t)\ \mathrm{d}t = Y_n \int_{-1}^{1} g(t)P_n(t)\,\mathrm{d}t.$$

Because this equality is valid for all $g \in L^2[-1,1]$, we can conclude

$$\mathcal{M}Y_n(\cdot,t) = P_n(t) Y_n \tag{3.6}$$

for almost all $t \in [-1,1]$. Since both sides of (3.6) are continuous with respect to t, the equality must hold for all t. Finally (3.3) follows by the Fourier decomposition of f. □

Theorem 3.1 can be seen as a kind of generalization of the Funk-Hecke formula (2.41). Formally, inserting the δ distribution into the Funk-Hecke formula, $\psi = \delta((\cdot) - t)$, would yield (3.3) immediately, but the delta distribution does not lie in $L^2[-1,1]$.

The singular value decomposition of the mean operator on the sphere follows from Theorem 3.1,

$$\mathcal{M}f(\boldsymbol{x},t) = \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \hat{f}(n,k)\, P_n(t)\, Y_{n,k}(\boldsymbol{x}).$$

Restrictions of the mean operator on the sphere. When considering the reconstruction of a function f where $\mathcal{M}f$ is given, we have more information available than we actually need. To see that, we assume that we know $\mathcal{M}f(\boldsymbol{x}, 1)$ for all $\boldsymbol{x} \in \mathbb{S}^2$. Then, by definition, we have already full information about the function $f \equiv \mathcal{M}f(\cdot, 1)$. So it makes sense to consider the inversion of $\mathcal{M}$ when we only know the spherical means on a submanifold of the image space $\mathbb{S}^2 \times [-1, 1]$.

A well-known example is the so-called spherical Radon or Funk transform, where we fix $t = 0$, cf. [16, 8]. In that case, we integrate f along all great circles of the sphere. Obviously, all odd functions lie in the nullspace of the Funk transform. It has been shown that any even function $f \in L^2(\mathbb{S}^2)$ can be reconstructed from its Funk transform. The translation operator is a slight generalization to the Funk transform, cf. [48]. Instead of $t = 0$, we fix an arbitrary $t_0 \in [-1, 1]$.

Another example was considered in [24, II.1.C]. For the spherical slice transform, a function is integrated along all circles passing through the north pole. Formally, the spherical slice transform of a function f is given by $\boldsymbol{x} \mapsto \mathcal{M}f(\boldsymbol{x}, \langle \boldsymbol{x}, \boldsymbol{e}_3 \rangle)$ for $\boldsymbol{x} \in \mathbb{S}^2$ with $x_3 \geq 0$.

In the following, we will take a closer look at the circular average transform, which computes the integrals along all circles that have their poles on the equator.

3.2 The circular average transform

In order to define the circular average transform, we introduce some notation. Let $\sigma \in \mathbb{T}$ be a point on the one-dimensional torus $\mathbb{T} = [0, 2\pi)$. Now we denote with

$$\boldsymbol{e}_\sigma = (\cos\sigma, \sin\sigma, 0)^\top \in \mathbb{S}^2$$

the point on the equator of the two-dimensional sphere $\mathbb{S}^2$ with longitude σ.

Definition 3.2. The *circular average transform* $\mathcal{T}f$ of a continuous function $f \in C(\mathbb{S}^2)$ is defined by

$$\mathcal{T}f(\sigma, t) = \mathcal{M}f(\boldsymbol{e}_\sigma, t), \quad (\sigma, t) \in \mathbb{T} \times [-1, 1], \tag{3.7}$$

where $\mathcal{M}$ is the mean operator from (3.1).

The circular average transform $\mathcal{T}f(\sigma, t)$ computes the mean values of a function f along the circles

$$C(\sigma, t) = \left\{ \boldsymbol{y} \in \mathbb{S}^2 : \langle \boldsymbol{y}, \boldsymbol{e}_\sigma \rangle = t \right\}. \tag{3.8}$$

We denote with $\mathscr{C} = \{C(\sigma, t) : (\sigma, t) \in \mathbb{T} \times [-1, 1]\}$ the set of all these circles. Then $\mathscr{C}$ contains exactly those circles of the sphere whose poles are located on the equator of the sphere.

For $(\sigma, t) \in \mathbb{T} \times [-1, 1]$, the small circles $C(\sigma, t)$ and $C(\sigma + \pi, -t)$ are equal. With that in mind, we define the equivalence relation $\sim$ on the domain $\mathbb{T} \times [-1, 1]$ by saying that two points (σ, t) and (σ', t') are equivalent if $t = -t'$ and $\sigma = \sigma' + \pi + 2k\pi$ for some integer k.

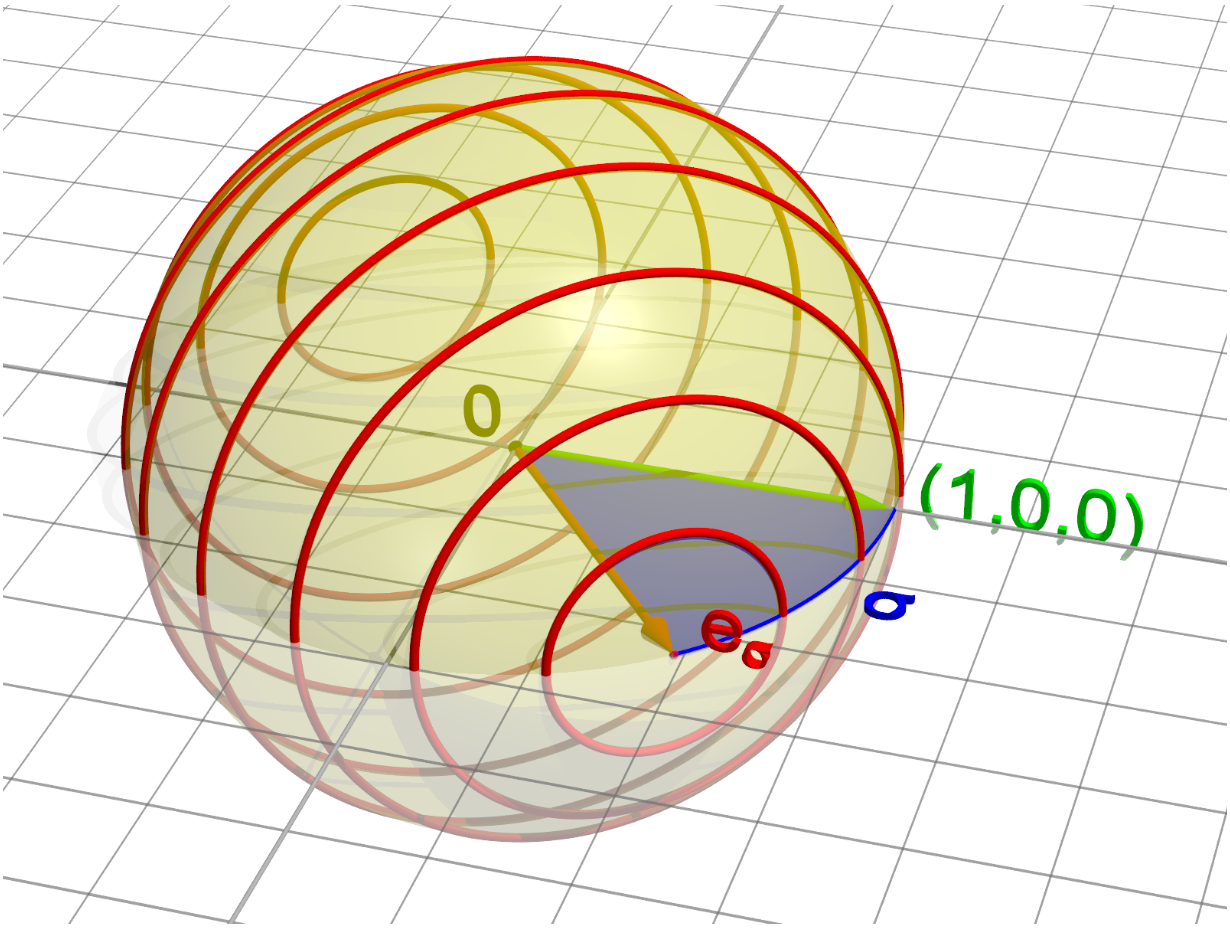

Figure 3.1: Red: Circles $C\,(\sigma, t)$ for a fixed $\sigma = 5\pi/18$. The plane with the grid is the x_1–x_2 plane.

So the set of circles $\mathscr{C}$ is topologically equivalent to the quotient set $(\mathbb{T} \times [-1,1])\,/\sim$. This set is called a Möbius band, which was first described in 1886, see [39]. It can be described as a rectangle $[0,\pi] \times [-1,1]$ whose lower and upper edge are identified via $(0,t) \sim (\pi,-t)$ for $t \in [-1,1]$, see Figure 3.2 on page 37.

Proposition 3.3. *Let* $(\sigma, t) \in \mathbb{T} \times [-1,1]$. *The small circles* $C(\sigma,t)$ *given in* (3.8) *have the parameterized representation*

$$C(\sigma,t) = \left\{ t \begin{pmatrix} \cos\sigma \\ \sin\sigma \\ 0 \end{pmatrix} + \sqrt{1-t^2} \begin{pmatrix} -\sin\sigma\cos\varphi \\ \cos\sigma\cos\varphi \\ \sin\varphi \end{pmatrix} : \varphi \in \mathbb{T} \right\}. \tag{3.9}$$

Proof. For $\boldsymbol{x} = (x_1, x_2, x_3)^\top \in \mathbb{S}^2$, we obtain from (3.8)

$$\boldsymbol{x} \in C(\sigma,t) \Leftrightarrow x_1\cos\sigma + x_2\sin\sigma = t.$$

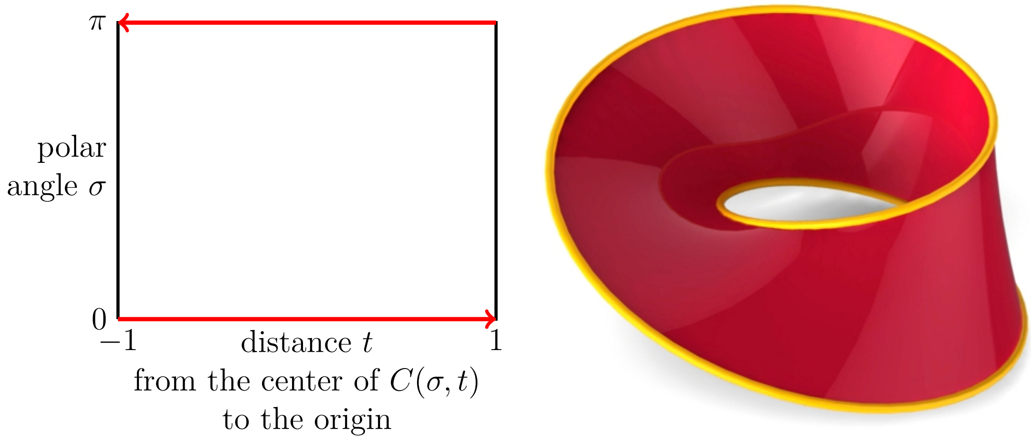

Figure 3.2: Left: Visualization of the Möbius band. The red arrows are identified with each other.
Right: The Möbius band embedded in $\mathbb{R}^3$. (image: CC-BY-SA 3.0 JoshDif)

At first, let $\sigma = 0$, then the circle $C(0, t)$ is parametrized by

$$C(0,t) = \left\{ \left(t, \sqrt{1-t^2}\cos\varphi, \sqrt{1-t^2}\sin\varphi \right) : \varphi \in \mathbb{T} \right\}.$$

The circle $C(\sigma, t)$ is just a rotation of $C(0, t)$ with the angle σ around the x_3 axis. This rotation can be described with the rotation matrix

$$Q_\sigma = \begin{pmatrix} \cos\sigma & -\sin\sigma & 0 \\ \sin\sigma & \cos\sigma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So we have $C(\sigma, t) = Q_\sigma C(0, t)$ which shows that

$$C(\sigma,t) = \left\{ \begin{pmatrix} t\cos\sigma - \sqrt{1-t^2}\cos\varphi\sin\sigma \\ t\sin\sigma + \sqrt{1-t^2}\cos\varphi\cos\sigma \\ \sqrt{1-t^2}\sin\varphi \end{pmatrix} : \varphi \in \mathbb{T} \right\}.$$

$\square$

The following basic properties of the circular average transform are easy to see.

1. As a special case of spherical means, the circular average transform preserves constant functions in the sense of

$$f \equiv c \in \mathbb{C} \Rightarrow \mathcal{T} f \equiv c.$$

Furthermore, it satisfies the symmetry

$$\mathcal{T} f(\sigma, t) = \mathcal{T} f(\sigma + \pi, -t), \quad \sigma \in [0, \pi], \ t \in [-1, 1].$$

2. Let Ref denote the reflection in the x_1–x_2 plane, i.e. $\mathrm{Ref}(x_1, x_2, x_3) = (x_3, x_2, -x_3)$. Then $\mathcal{T}f = \mathcal{T}(\mathrm{Ref}f)$ for all functions $f : \mathbb{S}^2 \to \mathbb{C}$. So, if the function f is odd in the third component, then its circular average transform vanishes.
3. Equation (3.9) implies the formula

$$\mathcal{T}f\,(\sigma, \cos\vartheta) = \frac{1}{2\pi}\int_0^{2\pi} f\,(\gamma_{\sigma,t}(\varphi))\,\mathrm{d}\varphi \tag{3.10}$$

with the integration path

$$\gamma_{\sigma,t} : \mathbb{T} \to \mathbb{S}^2 : \varphi \mapsto t\begin{pmatrix}\cos\sigma\\ \sin\sigma\\ 0\end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix}-\sin\sigma\cos\varphi\\ \cos\sigma\cos\varphi\\ \sin\varphi\end{pmatrix}.$$

One should note that the integration variable φ does not represent the arc length on a circle $C(\sigma, t)$. Instead, the length of the tangent vector is $\|\dot{\gamma}_{\sigma,t}\| = \sqrt{1-t^2}$. When we rewrite the integral (3.10) in terms of the arc length, it must be divided by $\sqrt{1-t^2}$.

Remark 3.4. The family $\mathscr{C}$ consists of all circles whose poles are located on the equator $\boldsymbol{e}_3^\perp$ of the sphere. One could also define the circular average transform for circles with poles on any great circle $\boldsymbol{x}^\perp$, $\boldsymbol{x} \in \mathbb{S}^2$. But since we are free to choose the coordinate system, we restrict our considerations to the equator.

The following theorem gives a necessary condition for a function to be in the range of the circular average transform. This range condition is very much like the range condition for the planar Radon transform that was proven in [24, Lemma 2.3].

Theorem 3.5. *For every function $f \in C\,(\mathbb{S}^2)$, the circular average transform $\mathcal{T}f$ satisfies the following condition: For every $n \in \mathbb{N}_0$, the function*

$$\sigma \mapsto \int_1^1 \mathcal{T}f\,(\sigma, t)\, t^n\,\mathrm{d}t, \quad \sigma \in \mathbb{T},$$

is a trigonometric polynomial in σ of degree n.

Proof. Using (3.2) and (3.7), we observe that that for $f \in C\,(\mathbb{S}^2)$ and $g \in C\,([-1, 1])$

$$\begin{aligned}\int_{-1}^1 \mathcal{T}f\,(\sigma, t)\, g(t)\,\mathrm{d}t &= \int_{-1}^1 \mathcal{M}f\,(\boldsymbol{e}_\sigma, t)\, g(t)\,\mathrm{d}t\\ &= \int_{-1}^1 \int_{\langle \boldsymbol{x}, \boldsymbol{e}_\sigma\rangle = t} f\,(\boldsymbol{x})\, g(t)\,\mathrm{d}s\,(\boldsymbol{x})\,\mathrm{d}t\\ &= \int_{\mathbb{S}^2} f\,(\boldsymbol{x})\, g\,(\langle \boldsymbol{x}, \boldsymbol{e}_\sigma\rangle)\,\mathrm{d}\boldsymbol{x}\\ &= g \star f\,(\boldsymbol{e}_\sigma)\,.\end{aligned}$$

Considering the Funk-Hecke formula (2.41), we see that if g is a polynomial of degree n, then $g \star f$ is a homogeneous polynomial of degree n on the sphere. So $g \star f$ can be written in terms of spherical harmonics,

$$g \star f = \sum_{k=-n}^{n} \alpha_k Y_{n,k}$$

for some coefficient vector $\boldsymbol{\alpha} \in \mathbb{C}^{2n+1}$. Inserting the representation (2.39) of the spherical harmonics shows that $f \star g\,(\boldsymbol{e}_\sigma)$ is a trigonometric polynomial of degree n in σ. □

3.3 Relation to planar Radon transforms

In this section, we present two well-known transforms that are defined on the two dimensional plane, namely the Radon transform and the circular Radon transform. We utilize two mappings from the sphere onto the plane to establish a connection from the circular average transform to these other transforms.

3.3.1 Relation to the circular Radon transform

Definition 3.6. Let $f : \mathbb{R}^2 \to \mathbb{C}$ be a function on the two-dimensional plane. Then its *circular Radon transform* is defined as (see [46])

$$\mathcal{R}_{\mathrm{c}} f(u,r) = \frac{1}{2\pi} \int_0^{2\pi} f\left(u + r\cos\varphi, r\sin\varphi\right) \mathrm{d}\varphi, \quad (u,r) \in \mathbb{R} \times \mathbb{R}_+. \tag{3.11}$$

The circular Radon transform of a function f computes the mean value of a function along circles having the center $(u,0)$, which lies on the x_1 axis of the two-dimensional plane, and the radius r. The circumference of such a circle is $2\pi r$. The circular Radon transform of a constant function $f_0 \equiv c \in \mathbb{C}$ is again constant,

$$\mathcal{R}_{\mathrm{c}} f_0(u,r) = \frac{1}{2\pi} \int_0^{2\pi} c \, \mathrm{d}\varphi = c, \quad (u,r) \in \mathbb{R} \times \mathbb{R}_+.$$

The following theorem shows the relation between the circular Radon transform and the circular average transform. A similar version of the theorem was proven in [59]. This method was also used in [58] as one step of inverting the circular Radon transform with the help of the Funk transform. Before stating the theorem, we define the stereographic projection of the two-sphere onto the x_2–x_3 plane by

$$\pi_{(-1,0,0)} : \mathbb{S}^2 \setminus \{(-1,0,0)\} \to \mathbb{R}^2,\ \pi_{(-1,0,0)}(\boldsymbol{x}) = \left(\frac{x_2}{1+x_1}, \frac{x_3}{1+x_1} \right).$$

The pushforward metric in $\mathbb{R}^2$ with respect to the stereographic projection is called the spherical metric. It is given by

$$\frac{\mathrm{d}x_1\mathrm{d}x_2}{1+x_1^2+x_2^2}, \tag{3.12}$$

see [54, Chapter 1].

Theorem 3.7. *Let $\pi_{(-1,0,0)}$ denote the stereographic projection as above and $(\sigma,t) \in \mathbb{T} \times [-1,1]$ satisfy $t + \cos\sigma \neq 0$. Then the following relation between the circular average transform on the sphere and the circular Radon transform in the plane holds:*

$$\begin{aligned} &\sqrt{1-t^2}\,[\mathcal{T}f]\,(\sigma,t) \\ =\;& [\mathcal{R}_{\mathrm{c}}\,(F)]\left(\frac{t\sin\sigma+\sqrt{1-t^2}\cos\sigma}{1+t\cos\sigma-\sqrt{1-t^2}\sin\sigma}-\frac{\sqrt{1-t^2}}{t+\cos\sigma},\frac{\sqrt{1-t^2}}{t+\cos\sigma}\right) \end{aligned} \tag{3.13}$$

where

$$F\,(\boldsymbol{\xi}) = \frac{\left[f\circ\pi_{(-1,0,0)}\right](\boldsymbol{\xi})}{2\left(1+\|\boldsymbol{\xi}\|^2\right)}, \quad \boldsymbol{\xi}\in\mathbb{R}^2.$$

Proof. Let $(\sigma,t) \in \mathbb{T} \times [-1,1]$. It is well-known that the stereographic projection maps circles onto circles. At first, we calculate the radius of the circle $\pi_{(-1,0,0)}C_{\sigma,t}$. Inserting $\varphi = 0$ and $\varphi = \pi$ into (3.9) gives us the points

$$\boldsymbol{x}_{1,2} = \left(t\cos\sigma \mp \sqrt{1-t^2}\sin\sigma, t\sin\sigma \pm \sqrt{1-t^2}\cos\sigma, 0\right)^\top$$

which are antipodal on the circle $C_{\sigma,t}$. Then the stereographic projection of these points is given by

$$\pi_{(-1,0,0)}\boldsymbol{x}_{1,2} = \left(\frac{t\sin\sigma \pm \sqrt{1-t^2}\cos\sigma}{1+t\cos\sigma\mp\sqrt{1-t^2}\sin\sigma}, 0\right).$$

Since the circle $C_{\sigma,t}$ is symmetric in the x_3 coordinate, the circle $\pi_{(-1,0,0)}C_{\sigma,t}$ is symmetric in its second coordinate. So the points $\pi_{(-1,0,0)}\boldsymbol{x}_1$ and $\pi_{(-1,0,0)}\boldsymbol{x}_2$ are also antipodal on the circle $\pi_{(-1,0,0)}C_{\sigma,t}$. Hence the radius $r_{\sigma,t}$ of the circle $\pi_{(-1,0,0)}C_{\sigma,t}$ is given by

$$\begin{aligned} r_{\sigma,t} &= \frac{1}{2}\left(\frac{t\sin\sigma+\sqrt{1-t^2}\cos\sigma}{1+t\cos\sigma-\sqrt{1-t^2}\sin\sigma} - \frac{t\sin\sigma-\sqrt{1-t^2}\cos\sigma}{1+t\cos\sigma+\sqrt{1-t^2}\sin\sigma}\right) \\ &= \frac{\sqrt{1-t^2}}{t+\cos\sigma}. \end{aligned}$$

Now we determine the center $(u_{\sigma,t}, 0)$ of the circle $\pi_{(-1,0,0)}C_{\sigma,t}$, which is just the midpoint of the two antipodal points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, so

$$u_{\sigma,t} = \frac{t\sin\sigma+\sqrt{1-t^2}\cos\sigma}{1+t\cos\sigma-\sqrt{1-t^2}\sin\sigma} - \frac{\sqrt{1-t^2}}{t+\cos\sigma}.$$

Equation (3.13) follows by inserting the formula (3.12) of the spherical metric in the plane. □

The circular Radon transform has been examined in many publications like [46] or [2]. The latter featured an inversion method utilizing Fourier transforms. The connection shown in Theorem 3.7 can be used to compute the inversion of the circular average transform. This approach has some drawbacks, like, it is necessary to deal with functions on the whole plane instead of the compact sphere.

3.3.2 Relation to the Radon transform

The circular average transform on the sphere is closely related to the Radon transform in the plane which was first described in [44] (for an English version, see [45]). The *Radon transform* $\mathcal{R}$ of a function $f : \mathbb{R}^2 \to \mathbb{C}$ is defined by

$$\mathcal{R}f\,(\sigma, t) = \int\limits_{\langle \boldsymbol{e}_\sigma, \boldsymbol{x}\rangle = t} f(\boldsymbol{x})\,\mathrm{d}s(\boldsymbol{x}), \quad (\sigma, t) \in \mathbb{T} \times \mathbb{R},$$

where the integral is carried out with respect to the arc length s. The domain of integration is the line $\{\boldsymbol{x} \in \mathbb{R}^2 : \langle \boldsymbol{e}_\sigma, \boldsymbol{x}\rangle = t\}$ which has the angle σ to the x_2 axis and the distance t to the origin. The Radon transform can be written equivalently in terms of a line integral as

$$\mathcal{R}f(\sigma, t) = \int\limits_{-\infty}^{\infty} f\left(t\begin{pmatrix}\cos\sigma \\ \sin\sigma\end{pmatrix} + \begin{pmatrix}-\sin\sigma \\ \cos\sigma\end{pmatrix} u\right) \mathrm{d}u, \quad (\sigma, t) \in \mathbb{T} \times \mathbb{R}.$$

Let $f : \mathbb{S}^2 \to \mathbb{C}$ be a function on the sphere. We define its corresponding planar function $\tilde{f} : \mathbb{R}^2 \to \mathbb{C}$ by

$$\tilde{f}\,(x_1, x_2) = \frac{1}{2\pi\sqrt{x_1^2 + x_2^2}} \begin{cases} f\begin{pmatrix} x_1 \\ x_2 \\ \sqrt{1 - x_1^2 - x_2^2}\end{pmatrix} + f\begin{pmatrix} x_1 \\ x_2 \\ -\sqrt{1 - x_1^2 - x_2^2}\end{pmatrix}, & x_1^2 + x_2^2 < 1 \\ 0, & \text{otherwise.}\end{cases}$$

The corresponding planar function $\tilde{f}$ is like an orthogonal projection of the function f onto the x_1–x_2 plane. For every point $(x_1, x_2) \in B_1(0) \subset \mathbb{R}^2$ in the open unit circle $B_1(0) = \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| < 1\}$, there exist exactly two points on the sphere whose first two coordinates are equal to those of (x_1, x_2), namely $\boldsymbol{x}^\pm = \left(x_1, x_2, \pm\sqrt{1 - x_1^2 - x_2^2}\right)^\top$. The corresponding planar function function evaluated at (x_1, x_2) is just the average of $f\,(\boldsymbol{x}^+)$ and $f\,(\boldsymbol{x}^-)$. If the function f is even with respect to the x_3 axis, then we have

$$\frac{f\,(\boldsymbol{x})}{\pi\sqrt{x_1^2 + x_2^2}} = \tilde{f}\,(x_1, x_2)$$

for $\boldsymbol{x} = (x_1, x_2, x_3)^\top \in \mathbb{S}^2$.

Theorem 3.8. *Let f be a function on the sphere which is even in the third component and $\widetilde{f}$ denote its corresponding planar function as defined above. Then the following relation between the circular average transform on the sphere and the Radon transform on the plane holds:*

$$\mathcal{T}f(\sigma,t) = \left[\mathcal{R}\tilde{f}\right](\sigma,t)\,, \quad (\sigma,t) \in \mathbb{T} \times (-1,1)\,.$$

Proof. According to (3.10), we have

$$\mathcal{T}f\,(\sigma,t) = \frac{1}{\pi}\int_0^\pi f\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\\ 0\end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix}-\cos\varphi\sin\sigma\\ \cos\varphi\cos\sigma\\ \sin\varphi\end{pmatrix}\right)\mathrm{d}\varphi.$$

because f is even in the third component. $\mathcal{T}f$ can be written in terms of the corresponding planar function $\tilde{f}$ as

$$\frac{\mathcal{T}f\,(\sigma,t)}{\sqrt{1-t^2-(1-t^2)\,(\cos\varphi)^2}} = \int_0^\pi \tilde{f}\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + \sqrt{1-t^2}\cos\varphi\begin{pmatrix}-\sin\sigma\\ \cos\sigma\end{pmatrix}\right)\mathrm{d}\varphi$$

since

$$\left\|t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix}-\cos\varphi\sin\sigma\\ \cos\varphi\cos\sigma\end{pmatrix}\right\|^2 = t^2 + \left(1-t^2\right)(\cos\varphi)^2\,.$$

Performing the substitution

$$u \quad = \quad \sqrt{1-t^2}\cos\varphi,$$

with

$$\begin{aligned}
\mathrm{d}u &= \sqrt{1-t^2}\sin\varphi\,\mathrm{d}\varphi\\
&= \sqrt{1-t^2}\sqrt{1-\frac{u^2}{1-t^2}}\,\mathrm{d}\varphi\\
&= \sqrt{1-t^2-u^2}\,\mathrm{d}\varphi
\end{aligned}$$

leads to

$$\frac{\mathcal{T}f\,(\sigma,\cos\vartheta)}{\sqrt{1-t^2-u^2}} = \frac{1}{2\pi}\int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}} \frac{\tilde{f}\left(t\begin{pmatrix}\cos\sigma\\ \sin\sigma\end{pmatrix} + u\begin{pmatrix}-\sin\sigma\\ \cos\sigma\end{pmatrix}\right)}{\sqrt{1-t^2-u^2}}\,\mathrm{d}u. \tag{3.14}$$

Since $\tilde{f}$ vanishes outside the unit circle $B_1(0)$ by definition, equation (3.14) becomes

$$\mathcal{T}f\,(\sigma,\cos\vartheta) = \int_{\left\langle \boldsymbol{x},\binom{\cos\sigma}{\sin\sigma}\right\rangle = t} \tilde{f}\,(\boldsymbol{x})\;\mathrm{d}s(\boldsymbol{x}).$$

□

The Radon transform and its inversion have been subject to many publications, since it is used to describe the mathematics of computerized tomography, cf. [24, 49]. Using the equivalence shown in Theorem 3.8 to compute the circular average transform is somehow problematic because the corresponding function usually has a singularity along the unit circle $\|\boldsymbol{x}\| = 1$. If, for instance, the function $f \equiv 1$ is constant on the two-sphere, its corresponding function

$$\tilde{f}(\boldsymbol{x}) = \frac{1}{\pi\sqrt{1-\|\boldsymbol{x}\|^2}}, \quad \boldsymbol{x} \in B_1(0) \subset \mathbb{R}^2$$

has a singularity along the unit circle because

$$\lim_{\|\boldsymbol{x}\|\uparrow 1} \tilde{f}(\boldsymbol{x}) = \infty \text{ whereas } \lim_{\|\boldsymbol{x}\|\downarrow 1} \tilde{f}(\boldsymbol{x}) = 0.$$

3.4 Singular value decomposition

The following theorem shows the singular value decomposition (SVD) of the circular average transform $\mathcal{T}$ in terms of the spherical harmonics $Y_{n,k}$ and the functions $B_{n,k}$ from (2.29). This will be the basis of our inversion algorithm.

Theorem 3.9. *The circular average transform, as defined in* (3.7)*, can be extended continuously to a compact linear operator* $\mathcal{T} : L^2(\mathbb{S}^2) \to L^2(\mathbb{T} \times [-1,1])$*. The singular system*

$$\{(Y_{n,k}, B_{n,k}, \lambda(n,k)) : \; n \in \mathbb{N}_0, \; k = -n, \ldots, n\}$$

of the circular average transform consists of the spherical harmonics $Y_{n,k}$*, cf.* (2.39)*, the basis functions*

$$B_{n,k} : \mathbb{T} \times [-1,1] \to \mathbb{C}, \; (\sigma, t) \mapsto \sqrt{\frac{2n+1}{4\pi}} \mathrm{e}^{\mathrm{i}k\sigma} P_n(t),$$

cf. (2.29)*, and the singular values*

$$\lambda(n,k) = \begin{cases} (-1)^{(n+k)/2} \, \frac{(n+k-1)!!}{(n-k)!!} \sqrt{\frac{(n-k)!}{(n+k)!}}, & n+k \text{ even} \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

The following singular value decomposition holds for all $f \in L^2(\mathbb{S}^2)$*,*

$$\langle \mathcal{T}f, B_{n,k} \rangle = \lambda(n,k) \langle f, Y_{n,k} \rangle, \quad n \in \mathbb{N}_0, \; k = -n, \ldots, n. \tag{3.16}$$

Notation. We are going to use the symbol $\mathcal{T}$ for the operator defined for continuous functions in (3.7) as well as for its extension on $L^2(\mathbb{S}^2)$.

Proof. Let $n \in \mathbb{N}_0$ and $k \in \{-n, \dots, n\}$. Initially, we compute the circular average transform applied to a spherical harmonic $Y_{n,k}$. By the Definition 3.2 of the circular average transform, the singular value decomposition of the mean operator in Theorem 3.1 and the definition of the spherical harmonics in (2.39),

$$\begin{aligned} \mathcal{T}Y_{n,k}(\sigma, t) &= \mathcal{M}Y_{n,k}(\boldsymbol{e}_\sigma, t) \\ &= P_n(t)\, Y_{n,k}(\boldsymbol{e}_\sigma) \\ &= P_n(t) \sqrt{\frac{2n+1}{4\pi}} \sqrt{\frac{(n-k)!}{(n+k)!}} \mathrm{e}^{ik\sigma} P_{n,k}(0). \end{aligned}$$

To derive a closed formula for $P_{n,k}(0)$, we start with the case $P_{n,0}(0) = P_n(0)$. According to the recurrence relation (2.38) for the Legendre polynomials, we have for $n \geq 1$

$$P_n(0) = \frac{-(n-1)P_{n-2}(0)}{n} = \begin{cases} (-1)^{n/2} \frac{(n-1)!!}{n!!} & : n \text{ even} \\ 0 & : n \text{ odd} \end{cases}$$

and $P_0(0) = 1$. With the recurrence relation (2.38) for the associated Legendre polynomials, we obtain for $n, k \geq 1$,

$$\begin{aligned} P_{n,k}(0) &= -(n+k-1)\, P_{n-1,k-1}(0) \\ &= (-1)^k \frac{(n+k-1)!!}{(n-k-1)!!} P_{n-k,0}(0) \\ &= (-1)^k \frac{(n+k-1)!!}{(n-k-1)!!} \frac{\left(1+(-1)^{n-k}\right)}{2} (-1)^{(n-k)/2} \frac{(n-k-1)!!}{(n-k)!!} \\ &= (-1)^{(n+k)/2} \frac{(n+k-1)!!}{(n-k)!!} \frac{\left(1+(-1)^{n-k}\right)}{2} \end{aligned}$$

and, by (2.37),

$$\begin{aligned} P_{n,-k}(0) &= (-1)^k \frac{(n-k)!}{(n+k)!} P_{n,k}(0) \\ &= \frac{(n-k-1)!!}{(n+k)!!} (-1)^{(n+k)/2+k} \frac{\left(1+(-1)^{n-k}\right)}{2} \end{aligned}$$

where we define $(-1)!! = 1$. Hence,

$$\begin{aligned} \mathcal{T}Y_{n,k}(\sigma, t) &= (-1)^{(n+k)/2} \frac{(n+k-1)!!}{(n-k)!!} \frac{\left(1+(-1)^{n-k}\right)}{2} \sqrt{\frac{2n+1}{4\pi}} \sqrt{\frac{(n-k)!}{(n+k)!}} \\ &\quad \cdot P_n(t)\, \mathrm{e}^{ik\sigma} \\ &= (-1)^{(n+k)/2} \frac{(n+k-1)!!}{(n-k)!!} \frac{\left(1+(-1)^{n-k}\right)}{2} \sqrt{\frac{(n-k)!}{(n+k)!}} B_{n,k}(\sigma, t). \end{aligned}$$

With $\lambda(n,k)$ from (3.15) and the orthonormality of $Y_{n,k}$ and $B_{n,k}$, this proves (3.16).
Let $f \in \operatorname{span}\{Y_{n,k} : n \in \mathbb{N}_0,\ k = -n, \ldots, n\}$, i.e. f is a polynomial. Then

$$\mathcal{T}f = \sum_{n=0}^{\infty} \sum_{k=-n}^{n} \hat{f}(n,k)\, \lambda(n,k)\, B_{n,k}(\sigma, t)$$

where the first summation is only over finitely many indexes n.
Since the singular values $\lambda(n,k)$ are bounded independently of n and k, the operator $\mathcal{T}$ is bounded on $\operatorname{span}\{Y_{n,k} : n \in \mathbb{N}_0,\ k = -n, \ldots, n\}$ with respect to the corresponding L^2 norms. So $\mathcal{T}$ is a bounded linear operator on a dense subspace of $L^2(\mathbb{S}^2)$. Hence, there exists a unique bounded linear operator $L^2(\mathbb{S}^2) \to L^2(\mathbb{T} \times [-1,1])$ extending $\mathcal{T}$. For simplicity, we denote the extension just by $\mathcal{T}$. Because the sequence $\lambda(n,k)$ converges to zero for $n \to \infty$, the operator $\mathcal{T}$ is also compact. □

Remark 3.10. Double factorials can be expressed in terms of factorials for n even by

$$(n-1)!! = \frac{n!}{2^{n/2}\left(\frac{n}{2}\right)!} \tag{3.17}$$

and

$$n!! = 2^{n/2}\left(\frac{n}{2}\right)!. \tag{3.18}$$

With these formulas we can derive a representation of the singular values (3.15) without double factorials. For $n+k$ even,

$$\begin{aligned}
\lambda(n,k) &= (-1)^{(n+k)/2} \frac{(n+k-1)!!}{(n-k)!!} \sqrt{\frac{(n-k)!}{(n+k)!}} \\
&= (-1)^{(n+k)/2} \frac{(n+k)!}{2^{(n+k)/2}\left(\frac{n+k}{2}\right)!} \frac{1}{2^{(n-k)/2}\left(\frac{n-k}{2}\right)!} \sqrt{\frac{(n-k)!}{(n+k)!}} \\
&= (-1)^{(n+k)/2} \frac{\sqrt{(n+k)!}\sqrt{(n-k)!}}{2^{n}\left(\frac{n+k}{2}\right)!\left(\frac{n-k}{2}\right)!}.
\end{aligned} \tag{3.19}$$

Asymptotics of the singular values. To describe the asymptotic behavior of the singular values $\lambda(n,k)$, we use the following version of Stirling's formula. It was shown in [47] that for $n = 1, 2, \ldots$

$$n! = \sqrt{2\pi}\, n^{n+1/2} \mathrm{e}^{-n} \mathrm{e}^{r(n)} \tag{3.20}$$

where r_n satisfies

$$\frac{1}{12n+1} < r_n < \frac{1}{12n}. \tag{3.21}$$

Theorem 3.11. *Let $n \in \mathbb{N}_0$, $k \in \{-n, \ldots, n\}$, and $n+k$ be even, then the singular values $\lambda(n,k)$ of the circular average transform, cf. Theorem 3.9, can be written as*

$$|\lambda(n,k)| = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt[4]{n^2 - k^2}} \mathrm{e}^{R(n,k)} \tag{3.22}$$

where

$$|R(n,k)| < \frac{1}{3 \min(n+k, n-k)}. \tag{3.23}$$

Furthermore, for n even, we have

$$|\lambda(n,n)| \simeq \frac{1}{\sqrt[4]{\pi n}}, \quad n \to \infty. \tag{3.24}$$

Proof. By Stirling's formula (3.20) and the representation (3.19) of the singular values,

$$\begin{aligned}
|\lambda(n,k)| &= \frac{\sqrt{(n+k)!}\sqrt{(n-k)!}}{2^n \left(\frac{n+k}{2}\right)! \left(\frac{n-k}{2}\right)!} \\
&= \frac{\sqrt{2\pi}\sqrt{(n+k)^{n+k+1/2}\,\mathrm{e}^{-(n+k)}\mathrm{e}^{r(n+k)}}\sqrt{(n-k)^{n-k+1/2}\,\mathrm{e}^{-(n-k)}\mathrm{e}^{r(n-k)}}}{2^n 2\pi \left(\left(\frac{n+k}{2}\right)^{\frac{n+k+1}{2}} \mathrm{e}^{-\frac{n+k}{2}} \mathrm{e}^{r\left(\frac{n+k}{2}\right)}\right)\left(\left(\frac{n-k}{2}\right)^{\frac{n-k+1}{2}} \mathrm{e}^{-\frac{n-k}{2}} \mathrm{e}^{r\left(\frac{n-k}{2}\right)}\right)} \\
&= \frac{1}{\sqrt{2\pi}} 2^{-n+\frac{n+k+1}{2}+\frac{n-k+1}{2}} (n+k)^{\frac{n-k+1/2}{2}-\frac{n-k+1}{2}} (n-k)^{\frac{n-k+1/2}{2}-\frac{n-k+1}{2}} \cdot \\
&\quad \mathrm{e}^{-\frac{n+k}{2}-\frac{n-k}{2}+\frac{n+k}{2}+\frac{n-k}{2}} \mathrm{e}^{\frac{r(n+k)}{2}+\frac{r(n-k)}{2}-r\left(\frac{n+k}{2}\right)-r\left(\frac{n-k}{2}\right)} \\
&= \frac{\sqrt{2}}{\sqrt{\pi}} (n+k)^{-1/4} (n-k)^{-1/4} \mathrm{e}^{R(n,k)}.
\end{aligned}$$

where

$$R(n,k) = \frac{r(n+k)}{2} + \frac{r(n-k)}{2} - r\left(\frac{n+k}{2}\right) - r\left(\frac{n-k}{2}\right).$$

This shows (3.22).
Next, we derive a bound for the error term $R(n,k)$. We set $\nu = \min(n+k, n-k)$, then

$$n+k, n-k \in \{\nu, \nu+1, \ldots, 2n-\nu\}.$$

With the inequality (3.21) for r, we derive the upper bound

$$\begin{aligned}
R(n,k) &= \frac{r(n+k)}{2} + \frac{r(n-k)}{2} - r\left(\frac{n+k}{2}\right) - r\left(\frac{n-k}{2}\right) \\
&< \frac{1}{24(n+k)} + \frac{1}{24(n-k)} - \frac{1}{6(n+k)+1} - \frac{1}{6(n+k)+1} \\
&\le \frac{1}{24\nu} + \frac{1}{24\nu} - \frac{1}{6(2n-\nu)+1} - \frac{1}{6(2n-\nu)+1} \\
&= \frac{1}{12\nu} - \frac{2}{6(2n-\nu)+1}
\end{aligned}$$

and, analogously, the lower bound

$$\begin{aligned} R(n,k) &= \frac{r(n+k)}{2} + \frac{r(n-k)}{2} - r\left(\frac{n+k}{2}\right) - r\left(\frac{n-k}{2}\right) \\ &> -\frac{1}{6(n+k)} - \frac{1}{6(n+k)} \\ &> -\frac{1}{6\nu} - \frac{1}{6\nu} \\ &= -\frac{1}{3\nu}. \end{aligned}$$

The last two calculations imply (3.23).
Now let $n \in \mathbb{N}$ be even. Using the same arguments as in the first part of the proof, we observe that

$$\begin{aligned} |\lambda(n,n)| &= \frac{\sqrt{(2n)!}}{2^n \left(\frac{2n}{2}\right)!} \\ &= \frac{\sqrt{\sqrt{2\pi}\,(2n)^{2n+1/2}\,\mathrm{e}^{-2n}\mathrm{e}^{r(2n)}}}{2^n\sqrt{2\pi}n^{n+1/2}\mathrm{e}^{-n}\mathrm{e}^{r(n)}} \\ &= \frac{1}{\sqrt[4]{\pi n}}\mathrm{e}^{r(2n)/2-r(n)}. \end{aligned}$$

Equation (3.24) follows using the truncated Taylor series of the exponential function

$$\mathrm{e}^{1/x} = 1 + \mathcal{O}\left(\frac{1}{x}\right), \quad x \to \infty.$$

□

Remark 3.12. To get a better understanding of the error bound (3.23) for the asymptotic expression (3.22) of the singular values of the circular average transform, let us fix some $\nu \in \mathbb{N}$. Then

$$|R(n,k)| < \frac{1}{3\nu}, \quad n \geq \nu,\ |k| \leq n - \nu.$$

Adjoint operator. Let $A \in \mathcal{L}(X \to Y)$ be a bounded linear operator between the two Hilbert spaces X and Y. Then the operator $A^\star \in \mathcal{L}(Y \to X)$ satisfying $\langle Af, g\rangle = \langle f, A^\star g\rangle$ for all $f \in X$ and $g \in Y$ is called the adjoint operator of A.

Theorem 3.13. *The adjoint operator of the circular average transform* $\mathcal{T} : L^2(\mathbb{S}^2) \to L^2(\mathbb{T} \times [-1,1])$ *is given by*

$$\mathcal{T}^\star g(\boldsymbol{x}) = \frac{1}{2\pi}\int_0^{2\pi} g(\sigma, \langle \boldsymbol{x}, \boldsymbol{e}_\sigma\rangle)\,\mathrm{d}\sigma. \tag{3.25}$$

Proof. Let $f \in C(\mathbb{S}^2)$ and $g \in C(\mathbb{T} \times [-1,1])$. Then, by (2.31),

$$\begin{aligned}
\langle \mathcal{T} f, g \rangle &= \int_0^{2\pi} \int_{-1}^{1} \mathcal{T} f(\sigma, t)\, g(\sigma, t)\, \mathrm{d}t\, \mathrm{d}\sigma \\
&= \int_0^{2\pi} \left(\int_{-1}^{1} \mathcal{M} f(\boldsymbol{e}_\sigma, t)\, g(\sigma, t)\, \mathrm{d}t \right) \mathrm{d}\sigma \\
&= \frac{1}{2\pi} \int_0^{2\pi} \left(\int_{\mathbb{S}^2} f(\boldsymbol{x})\, g(\sigma, \langle x, \boldsymbol{e}_\sigma \rangle)\, \mathrm{d}\boldsymbol{x} \right) \mathrm{d}\sigma = \langle f, \mathcal{T}^\star g \rangle .
\end{aligned}$$

Since the sets of continuous functions are dense in both the Hilbert spaces $L^2(\mathbb{S}^2)$ and $L^2(\mathbb{T} \times [-1,1])$, we can conclude (3.25). □

3.5 Computing the circular average transform

In this section, we describe an algorithm for the computation of the circular average transform based on its singular value decomposition from Theorem 3.9. Suppose that we are given a continuous function $f \in C(\mathbb{S}^2)$. For the derivation of an algorithm, we assume that f is a polynomial of degree up to $N \in \mathbb{N}$. In the first step, we compute the Fourier coefficients $\hat{f}(n,k)$ via the quadrature

$$\hat{f}(n,k) = \sum_{j=1}^{J} \omega_j f(\boldsymbol{x}_j)\, Y_{n,k}(\boldsymbol{x}_j), \quad n = 0, \ldots, N,\ k = -n, \ldots, n \tag{3.26}$$

with the nodes $\boldsymbol{x}_j$ and the corresponding weights ω_j for $j = 1, \ldots, J$. The computation of (3.26) can be done efficiently with the adjoint Nonequispaced Fast Spherical Fourier Transform (NFSFT) which was developed in [31]. The adjoint NFSFT is an approximative algorithm for computing the Fourier coefficients $\hat{f}(n,k)$ up to degree N given the function values $f(\boldsymbol{x}_j)$. When the number of nodes J is in $\mathcal{O}(N^2)$, the NFSFT needs $\mathcal{O}\left(N^2 \log^2 N\right)$ operations. The NFSFT is available as part of the open source NFFT software package [30].

We want to compute the circular average transform $\mathcal{T} f$ at some nodes

$$(\sigma_l, t_m) \in \mathbb{T} \times [-1,1], \quad l = 1, \ldots, L,\ m = 1, \ldots, M.$$

From the singular value decomposition in Theorem 3.9, we obtain that

$$\mathcal{T} f(\sigma_l, t_m) = \sqrt{\frac{2n+1}{4\pi}} \sum_{n=0}^{N} \sum_{k=-n}^{n} \hat{f}(n,k)\, \lambda(n,k)\, \mathrm{e}^{ik\sigma_l} P_n(t_m) . \tag{3.27}$$

First we examine the inner sum from (3.27). Therefore, we define the matrix $S \in \mathbb{C}^{(N+1)\times L}$ by

$$S(n,l) = \sum_{k=-n}^{n} \hat{f}\,(n,k)\,\lambda\,(n,k)\,\mathrm{e}^{ik\sigma_l}, \quad n = 0,\dots,N,\ l = 1,\dots,L.$$

We choose the equidistant nodes

$$\sigma_l = \frac{2\pi l}{L}, \quad l = 1,\dots,L$$

which is a pretty natural choice on $\mathbb{T}$. When we define $\hat{f}\,(n,k) = 0$ for $|k| > n$, then, for a fixed n, we can rewrite S with the discrete Fourier transform defined in (2.27):

$$\begin{aligned} S(n,l) &= \sum_{k=-n}^{n} \hat{f}\,(n,k)\,\lambda\,(n,k)\,\mathrm{e}^{2\pi ikl/L} \\ &= \sum_{k=0}^{2n+1} \hat{f}\,(n,N-k)\,\lambda\,(n,N-k)\,\mathrm{e}^{-2\pi ikl/L}\mathrm{e}^{2\pi iNl/L} \\ &= \left[\mathrm{DFT}\left(\hat{f}\,(n,N-\circ)\,\lambda\,(n,N-\circ)\right)\right](l)\mathrm{e}^{2\pi iNl/L}. \end{aligned} \tag{3.28}$$

So $[S(n,l)]_l$ is the discrete Fourier transform of the vector $\left[\hat{f}\,(n,k)\,\lambda\,(n,k)\right]_k$ multiplied with an exponential function. For analyzing the computational complexity, we assume that $L, M, N \in \mathcal{O}(N)$. In (3.28), we compute N fast Fourier transforms of size L which leads to a complexity of $\mathcal{O}\,(N^2 \log N)$. If the nodes σ_l are not equidistant, the FFT can be replaced by the NFFT (Nonequispaced Fast Fourier Transform) which is of the same complexity class, see [35].

Now, (3.27) can be written as

$$\mathcal{T}f\,(\sigma_l, t_m) = \sqrt{\frac{2n+1}{4\pi}} \sum_{n=0}^{N} S(n,l)P_n\,(t_m)\,. \tag{3.29}$$

Evaluating the sum in (3.29) means computing the product of the matrices $S^\top$ and $[P_n\,(t_m)]_{n=0,m=1}^{N,\ \ M}$. The computation of this matrix product is done in $\mathcal{O}\,(N^3)$ arithmetic operations. The matrix $[P_n\,(t_m)]_{n=0,m=1}^{N,\ \ M}$ can be evaluated using the three-term recurrence for the Legendre polynomials, which can be done using $\mathcal{O}\,(N^2)$ operations. The computation of (3.29) can be accelerated with the fast polynomial transform algorithms proposed in [43] which leads to a complexity of $\mathcal{O}\left(N^2\,(\log N)^2\right)$ steps. Hence, the whole computation of the circular average transform can be done with $\mathcal{O}\left(N^2\,(\log N)^2\right)$ operations. For the case that t_m are the Chebyshev nodes, we can alternatively use the algorithm [21] to compute (3.29) in $\mathcal{O}\left(N^2\,(\log N)^2\,/\log\log N\right)$ steps.

The following algorithm sums up the considerations in this section.

Algorithm 3.14. *Input: Function values $f\left(\boldsymbol{x}_j\right)$ of the function $f \in \Pi_N\left(\mathbb{S}^2\right)$, a quadrature rule*

$$\sum_{j=1}^{J} \omega_j f(\boldsymbol{x}_j)$$

using the weights ω_j and the nodes $\boldsymbol{x}_j \in \mathbb{S}^2$ that is exact for $\Pi_N\left(\mathbb{S}^2\right)$, and nodes

$$(\sigma_l, t_m) \in \mathbb{T} \times [-1,1], \quad l = 1, \ldots, L,\ m = 1, \ldots, M,$$

where $\mathcal{T} f$ should be calculated.

1. *Compute the Fourier coefficients*

$$\hat{f}(n,k) := \sum_{j=1}^{J} \omega_j f\left(\boldsymbol{x}_j\right) Y_{n,k}\left(\boldsymbol{x}_j\right)$$

   for $n = 0, \ldots, N,\ k = -n, \ldots, n$.

2. *Compute for each $n = 0, \ldots, N$ the fast Fourier transform*

$$S(n,l) := \left[\mathrm{DFT}\left(\hat{f}\left(n, N - \circ\right) \lambda\left(n, N - \circ\right)\right)\right](l) \mathrm{e}^{2\pi i N l / L}$$

   for $l = 1, \ldots, L$.

3. *Compute the discrete Legendre transform*

$$\mathcal{T} f\left(\frac{2\pi l}{L}, t_m\right) := \sqrt{\frac{2n+1}{4\pi}} \sum_{n=0}^{N} S(n,l) P_n\left(t_m\right)$$

   for $l = 1, \ldots, L,\ m = 1, \ldots, M$.

Output: Circular average transform

$$\mathcal{T} f\left(\frac{2\pi l}{L}, t_m\right), \quad l = 1, \ldots, L,\ m = 1, \ldots, M.$$

4

Inversion of the Circular Average Transform

We consider the inversion of the circular average transform $\mathcal{T}$ from Definition 3.2. Given a function $\mathcal{T}f$, we want to compute the original function $f \in L^2(\mathbb{S}^2)$. As shown in Theorem 3.9, the circular average transform $\mathcal{T} : L^2(\mathbb{S}^2) \to L^2(\mathbb{T} \times [-1,1])$ is a compact operator. Therefore, the inversion of the circular average transform is a so-called ill-posed problem. An ill-posed problem is, loosely speaking, a problem which is overly sensible to small deviations of the input data. For a more detailed view on ill-posed problems and their regularization, we refer to [29, 28].

4.1 Regularization

Pseudo-inverse. As a compact operator, the circular average transform is not invertible. That is why we introduce the concept of a pseudo-inverse. We denote by

$$\mathrm{R}(\mathcal{T}) = \left\{\mathcal{T}f : f \in L^2\left(\mathbb{S}^2\right)\right\}$$

the range of the operator $\mathcal{T}$ and by

$$\mathrm{N}(\mathcal{T}) = \left\{f \in L^2\left(\mathbb{S}^2\right) : \mathcal{T}f = 0\right\}$$

the nullspace of $\mathcal{T}$. The singular value decomposition from Theorem 3.9 shows that the nullspace of the circular average transform is non-trivial, in particular

$$\mathrm{N}(\mathcal{T}) = \overline{\operatorname{span}}\left\{Y_{n,k} : n \in \mathbb{N}_0,\, |k| \le n,\, n+k \text{ odd}\right\}, \tag{4.1}$$

which is equal to the set of all functions of $L^2(\mathbb{S}^2)$ that are odd in its third coordinate. The orthogonal complement of $\mathrm{N}(\mathcal{T})$ is denoted by

$$\mathrm{N}(\mathcal{T})^\perp = \left\{f \in L^2\left(\mathbb{S}^2\right) : \langle f, h\rangle = 0,\, h \in \mathrm{N}(\mathcal{T})\right\}.$$

We define the generalized inverse or Moore-Penrose pseudo-inverse

$$\mathcal{T}^\dagger : \mathrm{D}(\mathcal{T}^\dagger) = \mathrm{R}(\mathcal{T}) \oplus \mathrm{R}(\mathcal{T})^\perp \longrightarrow \mathrm{N}(\mathcal{T})^\perp$$

of the operator $\mathcal{T}$ as follows, cf. [29, Section 2.9]. Let $g \in \mathrm{D}(\mathcal{T}^\dagger)$, so $g = g_1 + g_2$ can be written as the sum of $g_1 \in \mathrm{R}(\mathcal{T})$ and $g_2 \in \mathrm{R}(\mathcal{T})^\perp$. The set $\{f \in L^2(\mathbb{S}^2) : \mathcal{T}f = g_1\}$ is a convex subset of the Hilbert space $L^2(\mathbb{S}^2)$. Hence there exists a unique function $f_0 \in L^2(\mathbb{S}^2)$ that minimizes the norm $\|\cdot\|_{L^2(\mathbb{S}^2)}$ on that set, i.e. $\|f_0\|_{L^2(\mathbb{S}^2)} \leq \|f\|_{L^2(\mathbb{S}^2)}$ for all $f \in L^2(\mathbb{S}^2)$ with $\mathcal{T}f = g_1$. We define

$$\mathcal{T}^\dagger g = f_0.$$

It is easy to see that $\mathcal{T}^\dagger$ is a linear operator. Another direct consequence from the definition is, that $\mathcal{T}^\dagger\mathcal{T}$ is equal to the orthogonal projection onto $\mathrm{N}(\mathcal{T})^\perp$. Since $\mathcal{T}$ is compact, its pseudo-inverse is unbounded.

Another characterization of the pseudo-inverse uses the singular value decomposition of the operator $\mathcal{T}$. For a function g that lies in the domain $\mathrm{D}\left(\mathcal{T}^\dagger\right)$ of the pseudo-inverse, we have

$$\widehat{\mathcal{T}^\dagger g}(n,k) = \begin{cases} \lambda(n,k)^{-1}\,\hat{g}(n,k), & \lambda(n,k) \neq 0 \\ 0, & \text{otherwise,} \end{cases}$$

and thus

$$\mathcal{T}^\dagger g = \sum_{n=0}^{\infty} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \frac{\hat{g}(n,k)}{\lambda(n,k)} Y_{n,k}, \quad g \in \mathrm{D}\left(\mathcal{T}^\dagger\right). \tag{4.2}$$

The mollifier method. We use the so-called mollifier method for invere problems which was introduced in [38]. The idea is to compute a smoothened version of f, namely

$$\psi \star f(\boldsymbol{x}) = \int_{\mathbb{S}^2} \psi(\langle \boldsymbol{x}, \boldsymbol{y} \rangle) f(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}, \quad \boldsymbol{x} \in \mathbb{S}^2 \tag{4.3}$$

for some mollifier $\psi : [-1,1] \to \mathbb{R}$.

Let $g \in C(\mathbb{T} \times [-1,1]) \cap \mathrm{D}(\mathcal{T}^\dagger)$, we want to compute $f = \mathcal{T}^\dagger g$. We assume that the mollifier ψ is a polynomial of degree $N \in \mathbb{N}$. As in the Funk-Hecke formula from Proposition 2.10, we denote the Fourier coefficients of the mollifier ψ with respect to the Legendre polynomials P_n by

$$\hat{\psi}(n) = 2\pi \int_{-1}^{1} \psi(t) P_n(t) \, \mathrm{d}t, \quad n = 0, \ldots, N.$$

So we can write the mollifier in terms of its Legendre decomposition

$$\psi(t) = \sum_{n=0}^{N} \frac{2n+1}{4\pi} \hat{\psi}(n) P_n(t), \quad t \in [-1,1]$$

(cf. Subsection 2.1.1). Thus, by the addition theorem (2.35),

$$\psi\left(\langle \boldsymbol{x},\boldsymbol{y}\rangle\right) = \sum_{n=0}^{N}\sum_{k=-n}^{n} \hat{\psi}\left(n\right) Y_{n,k}\left(\boldsymbol{x}\right) Y_{n,k}\left(\boldsymbol{y}\right), \quad \boldsymbol{x},\boldsymbol{y}\in\mathbb{S}^2.$$

With the Funk-Hecke formula (2.41), we see that

$$\begin{aligned} \psi\star f\left(\boldsymbol{x}\right) &= \int_{\mathbb{S}^2} \psi\left(\langle \boldsymbol{x},\boldsymbol{y}\rangle\right) f\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y} \\ &= \sum_{n=0}^{\infty}\sum_{k=-n}^{n} \hat{\psi}\left(n\right) \hat{f}\left(n,k\right) Y_{n,k}\left(\boldsymbol{x}\right), \quad \boldsymbol{x}\in\mathbb{S}^2. \end{aligned}$$

This formula also shows that the smoothened solution is a polynomial of degree N, i.e. $\psi\star f \in \mathit{\Pi}_N\left(\mathbb{S}^2\right)$. The mollified solution $\psi\star f$ can be expressed in terms of the given function g, by (4.2) we have

$$\psi\star\mathcal{T}^{\dagger}g = \sum_{n=0}^{\infty}\hat{\psi}\left(n\right)\sum_{\substack{k=-n\\2|(n+k)}}^{n}\frac{\hat{g}(n,k)}{\lambda(n,k)}Y_{n,k}. \tag{4.4}$$

Discrete data. Suppose that we are given the function values of $\mathcal{T}f$ on the discrete data points $(\sigma_l,t_m)\in\mathbb{T}\times[-1,1]$,

$$\mathcal{T}f\left(\sigma_l,t_m\right), \quad l=1,\dots,L,\ m=1,\dots,M$$

We are going to use a quadrature rule on $\mathbb{T}\times[-1,1]$, namely,

$$Qh = \sum_{l=1}^{L}\sum_{m=1}^{M} w_{l,m}h\left(\sigma_l,t_m\right), \quad h\in C\left(\mathbb{T}\times[-1,1]\right), \tag{4.5}$$

cf. Subsection 2.1.3.

For the mollified solution (4.4), we compute the Fourier coefficients

$$\hat{g}\left(n,k\right) = \langle g, B_{n,k}\rangle = \int_{-1}^{1}\int_{0}^{2\pi} g\left(\sigma,t\right) B_{n,k}(\sigma,t)\,\mathrm{d}\sigma\,\mathrm{d}t.$$

We replace the integral from the inner product by the quadrature rule Q and obtain

$$\langle g, B_{n,k}\rangle_Q = \sum_{l=1}^{L}\sum_{m=1}^{M} w_{l,m}g\left(\sigma_l,t_m\right) B_{n,k}(\sigma,t).$$

As in (2.6) for the general case, we denote by

$$\mathcal{L}_N g\left(\sigma,t\right) = \sum_{n=0}^{N}\sum_{k=-n}^{n}\langle g, B_{n,k}\rangle_Q\, B_{n,k}(\sigma,t) \tag{4.6}$$

the hyperinterpolation onto

$$\Pi_N\left(\mathbb{T}\times[-1,1]\right)=\operatorname{span}\left\{B_{n,k}(\sigma,t):\ n=0,\dots,N,\ k=-n,\dots,n\right\}.$$

Note that, in the definition of the hyperinterpolation, we have made the restriction $|k|\le n$ because the functions $B_{n,k}$ where $|k|>n$ do not lie in the range of the circular average transform, see Theorem 3.5.
We define the discretized regularized solution

$$f_\psi=\psi\star\mathcal{T}^\dagger\mathcal{L}_N g. \tag{4.7}$$

Noisy data. We assume that the given data g is disrupted by some white noise ε which is specified in the next definition.

Definition 4.1. Let $L,M\in\mathbb{N}_0$ and $(\sigma_l,t_m)\in\mathbb{T}\times[-1,1]$, $l=1,\dots,L,\ m=1,\dots,M$, be the nodes of a quadrature as in (4.5). Furthermore, let the random variable $\varepsilon:\mathbb{C}^{L\times M}\to\mathbb{C}$ be

1. unbiased, i.e. $\mathbb{E}\left(\varepsilon\left(l,m\right)\right)=0$ for all (l,m), and
2. uncorrelated, i.e. $\mathbb{E}\left(\varepsilon\left(l,m\right)\varepsilon\left(l',m'\right)\right)=0$ for $(l,m)\neq(l',m')$, and
3. has a standard deviation of $\delta>0$, i.e. $\mathbb{E}\left(\varepsilon\left(l,m\right)^2\right)=\delta^2$ for all (l,m).

Then we denote the noisy data by

$$g^\varepsilon\left(\sigma_l,t_m\right)=\mathcal{T}f\left(\sigma_l,t_m\right)+\varepsilon\left(l,m\right)$$

and we define the estimator

$$f_\psi^\varepsilon=\psi\star\left(\mathcal{T}^\dagger\mathcal{L}_N g^\varepsilon\right) \tag{4.8}$$

where we use the convention that

$$\mathcal{L}_N g^\varepsilon=\sum_{n=0}^{N}\sum_{k=-n}^{n}\sum_{l=1}^{L}\sum_{m=1}^{M}g^\varepsilon\left(\sigma_l,t_m\right)B_{n,k}\left(\sigma_l,t_m\right)B_{n,k}.$$

This is a standard statistical model for inverse problems, cf. [5]. Note that the estimator f_ψ^ε also depends on the quadrature Q and the degree N.

4.2 Decomposition of the error

As a measure for the accuracy of the estimator f_ψ^ε of the actual solution f, we use the mean integrated squared error (MISE). The MISE is defined as

$$\operatorname{MISE}\left(f_\psi^\varepsilon,f\right)=\mathbb{E}\left\|f_\psi^\varepsilon-f\right\|_{L^2(\mathbb{S}^2)}^2=\mathbb{E}\left(\int_{\mathbb{S}^2}\left|f_\psi^\varepsilon(\boldsymbol{x})-f(\boldsymbol{x})\right|^2\,\mathrm{d}\boldsymbol{x}\right).$$

We want to bound the MISE over some Sobolev balls

$$\mathscr{F}(s,S) = \left\{ f \in H^s\left(\mathbb{S}^2\right) \cap \mathrm{N}\left(\mathcal{T}\right)^{\perp} : \|f\|_{H^s(\mathbb{S}^2)} \le S \right\}$$

for some positive constants $s, S > 0$. To be more precise, we are interested in the maximum risk

$$\sup_{f \in \mathscr{F}(s,S)} \mathrm{MISE}\left(f_\psi^\varepsilon, f\right). \tag{4.9}$$

The condition that f lies in the orthogonal complement of the nullspace of the circular average transform is important since the nullspace is nontrivial, cf. (4.1). If this condition were dropped, we could replace f by $f + \alpha f_0$ where $f_0 \in \mathrm{N}(\mathcal{T}) \setminus \{0\}$ and $\alpha \in \mathbb{R}$. Then the data $\mathcal{T}(f + \alpha f_0)$ would be exactly the same as $\mathcal{T} f$ and, no matter what estimator we use, the MISE would grow infinitely for $\alpha \to \infty$. Thus, the maximum risk could not be bounded.

Now it is pretty reasonable to ask for mollifiers ψ that minimize the maximum risk. The maximum risk for the "best" choice of the mollifier is called the minimax risk

$$\inf_{\psi \in L^2[-1,1]} \sup_{f \in \mathscr{F}(s,S)} \mathrm{MISE}\left(f_\psi^\varepsilon, f\right), \tag{4.10}$$

cf. [53]. We are looking for mollifiers minimizing the maximum risk asymptotically as N goes to infinity.

In the rest of this section, we derive bounds for the MISE.

Lemma 4.2. *Let N be a positive integer and ε fulfill the conditions from Definition 4.1. Then the expected value of the estimator f_ψ^ε given in (4.8) is equal to the regularized solution f_ψ with exact data from (4.7), i.e.*

$$\mathbb{E} f_\psi^\varepsilon = f_\psi.$$

Proof. We have for $\boldsymbol{x} \in \mathbb{S}^2$,

$$\begin{aligned} \mathbb{E} f_\psi^\varepsilon(\boldsymbol{x}) &= \mathbb{E}\left(\left[\psi \star \left(\mathcal{T}^\dagger \mathcal{L}_N g^\varepsilon\right)\right](\boldsymbol{x})\right) \\ &= \left(\left[\psi \star \left(\mathcal{T}^\dagger \mathcal{L}_N g\right)\right](\boldsymbol{x})\right) + \mathbb{E}\left(\left[\psi \star \left(\mathcal{T}^\dagger \mathcal{L}_N \varepsilon\right)\right](\boldsymbol{x})\right) \\ &= f_\psi(\boldsymbol{x}) + \left[\psi \star \mathcal{T}^\dagger \left(\sum_{n=0}^{N} \sum_{k=-n}^{n} \sum_{l=1}^{L} \sum_{m=1}^{M} \mathbb{E}\varepsilon(l,m) B_{n,k}(\sigma_l, t_m) B_{n,k}\right)\right](\boldsymbol{x}) \\ &= f_\psi(\boldsymbol{x}) \end{aligned}$$

because $\mathbb{E}\varepsilon(l,m) = 0$ as assumed in Definition 4.1. □

Theorem 4.3. *Let $f \in C(\mathbb{S}^2)$, f_ψ be as given in (4.7), and f_ψ^ε be the estimator of f from Definition 4.1 where ε satisfies the conditions from there. Then the following decomposition of the MISE holds:*

$$\mathbb{E}\left\|f - f_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2 = \left\|f - f_\psi\right\|_{L^2(\mathbb{S}^2)}^2 + \mathbb{E}\left\|f_\psi - f_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2. \tag{4.11}$$

The MISE consists of the bias term

$$\|f - f_\psi\|^2_{L^2(\mathbb{S}^2)}$$

and the variance term

$$\mathbb{E}\left\|f_\psi - f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)}.$$

Proof. We have

$$\begin{aligned}
\mathbb{E}\left\|f - f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} &= \mathbb{E}\left\|\left(f - \mathbb{E}f_\psi^\varepsilon\right) - \left(f_\psi^\varepsilon - \mathbb{E}f_\psi^\varepsilon\right)\right\|^2_{L^2(\mathbb{S}^2)} \\
&= \mathbb{E}\left\|f - \mathbb{E}f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} + \mathbb{E}\left\|f_\psi^\varepsilon - \mathbb{E}f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} \\
&\quad -\mathbb{E}\left(2\mathrm{Re}\left\langle f - \mathbb{E}f_\psi^\varepsilon, f_\psi^\varepsilon - \mathbb{E}f_\psi^\varepsilon\right\rangle\right)
\end{aligned}$$

where Re denotes the real part of a complex number. Because the expected value and the inner product are linear, we observe that

$$\begin{aligned}
\mathbb{E}\left\langle f - \mathbb{E}f_\psi^\varepsilon, f_\psi^\varepsilon - \mathbb{E}f_\psi^\varepsilon\right\rangle &= \left\langle f - \mathbb{E}f_\psi^\varepsilon, \mathbb{E}\left(f_\psi^\varepsilon - \mathbb{E}f_\psi^\varepsilon\right)\right\rangle \\
&= \left\langle f - \mathbb{E}f_\psi^\varepsilon, \mathbb{E}f_\psi^\varepsilon - \mathbb{E}f_\psi^\varepsilon\right\rangle \\
&= 0.
\end{aligned}$$

The lemma above shows that $\mathbb{E}f_\psi^\varepsilon = f_\psi$ which finally implies (4.11). □

We are going to analyze the variance and the bias term of (4.11) separately.

4.2.1 Variance error

Theorem 4.4. *Let $f \in \mathrm{N}(\mathcal{T})^\perp$, $N \in \mathbb{N}$, $\psi \in L^2([-1,1])$, and the estimator f_ψ^ε as in Definition 4.1 using the hyperinterpolation $\mathcal{L}_N$ from* (4.6). *Then the variance term of* (4.11) *is given by*

$$\mathbb{E}\left\|f_\psi - f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} = \delta^2 \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\frac{\hat{\psi}(n)}{\lambda(n,k)}\right|^2 \sum_{l=1}^{L} \sum_{m=1}^{M} \left|w_{l,m} B_{n,k}(\sigma_l, t_m)\right|^2. \tag{4.12}$$

Proof. By Parseval's equality and the uncorrelatedness of ε from Definition 4.1,

$$\begin{aligned}
\mathbb{E}\left\|f_\psi - f_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2 &= \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \mathbb{E}\left| \frac{\hat{\psi}(n)}{\lambda(n,k)} \sum_{l=1}^{L}\sum_{m=1}^{M} w_{l,m}\varepsilon(l,m)\,\overline{B_{n,k}(\sigma_l,t_m)} \right|^2 \\
&= \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}(n)}{\lambda(n,k)} \right|^2 \cdot \\
&\qquad \sum_{l,l'=1}^{L}\sum_{m,m'=1}^{M} w_{l,m}w_{l',m'} B_{n,k}(\sigma_l,t_m)\,\overline{B_{n,k}(\sigma_{l'},t_{m'})}\mathbb{E}\varepsilon(l,m)\,\varepsilon(l',m') \\
&= \delta^2 \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}(n)}{\lambda(n,k)} \right|^2 \sum_{l=1}^{L}\sum_{m=1}^{M} \left|w_{l,m}B_{n,k}(\sigma_l,t_m)\right|^2
\end{aligned} \tag{4.13}$$

which proves (4.12). □

This theorem shows that the variance error (4.12) is independent of the particular function f. In fact, the variance error depends only on the variance δ of the data error, the hyperinterpolation $\mathcal{L}_N$ and the choice of the mollifier ψ.

Proposition 4.5. *Let the conditions of Theorem 4.4 hold. If the quadrature Q, on which the hyperinterpolation $\mathcal{L}_N$ is based, is exact for the functions $|B_{n,k}|^2$, $n \le N$, $|k| \le n$, and has constant weights, i.e. $w_{l,m} = 4\pi/(LM)$, then*

$$\mathbb{E}\left\|f_\psi - f_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2 = \frac{4\pi\delta^2}{LM} \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}(n)}{\lambda(n,k)} \right|^2 . \tag{4.14}$$

Proof. Let $w_{l,m} = 4\pi/(LM)$ and the quadrature be exact of degree $2N$. Since the functions $B_{n,k}$ are an orthonormal system and $|B_{n,k}|^2$ is a polynomial of degree $2n \le 2N$ for which Q is exact, we have by Theorem 4.4

$$\begin{aligned}
\sum_{l=1}^{L}\sum_{m=1}^{M} \left|w_{l,m}B_{n,k}(\sigma_l,t_m)\right|^2 &= \frac{4\pi}{LM} \sum_{l=1}^{L}\sum_{m=1}^{M} w_{l,m} \left|B_{n,k}(\sigma_l,t_m)\right|^2 \\
&= \frac{4\pi}{LM} \int_0^{2\pi}\int_{-1}^{1} \left|B_{n,k}(\sigma,t)\right|^2 \,\mathrm{d}t\,\mathrm{d}\sigma \\
&= \frac{4\pi}{LM}
\end{aligned}$$

and thus

$$\mathbb{E}\left\|f_\psi - f_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2 = \frac{4\pi\delta^2}{LM} \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}(n)}{\lambda(n,k)} \right|^2 .$$

□

Remark 4.6. The assumption in Proposition 4.5, that the quadrature weights should be constant, is rather strict. Let Q be the usual equidistant quadrature with respect to σ on the unit arc, and some quadrature on the unit interval as defined in Subsection 2.1.1.1 with respect to t. So $\sigma_l = 2\pi l/L$. We denote with t_m the nodes and with ω_m the weights of the quadrature in the t variable, so that $w_{l,m} = 2\pi\omega_m/L$. Using the definition of $B_{n,k}$ in (2.29), the inner sums from (4.13) become

$$\begin{aligned}\sum_{l=1}^{L}\sum_{m=1}^{M}\left|w_{l,m}B_{n,k}\left(\sigma_l,t_m\right)\right|^2 &= \frac{n+\frac{1}{2}}{2\pi}\sum_{m=1}^{M}\left(\frac{2\pi}{L}\right)^2\omega_m^2\left|\mathrm{e}^{\mathrm{i}k\sigma_l}P_n\left(t_m\right)\right|^2\\ &= \frac{2\pi}{L}\left(n+\frac{1}{2}\right)\sum_{m=1}^{M}\omega_m^2P_n\left(t_m\right)^2. \qquad (4.15)\end{aligned}$$

The numerical computation of (4.15) for $n \le N$ suggests that

$$\left(n+\frac{1}{2}\right)\sum_{m=1}^{M}\omega_m^2P_n\left(t_m\right)^2 \approx \frac{2}{M} \qquad (4.16)$$

for n and M being sufficiently large which can be seen in Figure 4.1 on page 59. So (4.16) gives a justification to believe that (4.14) is approximately true for the Gauss-Legendre and Clenshaw-Curtis quadrature. From Theorem 4.4, we know that (4.16) holds with "=" if the quadrature has equal weights.

In the following definition, we will make an assumption on $\mathcal{L}_N$. That assumption is a weaker version of (4.16).

Definition 4.7. For $N \in \mathbb{N}$, let $\mathcal{L}_N$ be an $L \times M$-point hyperinterpolation on $\mathbb{T} \times [-1,1]$ using the quadrature Q_N with positive weights $w_{l,m} > 0$ and the nodes (σ_l, t_m), as defined in (4.6). We say that $\mathcal{L}_N$ is applicable if there exist constants $\underline{\gamma}, \overline{\gamma} \in (0,\infty)$ such that the equation

$$\underline{\gamma}\frac{4\pi}{LM} \le \sum_{l=1}^{L}\sum_{m=1}^{M}\left|w_{l,m}B_{n,k}\left(\sigma_l,t_m\right)\right|^2 \le \overline{\gamma}\frac{4\pi}{LM}, \quad n=0,\dots,N,\ k=-n,\dots,n, \qquad (4.17)$$

holds for every sufficiently large N, Q_N is exact for $\mathit{\Pi}_{2N}\left(\mathbb{T}\times[-1,1]\right)$, and $LM \in \mathcal{O}\left(N^2\right)$ for $N \to \infty$.

Proposition 4.5 shows that equal-weight quadratures are applicable. In the following, we are going to prove the applicability of a hyperinterpolation based on the Fejér quadrature.

Lemma 4.8 (Fejér quadrature)**.** *Let t_m and ω_m, $m = 1,\dots,M$, be the nodes and weights of the the Fejér quadrature as defined in Subsection 2.1.1.1 and let P_n denote the n-th Legendre polynomial, as usual. Then*

$$\sum_{m=1}^{M}\omega_m^2\left|P_n\left(t_m\right)\right|^2 = \gamma_M^2\frac{\pi}{M}\int_{-1}^{1}\sqrt{1-t^2}\left|P_n\left(t\right)\right|^2\,\mathrm{d}t, \quad M\in\mathbb{N},\ n\le M-1,$$

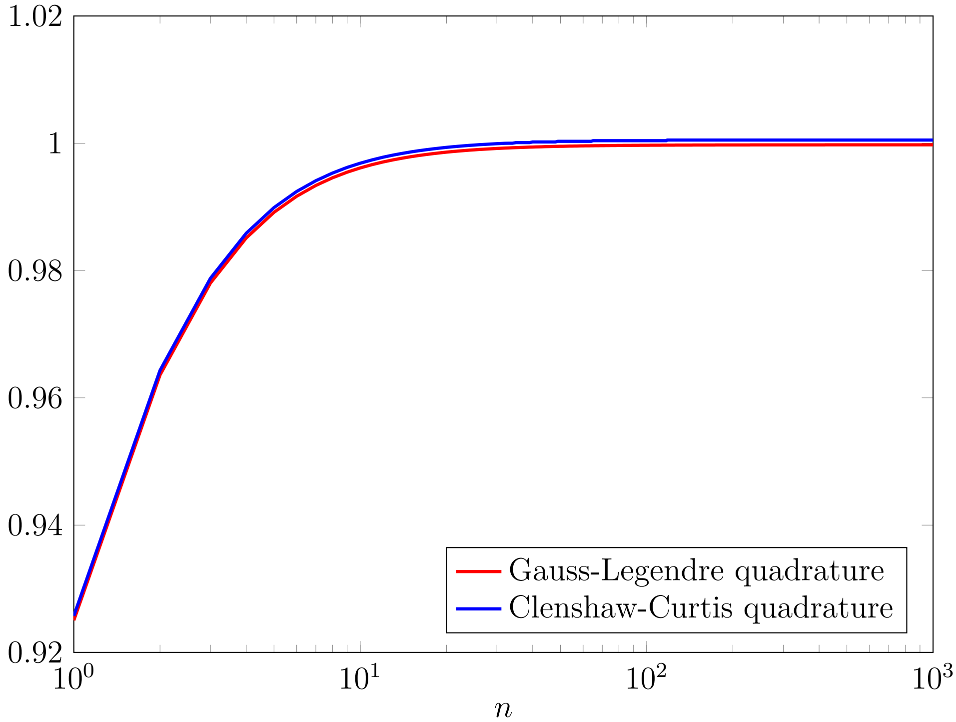

Figure 4.1: Numerical computation of $\frac{N}{2}\sum_{m=1}^{M}\left(n+\frac{1}{2}\right)\omega_m^2 P_n\left(t_m\right)^2$ from (4.15) with $N = 1000$ and $M = 2N$, where t_m, ω_m denote the nodes and weights of either the Gauss-Legendre or Clenshaw-Curtis quadrature. This figure suggests that (4.16) holds for large n. Note that the factor $N/2$ is equal to the right side of that equation.

where

$$0.9028233 = \underline{\gamma} < \gamma_M < \overline{\gamma} = 1.1789797, \quad M \in \mathbb{N}. \tag{4.18}$$

Proof. The Fejér quadrature uses the nodes (2.20)

$$t_m = \cos\left(\theta_m\right), \quad \theta_m = \frac{m\pi}{M+1}, \quad m = 1, \ldots, M$$

and the weights (2.22)

$$\omega_m = \frac{4}{M+1}\sin\left(\theta_m\right)\sum_{j=1}^{\lfloor M/2\rfloor}\frac{\sin\left(\left(2j-1\right)\theta_m\right)}{2j-1}, \quad m = 1, \ldots, M.$$

We want to derive an approximation of ω_m, therefore we define for $M \in \mathbb{N}$ the function

$$S_M(\theta) = \sum_{j=1}^{\lfloor M/2\rfloor}\frac{\sin\left(\left(2j-1\right)\theta\right)}{2j-1}, \quad \theta \in \left(-\pi, \pi\right), \tag{4.19}$$

which is equal to the sum in (2.22) where we replaced θ_m by θ. The function S_M is the $\lfloor M/2 \rfloor$-th partial sum of the Fourier series of the Heaviside step function

$$\theta \mapsto \frac{\pi}{4}\mathrm{sgn}\theta, \quad \theta \in (-\pi, \pi),$$

where sgn denotes the sign function, cf. [20, 1.442]. So $S_M(\theta)$ converges for $M \to \infty$ to the constant $\pi/4$ for all $\theta \in (0, \pi)$, but this convergence is not uniform. In fact, $S_M(\theta)$ oscillates near the boundary $\theta \in \{0, \pi\}$ of the interval. The fact that this oscillation does not decrease even for large M is known as the Gibbs phenomenon. In [26], it was shown that

$$\sum_{j=1}^{\lfloor M/2 \rfloor} \frac{\sin((2j-1)\theta_m)}{2j-1} = \frac{\pi}{4}\gamma_{M,m} \tag{4.20}$$

with some constants $\gamma_{M,m}$ satisfying

$$\underline{\gamma} < \gamma_{M,m} < \overline{\gamma}, \quad m = 1, \ldots, M, \ M \in \mathbb{N}.$$

Combining (2.22) and (4.20) yields

$$\omega_m = \frac{\pi}{M+1}\gamma_{M,m} \sin(\theta_m). \tag{4.21}$$

Up to the factor $\gamma_{M,m}$, the weights ω_m are equal to the weights of the Gauss-Chebyshev quadrature of the second kind (2.21), which uses the same nodes as the Fejér quadrature. Because the M-point Gauss-Chebyshev quadrature of the second kind is exact of degree $2M-1$ for the integral with respect to the measure $\sqrt{1-t^2}\,\mathrm{d}t$, and P_n is the Legendre polynomial of degree n, we have for $M \geq n+1$

$$\int_{-1}^{1} |P_n(t)|^2 \sqrt{1-t^2}\,\mathrm{d}t = \frac{\pi}{M+1}\sum_{m=1}^{M} (\sin(\theta_m))^2 |P_n(\cos\theta_m)|^2.$$

Considering this and (4.21) leads to

$$\begin{aligned}
\sum_{m=1}^{M} \omega_m^2 |P_n(t_m)|^2 &= \left(\frac{\pi}{M+1}\right)^2 \sum_{m=1}^{M} (\sin\theta_m)^2 \gamma_{M,m}^2 |P_n(\cos\theta_m)|^2 \\
&= \gamma_M^2 \frac{\pi}{M+1} \sum_{m=1}^{M} \frac{\pi}{M+1} (\sin\theta_m)^2 |P_n(\cos\theta_m)|^2 \\
&= \gamma_M^2 \frac{\pi}{M+1} \int_{-1}^{1} \sqrt{1-t^2} |P_n(t)|^2 \,\mathrm{d}t
\end{aligned}$$

where γ_M satisfies the same inequality as $\gamma_{M,m}$. □

Remark 4.9. The inequality (4.18) is asymptotically sharp for $m \in \{1, 2, M-1, M\}$ and $M \to \infty$ while $\gamma_{M,m}$ is considerably nearer to one when m is farther away from the extremal points 1 and M, which can be seen in [26]. Thus, also γ_M is closer to one than the lower and upper bounds $\underline{\gamma}$ and $\overline{\gamma}$, respectively.

Lemma 4.10. *For $n \to \infty$,*

$$\left(n+\frac{1}{2}\right)\int_{-1}^{1} |P_n(t)|^2 \sqrt{1-t^2}\,\mathrm{d}t = \frac{2}{\pi} + \mathcal{O}\left(n^{-1/2}\right). \tag{4.22}$$

Proof. By Stieltjes' formula (2.15), we have

$$\begin{aligned}
\int_{-1}^{1} |P_n(t)|^2 \sqrt{1-t^2}\,\mathrm{d}t &= \int_0^{\pi} \left(P_n(\cos\theta)\right)^2 (\sin\theta)^2\,\mathrm{d}\theta \\
&= \int_0^{\pi} \left(\frac{2}{\pi n}\frac{\left(\cos\left(\left(n+\frac{1}{2}\right)\theta-\frac{\pi}{4}\right)\right)^2}{\sin\theta} + \mathcal{O}\left((n\sin\theta)^{-3/2}\right)\right)(\sin\theta)^2\,\mathrm{d}\theta \\
&= \frac{2}{\pi n}\int_0^{\pi} \left(\cos\left(\left(n+\frac{1}{2}\right)\theta-\frac{\pi}{4}\right)\right)^2 \sin\theta\,\mathrm{d}\theta + \mathcal{O}\left(n^{-3/2}\right). \qquad (4.23)
\end{aligned}$$

It is left to evaluate the integral in (4.23). With the formula $(\cos x)^2 = (\cos 2x + 1)/2$, we observe that

$$\begin{aligned}
\left(\cos\left(\left(n+\frac{1}{2}\right)\theta-\frac{\pi}{4}\right)\right)^2 &= \frac{\cos\left((2n+1)\theta-\frac{\pi}{2}\right)+1}{2} \\
&= \frac{\sin\left((2n+1)\theta\right)+1}{2}.
\end{aligned}$$

Using this and the integral formula [20, 2.532.1], we do for $n \geq 1$ the calculation

$$\begin{aligned}
\int \left(\cos\left(\left(n+\frac{1}{2}\right)\theta-\frac{\pi}{4}\right)\right)^2 \sin\theta\,\mathrm{d}\theta &= \frac{1}{2}\int \left(\sin\left((2n+1)\theta\right)\sin\theta + \sin\theta\right)\,\mathrm{d}\theta \\
&= \frac{1}{2}\left(\frac{\cos(2n\theta)}{4n} - \frac{\sin\left((2n+2)\theta\right)}{2(2n+2)} - \cos\theta\right)
\end{aligned}$$

Finally, we calculate the definite integral

$$\begin{aligned}
\int_0^{\pi} \left(\cos\left(\left(n+\frac{1}{2}\right)\theta-\frac{\pi}{4}\right)\right)^2 \sin\theta\,\mathrm{d}\theta &= \frac{1}{2}\left[\frac{\cos(2n\theta)}{4n} - \frac{\sin\left((2n+2)\theta\right)}{2(2n+2)} - \cos\theta\right]_0^{\pi} \\
&= 1
\end{aligned}$$

from which the claimed formula (4.22) follows. □

Corollary 4.11. *Let, for $N \in \mathbb{N}$, $\mathcal{L}_N$ be the hyperinterpolation on $\mathbb{T} \times [-1,1]$ based on the quadrature that is the tensor product of the N–point equidistant quadrature on the torus and the $(N+1)$–point Fejér quadrature on the unit interval. Then $\mathcal{L}_N$ is applicable in the sense of Definition 4.7.*

Proof. This follows directly from the previous two lemmas and Remark 4.6. □

4.2.2 Bias error

With the triangle inequality, we split up the bias error into a smoothing and an aliasing error,

$$\|f - f_\psi\|_{L^2(\mathbb{S}^2)} \leq \|f - \psi \star f\|_{L^2(\mathbb{S}^2)} + \left\|\psi \star \left(f - \mathcal{T}^\dagger \mathcal{L}_N g\right)\right\|_{L^2(\mathbb{S}^2)}.$$

Smoothing error. If $f \in H^s(\mathbb{S}^2)$ for $s \geq 0$, then Proposition 2.11 implies that the smoothing error is bounded by

$$\|f - \psi \star f\|_{L^2(\mathbb{S}^2)} \leq \sup_{n \in \mathbb{N}_0} \left(\frac{\left|1 - \hat{\psi}(n)\right|}{\left(n + \frac{1}{2}\right)^s} \right) \|f\|_{H^s(\mathbb{S}^2)}. \tag{4.24}$$

Aliasing error. If the quadrature Q, on which $\mathcal{L}_N$ is based, is exact of degree $2N$, then by Theorem 2.2

$$\|g - \mathcal{L}_N g\|_{L^2(\mathbb{T} \times [-1,1])} \leq 2\sqrt{4\pi} \inf_{g \in \Pi_N(\mathbb{T} \times [-1,1])} \|f - g\|_{C(\mathbb{T} \times [-1,1])}.$$

To derive an upper bound for the quadrature error, we define the Sobolev-like space $H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1])$ as the closure of the set of polynomials

$$\operatorname{span}\{B_{n,k} : n \in \mathbb{N}_0,\ k = -n, \ldots, n,\ 2|\,(n+k)\}$$

with respect to the Sobolev-like norm

$$\|g\|_{H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1])} = \sum_{n=0}^{\infty} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \hat{g}(n,k) \frac{\left(n + \frac{1}{2}\right)^s}{\lambda(n,k)} \right|^2.$$

As a direct consequence of this definition,

$$\|f\|_{H^s(\mathbb{S}^2)} = \|\mathcal{T} f\|_{H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1])} \text{ for all } f \in H^s\left(\mathbb{S}^2\right) \cap \mathrm{N}(\mathcal{T})^\perp. \tag{4.25}$$

The following two lemmas are taken from [25] where they were proven for the Sobolev spaces $H^s(\mathbb{S}^2)$ on the two-sphere.

Lemma 4.12. *For $s > 1$, the Sobolev-like space $H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1])$ can be embedded continuously into the space of continuous functions $C(\mathbb{T} \times [-1,1])$. Let $N \in \mathbb{N}$, then there exists a constant $c > 0$ independent of N such that following estimates are valid.*

- *For $g \in H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1])$,*

$$\|g\|_{C(\mathbb{T}\times[-1,1])} \leq c\,\|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}, \quad g \in H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1]). \tag{4.26}$$

- *For $g \in \Pi_N(\mathbb{T} \times [-1,1])^{\perp}$,*

$$\|g\|_{C(\mathbb{T}\times[-1,1])} \leq c\left(N + \frac{1}{2}\right)^{1-s} \|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}. \tag{4.27}$$

  and

$$\|g\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \leq \left(N + \frac{1}{2}\right)^{-s} \|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}. \tag{4.28}$$

- *For $g \in \Pi_N(\mathbb{T} \times [-1,1])$,*

$$\|g\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \leq c\left(N + \frac{1}{2}\right)^{1/2} \|g\|_{L^2(\mathbb{T}\times[-1,1])}. \tag{4.29}$$

Proof. By (3.15),

$$\lambda(n,k) = \sqrt{\frac{\left(n+\frac{1}{2}\right)(n-k)!}{(n+k)!}}\,P_{n,k}(0), \quad n \in \mathbb{N}_0,\ k = -n,\ldots,n,\ 2|\,(n+k)\,.$$

From the definition of the spherical harmonics in (2.39) and the addition formula (2.35), we observe that

$$\begin{aligned}
\sum_{\substack{k=-n\\2|(n+k)}}^{n} |\lambda(n,k)|^2 &= \sum_{\substack{k=-n\\2|(n+k)}}^{n} \frac{\left(n+\frac{1}{2}\right)(n-k)!}{(n+k)!}\,|P_{n,k}(0)|^2 \\
&= \frac{4\pi}{2n+1}\sum_{\substack{k=-n\\2|(n+k)}}^{n} \left|Y_{n,k}(0,\frac{\pi}{2})\right|^2 \\
&\leq \frac{4\pi}{2n+1}\sum_{k=-n}^{n} \left|Y_{n,k}(0,\frac{\pi}{2})\right|^2 \\
&= 1.
\end{aligned}$$

Let $(\sigma,t) \in \mathbb{T} \times [-1,1]$ and $g \in H^s_{\mathcal{T}}(\mathbb{T} \times [-1,1])$, then, by the Cauchy-Schwartz inequality and the definition of $B_{n,k}$,

$$\begin{aligned}
|g(\sigma,t)| &= \left|\sum_{n=0}^{\infty}\sum_{\substack{k=-n\\2|(n+k)}}^{n}\hat{g}(n,k)\,B_{n,k}(\sigma,t)\right| \\
&= \sum_{n=0}^{\infty}\sum_{\substack{k=-n\\2|(n+k)}}^{n}\left|\hat{g}(n,k)\,\frac{\left(n+\frac{1}{2}\right)^{s}}{\lambda(n,k)}\right|\left|\sqrt{\frac{2n+1}{4\pi}}\mathrm{e}^{\mathrm{i}k\sigma}P_n(t)\lambda(n,k)\right|\left(n+\frac{1}{2}\right)^{-s} \\
&\leq \|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}\sqrt{\sum_{n=0}^{\infty}\frac{2n+1}{4\pi}\,|P_n(t)|^2\left(n+\frac{1}{2}\right)^{-2s}\sum_{\substack{k=-n\\2|(n+k)}}^{n}|\lambda(n,k)|^2} \\
&\leq \|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}\sqrt{\sum_{n=0}^{\infty}\left(n+\frac{1}{2}\right)^{1-2s}}.
\end{aligned}$$

This sum converges if and only if $1-2s<-1$ which is equivalent to $s>1$. Hence, there exists a constant $c>0$ depending only on s such that

$$\|g\|_{\infty}\leq c\,\|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}\,,\quad g\in H^s_{\mathcal{T}}\left(\mathbb{T}\times[-1,1]\right).$$

This shows the validity of (4.26).
Like in Section 2.1, we denote with $\mathcal{P}_N$ the L^2-orthogonal projection onto $\Pi_N\left(\mathbb{T}\times[-1,1]\right)$. Similarly to the first part of the proof, we observe that there exists a constant $c_1>0$ independent of N and g such that

$$\begin{aligned}
\|g-\mathcal{P}_N g\|_{C(\mathbb{T}\times[-1,1])} &= \left\|\sum_{n=N+1}^{\infty}\sum_{\substack{k=-n\\2|(n+k)}}^{n}\hat{g}(n,k)\,B_{n,k}\right\|_{C(\mathbb{T}\times[-1,1])} \\
&\leq \|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}\sqrt{\sum_{n=N+1}^{\infty}\left(n+\frac{1}{2}\right)^{1-2s}} \\
&\leq c_1\left(N+\frac{1}{2}\right)^{1-s}\|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}
\end{aligned}$$

which proves (4.27). Analogously, the calculation

$$\begin{aligned}
\|g-\mathcal{P}_N g\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} &= \sum_{n=N+1}^{\infty}\sum_{\substack{k=-n\\2|(n+k)}}^{n}\hat{g}(n,k)\,B_{n,k} \\
&\leq \left(N+\frac{1}{2}\right)^{-s}\|g-\mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}
\end{aligned}$$

shows (4.28).
By (3.22), we have for $g\in\Pi_N\left(\mathbb{T}\times[-1,1]\right)$,

$$\begin{aligned}
\|g\|^2_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} &= \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \hat{g}(n,k) \frac{1}{\lambda(n,k)} \right|^2 \\
&\leq \|g\|^2_{L^2(\mathbb{T}\times[-1,1])} \sup \left\{ \frac{1}{|\lambda(n,k)|^2} : n = 0,\ldots,N,\ |k| \leq n,\ 2|\,(n+k) \right\} \\
&\leq c_2 \left(N + \frac{1}{2} \right) \|g\|^2_{L^2(\mathbb{T}\times[-1,1])}
\end{aligned}$$

which shows (4.29). □

Theorem 4.13. *Let $s > 1$, $f \in H^s(\mathbb{S}^2) \cap \mathrm{N}(\mathcal{T})^{\perp}$, and $g = \mathcal{T} f$. Furthermore, let for every $N \in \mathbb{N}$, the quadrature used for the hyperinterpolation $\mathcal{L}_N$ be exact for $\Pi_N(\mathbb{T}\times[-1,1])$. Then there exists a constant $c > 0$ independent of f and N so that*

$$\left\| \mathcal{T}^{\dagger} (\mathcal{L}_N g - g) \right\|_{L^2(\mathbb{S}^2)} \leq c \left(N + \frac{1}{2} \right)^{3/2-s} \|f\|_{H^s(\mathbb{S}^2)}.$$

Proof. We first note that, because $s > 1$, g is continuous and hence $\mathcal{L}_N g$ exists. Since $\mathcal{P}_N g \in \Pi_N(\mathbb{T}\times[-1,1])$ and $\mathcal{L}_N p = p$ for $p \in \Pi_N(\mathbb{T}\times[-1,1])$, we have $\mathcal{L}_N \mathcal{P}_N = \mathcal{P}_N$ and

$$\begin{aligned}
\left\| \mathcal{T}^{\dagger} (\mathcal{L}_N g - g) \right\|_{L^2(\mathbb{S}^2)} &= \|\mathcal{L}_N g - g\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \\
&= \|\mathcal{L}_N (g - \mathcal{P}_N g) - (g - \mathcal{P}_N g)\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \\
&\leq \|\mathcal{L}_N (g - \mathcal{P}_N g)\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} + \|g - \mathcal{P}_N g\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])}. \quad (4.30)
\end{aligned}$$

Now we examine both terms of the equation (4.30). From (4.28), we observe that that

$$\|g - \mathcal{P}_N g\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \leq \left(N + \frac{1}{2} \right)^{-s} \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}. \quad (4.31)$$

For the first term of (4.30), we use the Lemma 4.12. Equation (4.29) yields that

$$\|\mathcal{L}_N (g - \mathcal{P}_N g)\|_{H^0_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \leq \left(N + \frac{1}{2} \right)^{1/2} \|\mathcal{L}_N (g - \mathcal{P}_N g)\|_{L^2(\mathbb{T}\times[-1,1])}. \quad (4.32)$$

By Theorem 2.2,

$$\|\mathcal{L}_N (g - \mathcal{P}_N g)\|_{L^2(\mathbb{T}\times[-1,1])} \leq \sqrt{4\pi}\, \|g - \mathcal{P}_N g\|_{C(\mathbb{T}\times[-1,1])}. \quad (4.33)$$

By (4.27), there exists a constant c_1 independent of g and N such that

$$\|g - \mathcal{P}_N g\|_{C(\mathbb{T}\times[-1,1])} \leq c_1 \left(N + \frac{1}{2} \right)^{1-s} \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}. \quad (4.34)$$

Combining (4.30), (4.31), (4.32), (4.33), and (4.34) yields

$$\left\| \mathcal{T}^{\dagger} (\mathcal{L}_N g - g) \right\|_{L^2(\mathbb{S}^2)} \leq c \left(N + \frac{1}{2} \right)^{3/2-s} \|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}$$

where c is some positive constant. Since $\mathcal{P}_N$ is a projection, $\|g - \mathcal{P}_N g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])} \leq \|g\|_{H^s_{\mathcal{T}}(\mathbb{T}\times[-1,1])}$, which finally proves the theorem. □

With this result, we are able to show the following estimate for the aliasing error.

Corollary 4.14. *Let the conditions from Theorem 4.13 hold. If* $\left|\hat{\psi}(n)\right| \leq 1$ *for all* $n \in \mathbb{N}$ *and* $f \in H^s(\mathbb{S}^2) \cap \mathrm{N}(\mathcal{T})^\perp$, *then the aliasing error is bounded by*

$$\left\|\psi \star \mathcal{T}^\dagger (g - \mathcal{L}_N g)\right\|_{L^2(\mathbb{S}^2)} \leq c\left(N + \frac{1}{2}\right)^{3/2-s} \|f\|_{H^s(\mathbb{S}^2)}. \tag{4.35}$$

Proof. This is a direct consequence of Theorem 4.13 and the Funk-Hecke formula (2.41). □

Combining the decomposition from Theorem 4.3 with the bounds for the variance error, smoothing error and aliasing error from (4.14), (4.24) and (4.35), respectively, yields that for $f \in \mathscr{F}(s, S)$,

$$\begin{aligned} \mathbb{E}\left\|f - f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} \leq & \left(\sup_{n\in 2\mathbb{N}_0} \frac{\left|1-\hat{\psi}(n)\right|}{\left(n+\frac{1}{2}\right)^s} S + c\left(N+\frac{1}{2}\right)^{3/2-s} S\right)^2 \\ & + \delta^2 \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\frac{\hat{\psi}(n)}{\lambda(n,k)}\right|^2 \sum_{l=1}^{L}\sum_{m=1}^{M} \left|w_{l,m} B_{n,k}(\sigma_l, t_m)\right|^2. \end{aligned} \tag{4.36}$$

4.3 Optimal mollifiers

For a positive number s, we define the following family of mollifiers

$$\psi_{\tilde{N}} = \sum_{n=0}^{\tilde{N}} \frac{2n+1}{4\pi} \hat{\psi}_{\tilde{N}}(n) P_n, \quad \tilde{N} \geq 0 \tag{4.37}$$

by its Fourier coefficients

$$\hat{\psi}_{\tilde{N}}(n) = \left(1 - \left(\frac{n+\frac{1}{2}}{\tilde{N}+\frac{1}{2}}\right)^s\right), \quad n = 0, \ldots, \left\lfloor \tilde{N} \right\rfloor. \tag{4.38}$$

The so-defined mollifiers $\psi_{\tilde{N}}$ are polynomials of degree $\left\lfloor \tilde{N} \right\rfloor$. Later, in Theorem 4.17, we will show that this family of mollifiers is in a certain way optimal for the reconstruction of functions $f \in \mathscr{F}(s, S)$.

Similar to these mollifiers $\psi_{\tilde{N}}$, we can define the Dirichlet kernel

$$\psi_{\tilde{N}}^{\mathrm{Dir}} = \sum_{n=0}^{\tilde{N}} \frac{2n+1}{4\pi} P_n$$

of order $\tilde{N} \in \mathbb{N}_0$. The Dirichlet kernels are often used for this kind of regularization introduced in the previous section, but that the Dirichlet kernel oscillates more than that from (4.37), which can be seen in Figure 4.2 on page 67.

In the following, we will prove the asymptotic optimality of the family $\{\psi_{\tilde{N}}\}_{\tilde{N}}$.

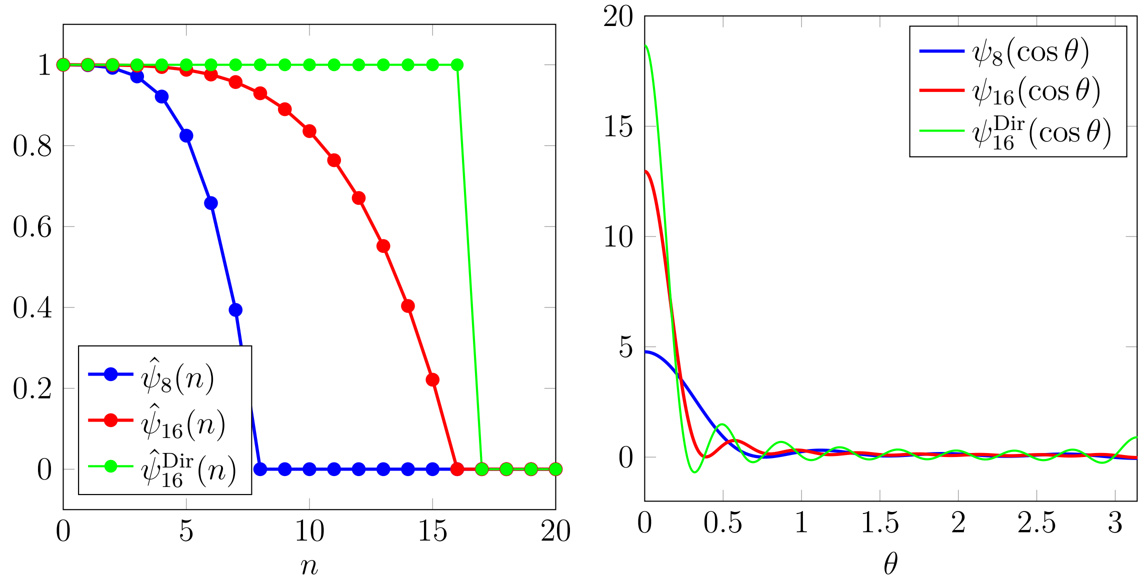

Figure 4.2: The mollifier $\psi_{\tilde{N}}(\cos\theta)$ as defined in (4.38) and its Fourier coefficients, with the parameters $\tilde{N} = 8$ (in blue), $\tilde{N} = 16$ (in red), and $s = 4$. We have also plotted the Dirichlet kernel for $\tilde{N} = 16$ in green. It can be seen that the Dirichlet kernel oscillates way more than $\psi_{\tilde{N}}$ on the whole interval. For $\tilde{N} \to \infty$, both the mollifiers $\psi_{\tilde{N}}$ and $\psi^{\mathrm{Dir}}_{\tilde{N}}$ behave like an approximation of the Dirac delta distribution centered at $\theta = 0$.

Lemma 4.15. *Let $\psi \in L^2[-1,1]$, $s > 1$ and $S > 0$. Furthermore, let the estimator f^{ε}_{ψ} of the function $f \in C(\mathbb{S}^2)$ and the noise ε be as in Definition 4.1, and the quadrature used for the hyperinterpolation $\mathcal{L}_N$ be exact for $\Pi_N(\mathbb{T} \times [-1,1])$. If N is sufficiently large, then there exists an index $\tilde{N}^* \geq 0$ for which the maximum risk (4.9) of the estimator f^{ε}_{ψ} is bounded from below by*

$$\sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\left\|f - f^{\varepsilon}_{\psi}\right\|^2_{L^2(\mathbb{S}^2)} \geq \delta^2 \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\frac{\hat{\psi}(n)}{\lambda(n,k)}\right|^2 \sum_{l=1}^{L} \sum_{m=1}^{M} \left|w_{l,m} B_{n,k}(\sigma_l, t_m)\right|^2 + \frac{S^2}{\left(\tilde{N}^* + \frac{1}{2}\right)^{2s}}. \tag{4.39}$$

The lower bound from (4.39) is the same as the upper bound from (4.36) without the aliasing error.

Proof. We define the set

$$\mathscr{N} = \left\{\tilde{N} \geq 0 : \hat{\psi}_{\tilde{N}}(n) \leq \left|\hat{\psi}(n)\right| \forall n \in \mathbb{N}_0\right\} \tag{4.40}$$

which is not empty because $0 \in \mathscr{N}$. It is easy to see that $\mathscr{N}$ is closed. Since the mollifier $\psi \in L^2[-1,1]$ is integrable, its Fourier coefficients $\hat{\psi}(n)$ vanish as $n \to \infty$, so $\mathscr{N}$ is bounded. Hence, there exists an $\tilde{N}^* = \max \mathscr{N}$. Now we choose an integer $n^* \in \mathbb{N}_0$ such that $\hat{\psi}_{\tilde{N}^*}(n^*) = \left|\hat{\psi}(n^*)\right|$. If n^* is even, we define the function

$$f^* = \frac{S}{\left(n^* + \frac{1}{2}\right)^s} Y_{n^*,0} \in \mathscr{P}_{\lfloor \tilde{N}^* \rfloor}\left(\mathbb{S}^2\right),$$

if n^* is odd then we replace $Y_{n^*,0}$ by $Y_{n^*,1}$ in the definition of f^*. In both cases, f^* is in the orthogonal complement of $\mathrm{N}(\mathcal{T})$ and has an H^s norm of S. Then, by (4.11),

$$\sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\left\|f - f_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2 \geq \mathbb{E}\left\|f^* - (f^*)_\psi^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2 = \left\|f^* - f_\psi^*\right\|_{L^2(\mathbb{S}^2)}^2 + \mathbb{E}\left\|f_{\psi_{\tilde{N}^*}} - f_{\psi_{\tilde{N}^*}}^\varepsilon\right\|_{L^2(\mathbb{S}^2)}^2.$$

The bias term from (4.11) with respect to f^* reads

$$\left\|f^* - f_\psi^*\right\|_{L^2(\mathbb{S}^2)}^2 = \left\|f^* - \psi \star f^*\right\|_{L^2(\mathbb{S}^2)}^2 = \left\|f^* - \psi_{\tilde{N}^*} \star f^*\right\|_{L^2(\mathbb{S}^2)}^2$$

The aliasing error $\left\|\psi \star \mathcal{T}^\dagger\left(\mathcal{T} f^* - \mathcal{L}_N \mathcal{T} f^*\right)\right\|_{L^2(\mathbb{S}^2)}$ vanishes if N is sufficiently large since f^* is a polynomial. Using the Funk-Hecke formula (2.41) and the definition of $\psi_{\tilde{N}^*}$, we obtain

$$\begin{aligned}
\left\|f^* - \psi_{\tilde{N}^*} \star f^*\right\|_{L^2(\mathbb{S}^2)}^2 &= \sum_{n=0}^{N} \sum_{k=-n}^{n} \left|\left(1 - \hat{\psi}_{\tilde{N}^*}(n^*)\right) \hat{f}(n,k)\right|^2 \\
&= \left(1 - \hat{\psi}_{\tilde{N}^*}(n^*)\right)^2 \frac{S^2}{\left(n^* + \frac{1}{2}\right)^{2s}} \\
&= \left(1 - \hat{\psi}(n^*)\right)^2 \frac{S^2}{\left(\tilde{N}^* + \frac{1}{2}\right)^{2s}} \\
&= \left\|f^* - \psi \star f^*\right\|_{L^2(\mathbb{S}^2)}^2.
\end{aligned}$$

By the definition of $\tilde{N}^*$, we have $\left|\hat{\psi}(n)\right| \geq \hat{\psi}_{\tilde{N}^*}(n)$ for all $n \in \mathbb{N}_0$. Hence, Theorem 4.4 implies that the variance error of f^* with the mollifier $\psi_{\tilde{N}^*}$ is not smaller than with ψ. □

The following lemma gives us an asymptotic expression of the sum in the variance error (4.14) for the mollifier $\psi_{\tilde{N}}$ from (4.37).

Lemma 4.16. *Let $\psi_{\tilde{N}}$ the mollifier from* (4.38)*. For $\tilde{N} \to \infty$,*

$$\sum_{n=0}^{\tilde{N}} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\frac{\hat{\psi}_{\tilde{N}}(n)}{\lambda(n,k)}\right|^2 \simeq d_s \tilde{N}^3. \tag{4.41}$$

where

$$d_s = \frac{\pi^3 s^2}{3(2s+3)(s+3)}, \tag{4.42}$$

Proof. At first, we compute the sum

$$\begin{aligned}\sum_{\substack{k=-n\\2|(n+k)}}^{n}\left|\frac{1}{\lambda(n,k)}\right|^2 &\simeq \frac{\pi}{2}\sum_{\substack{k=-n\\2|(n+k)}}^{n}\sqrt{n^2-k^2}\\ &\simeq \frac{\pi}{4}\int_{-n}^{n}\sqrt{n^2-k^2}\,\mathrm{d}k\\ &= \frac{\pi^2}{8}n^2\end{aligned}$$

where we have used the asymptotic expression of the singular values $\lambda(n,k)$ in Theorem 3.11. Now we can insert the Fourier coefficients of the mollifier (4.38) and see that, for $\tilde{N}\to\infty$,

$$\begin{aligned}\sum_{n=0}^{\tilde{N}}\left|\psi_{\tilde{N}}(n)\right|^2\sum_{\substack{k=-n\\2|(n+k)}}^{n}\left|\frac{1}{\lambda(n,k)}\right|^2 &\simeq \frac{\pi^2}{8}\sum_{n=0}^{\tilde{N}}n^2\left(1-\left(\frac{n+\frac{1}{2}}{\tilde{N}+\frac{1}{2}}\right)^s\right)^2\\ &\simeq \frac{\pi^2}{8}\int_{0}^{\tilde{N}}n^2\left(1-\left(\frac{n+\frac{1}{2}}{\tilde{N}+\frac{1}{2}}\right)^s\right)^2\mathrm{d}n\\ &\simeq \frac{\pi^2 s^2}{12(2s+3)(s+3)}\tilde{N}^3.\end{aligned}$$

□

After all that preparation, we can finally state our main theorem.

Theorem 4.17. *Let $s>(1+\sqrt{10})/2$, $S>0$, and $\delta>0$. For every $N\in\mathbb{N}$, let $\mathcal{L}_N$ be an $L\times M$-point hyperinterpolation, as defined in* (4.6)*, that is applicable in the sense of Definition 4.7. Furthermore, let the data error ε satisfy the conditions from Definition 4.1 and have the standard deviation δ. Then the family $\{\psi_{\tilde{N}}\}_{\tilde{N}}$, that was defined in* (4.38)*, is an asymptotically optimal family of mollifiers for the estimator f_ψ^ε for the inversion of the circular average transform $\mathcal{T}$ of functions f belonging to the class $\mathscr{F}(s,S)$. In particular, for $L,M\to\infty$, there exist parameters $\tilde{N}(L,M)$ such that*

$$\sup_{f\in\mathscr{F}(s,S)}\mathbb{E}\left\|f-f^\varepsilon_{\psi_{\tilde{N}(L,M)}}\right\|^2_{L^2(\mathbb{S}^2)}\simeq\inf_{\psi\in L^2[-1,1]}\sup_{f\in\mathscr{F}(s,S)}\mathbb{E}\left\|f-f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)}.$$

Furthermore, the minimax risk of the MISE is asymptotically for $L, M \to \infty$ bounded by

$$\begin{aligned}
&\left(\underline{\gamma}\frac{\delta^2 d_s}{LM}\left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{-2s}{2s+3}}\right)(LM)^{-\frac{2s}{2s+3}} \\
\lesssim\ & \inf_{\psi \in L^2[-1,1]} \sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\left\|f - f_\psi^\varepsilon\right\|^2 \qquad (4.43)\\
\lesssim\ & \left(\overline{\gamma}\delta^2 d_s\left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\delta^2 d_s}\right)^{\frac{-2s}{2s+3}}\right)(LM)^{-\frac{2s}{2s+3}}.
\end{aligned}$$

Proof. If L, M are sufficiently large, then, by Lemma 4.15, there exists an $\tilde{N} > 0$ such that

$$\begin{aligned}
&\inf_{\psi \in L^2[-1,1]} \sup_{f \in \mathscr{F}(s,S)} \mathbb{E}\left\|f - f_\psi^\varepsilon\right\|^2_{L^2(\mathbb{S}^2)} \qquad (4.44)\\
\geq\ & \delta^2 \sum_{n=0}^{\tilde{N}} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\frac{\hat{\psi}_{\tilde{N}}(n)}{\lambda(n,k)}\right|^2 \sum_{l=1}^{L}\sum_{m=1}^{M} \left|w_{l,m} B_{n,k}(\sigma_l, t_m)\right|^2 + \frac{S^2}{\left(\tilde{N} + \frac{1}{2}\right)^{2s}}.
\end{aligned}$$

The main idea behind the proof of this theorem is, to calculate the the parameter $\tilde{N}$ such that the variance and smoothing error are about equal and to show that the aliasing error is asymptotically lower when s is sufficiently large.

We minimize the right side of (4.44) with respect to $\tilde{N}$ and denote the optimal argument with $\tilde{N}(L, M)$.

Plugging the conditions (4.17) and (4.41) into the equation (4.12) for the variance term yields

$$\begin{aligned}
&\delta^2 \sum_{n=0}^{N} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left|\frac{\hat{\psi}(n)}{\lambda(n,k)}\right|^2 \sum_{l=1}^{L}\sum_{m=1}^{M} \left|w_{l,m} B_{n,k}(\sigma_l, t_m)\right|^2 + \frac{S^2}{\left(\tilde{N} + \frac{1}{2}\right)^{2s}} \\
\gtrsim\ & \underline{\gamma}\frac{\delta^2 d_s \tilde{N}^3}{LM} + \frac{S^2}{\left(\tilde{N} + \frac{1}{2}\right)^{2s}}. \qquad (4.45)
\end{aligned}$$

In order to minimize the right side of (4.45) with respect to $\tilde{N}$, we make its derivative with respect to $\tilde{N}$ vanish:

$$\begin{aligned}
(-2s)\, S^2 \tilde{N}^{-2s-1} + 3\underline{\gamma}\frac{\delta^2 d_s}{LM}\tilde{N}^2 &= 0 \\
\Leftrightarrow 3\underline{\gamma}\frac{3\delta^2 d_s}{LM}\tilde{N}^{2s+3} &= 2sS^2.
\end{aligned}$$

Hence, the right side of (4.45) is asymptotically minimized for $L, M \to \infty$ by

$$\tilde{N}_1(L, M) = \left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s} LM\right)^{\frac{1}{2s+3}}.$$

Plugging this value for $\tilde{N}_1(L,M)$ into (4.39) yields the following asymptotically lower bound of the maximum risk

$$\begin{aligned}
&\delta^2 \sum_{n=0}^{\tilde{N}_1(L,M)} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}_{\tilde{N}_1(L,M)}(n)}{\lambda(n,k)} \right|^2 \sum_{l=1}^{L} \sum_{m=1}^{M} |w_{l,m} B_{n,k}(\sigma_l, t_m)|^2 + \frac{S^2}{\left(\tilde{N}_1(L,M) + \frac{1}{2}\right)^{2s}} \\
\gtrsim\ & \left(\underline{\gamma} \delta^2 d_s \left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s} \right)^{\frac{3}{2s+3}} + S^2 \left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s} \right)^{\frac{-2s}{2s+3}} \right) (LM)^{\frac{-2s}{2s+3}} . \qquad (4.46)
\end{aligned}$$

Plugging this equation into (4.44) proves the first inequality of (4.43). From this, we also see that, for the choice $\tilde{N}_1(L,M)$, both terms of the sum in the right side of (4.45) decrease of the order $(LM)^{\frac{-2s}{2s+3}}$.

As we did above to derive (4.45), we can insert the upper bound for the variance error and obtain

$$\begin{aligned}
&\delta^2 \sum_{n=0}^{\tilde{N}} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}_{\tilde{N}}(n)}{\lambda(n,k)} \right|^2 \sum_{l=1}^{L} \sum_{m=1}^{M} |w_{l,m} B_{n,k}(\sigma_l, t_m)|^2 + \frac{S^2}{\left(\tilde{N} + \frac{1}{2}\right)^{2s}} \\
\lesssim\ & \overline{\gamma} \frac{\delta^2 d_s \tilde{N}^3}{LM} + \frac{S^2}{\left(\tilde{N} + \frac{1}{2}\right)^{2s}}, \qquad (4.47)
\end{aligned}$$

which is exactly like the equation (4.45) for the lower bound except that $\underline{\gamma}$ is replaced by $\overline{\gamma}$. Minimizing the right side of (4.47) with respect to $\tilde{N}$ gives us the argument

$$\tilde{N}_2(L,M) = \left(\frac{2sS^2}{3\overline{\gamma}\delta^2 d_s} LM \right)^{\frac{1}{2s+3}} .$$

Analogously to the derivation of (4.46), we plug $\tilde{N}_2(L,M)$ into (4.47) and observe that

$$\begin{aligned}
&\delta^2 \sum_{n=0}^{\tilde{N}_2(L,M)} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}_{\tilde{N}_2(L,M)}(n)}{\lambda(n,k)} \right|^2 \sum_{l=1}^{L} \sum_{m=1}^{M} |w_{l,m} B_{n,k}(\sigma_l, t_m)|^2 + \frac{S^2}{\left(\tilde{N}_2(L,M) + \frac{1}{2}\right)^{2s}} \\
\lesssim\ & c\,(LM)^{\frac{2s}{2s+3}} \qquad (4.48)
\end{aligned}$$

where c is some positive constant. According to (4.46) and (4.48), we have for the real minimizer $\tilde{N}(L,M)$ that

$$\begin{aligned}
&\delta^2 \sum_{n=0}^{\tilde{N}(L,M)} \sum_{\substack{k=-n \\ 2|(n+k)}}^{n} \left| \frac{\hat{\psi}_{\tilde{N}(L,M)}(n)}{\lambda(n,k)} \right|^2 \sum_{l=1}^{L} \sum_{m=1}^{M} |w_{l,m} B_{n,k}(\sigma_l, t_m)|^2 + \frac{S^2}{\left(\tilde{N}(L,M) + \frac{1}{2}\right)^{2s}} \\
\in\ & \Theta\left((LM)^{\frac{2s}{2s+3}} \right)
\end{aligned}$$

and

$$\tilde{N}(L,M) \in \Theta\left((LM)^{\frac{2s}{2s+3}}\right) \tag{4.49}$$

as $L, M \to \infty$.

In the second part of the proof, we derive the upper bound for the maximum risk. By (4.36), (4.41), and (4.35), we have for $f \in \mathscr{F}(s,S)$

$$\begin{aligned}
\mathbb{E}\left\|f - f^{\varepsilon}_{\psi_{\tilde{N}}}\right\|^2 &\leq \left(\sup_{n\in\mathbb{N}_0} \frac{\left|1-\hat{\psi}_{\tilde{N}}(n)\right|}{\left(n+\frac{1}{2}\right)^s} S + \left\|\psi \star \mathcal{T}^{\dagger}\left(g - \mathcal{L}_N g\right)\right\|_{L^2(\mathbb{S}^2)}\right)^2 \\
&\quad + \mathbb{E}\left\|f_{\psi_{\tilde{N}}} - f^{\varepsilon}_{\psi_{\tilde{N}}}\right\|^2_{L^2(\mathbb{S}^2)} && (4.50) \\
&\lesssim \left(\frac{S}{\tilde{N}^s} + c\left(N+\frac{1}{2}\right)^{3/2-s} S\right)^2 + \overline{\gamma}\frac{\delta^2 d_s \tilde{N}^3}{LM}. && (4.51)
\end{aligned}$$

Now we plug $\tilde{N}_1(L,M)$ from the first part of the proof into (4.51) and we see that

$$\begin{aligned}
\mathbb{E}\left\|f - f^{\varepsilon}_{\psi_{\tilde{N}}}\right\|^2_{L^2(\mathbb{S}^2)} &\lesssim \left(S\left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s}\right)^{\frac{-s}{2s+3}} (LM)^{\frac{-s}{2s+3}} + c\left(N+\frac{1}{2}\right)^{3/2-s} S\right)^2 && (4.52) \\
&\quad + \overline{\gamma}\delta^2 d_s \left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s}\right)^{\frac{3}{2s+3}} (LM)^{\frac{-2s}{2s+3}}.
\end{aligned}$$

We have assumed that $LM \in \mathcal{O}(N^2)$ and

$$s > \frac{1+\sqrt{10}}{2} \Rightarrow \frac{3}{2} - s < \frac{-s}{2s+3}.$$

Hence, only those terms in the sum in (4.52) that grow like $(LM)^{-2s/(2s+3)}$ play a significant role as L and M go to infinity. Thus, the aliasing error is asymptotically negligible. This yields to the upper bound

$$\begin{aligned}
&\lim_{L,M\to\infty} \sup_{f\in\mathscr{F}(s,S)} \mathbb{E}\left\|f - f^{\varepsilon}_{\psi_{\tilde{N}_1(L,M)}}\right\|^2_{L^2(\mathbb{S}^2)} \cdot (LM)^{\frac{2s}{2s+3}} \\
\leq\ & \overline{\gamma}\delta^2 d_s \left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s}\right)^{\frac{3}{2s+3}} + S^2\left(\frac{2sS^2}{3\underline{\gamma}\delta^2 d_s}\right)^{\frac{-2s}{2s+3}}, && (4.53)
\end{aligned}$$

which proves the second inequality of (4.43).

It is left to show the optimality of the family of mollifiers $\psi_{\tilde{N}}$ to complete the proof. When we insert the optimal value $\tilde{N}(L,M)$ into both (4.44) and (4.51), those two bounds coincide except for the aliasing term, but since $\tilde{N}(L,M)$ grows of the same order as $\tilde{N}_1(L,M)$, the aliasing part is again negligible. □

This theorem is somehow remarkable. It shows that, if we want to minimize the maximum risk asymptotically over all mollifiers $\psi \in L^2([-1,1])$, it suffices to minimize it over the family $\psi_{\tilde{N}}$ which only depends on one real parameter $\tilde{N} \geq 0$. Additionally, this theorem shows that the MISE decreases of the order $(LM)^{\frac{-2s}{2s+3}}$ for a good choice of $\tilde{N}(L,M)$. So, the higher the smoothness parameter s is, the faster the MISE decreases. For $s \to \infty$, we have $\frac{-2s}{2s+3} \to -1$.

Remark 4.18. Taking a closer look at the last proof shows the following. If we drop the condition for s and replace it by $s > 3/2$, then this method is still converging, i.e.

$$\lim_{L,M\to\infty} \sup_{f\in\mathscr{F}(s,S)} \mathbb{E} \left\| f - f^{\varepsilon}_{\psi_{\tilde{N}(L,M)}} \right\|^2_{L^2(\mathbb{S}^2)} = 0.$$

This can be seen in (4.52) where it follows from the fact that the aliasing error is in $\mathcal{O}\left((LM)^{3/2-s}\right)$ and thus the MISE vanishes for $L, M \to \infty$. But the aliasing error may be the dominating part of the MISE if s is smaller than it was required in the theorem.

Remark 4.19. Theorem 4.17 gives us a near-optimal choice $\tilde{N}_1(L,M)$ for the regularization parameter $\tilde{N}$ for a function f with $\|f\|_{H^s(\mathbb{S}^2)} = S$. In practice, the Sobolev norm of the solution f is usually unknown, so $\tilde{N}_1(L,M)$ cannot be computed as in the theorem. However, there exist a variety of heuristic methods to choose the regularization parameter $\tilde{N}$. A famous one is the L-curve method which was introduced in [23], that consists of creating a log-log plot of the norm of the regularized solution $f^{\varepsilon}_{\psi_{\tilde{N}}}$ versus the norm of the residual $f - f^{\varepsilon}_{\psi_{\tilde{N}}}$ for different values of $\tilde{N}$. The resulting graph is usually shaped like the letter “L” and the the regularization parameter $\tilde{N}$ is chosen to be at the “corner” of the graph. An overview on different parameter choice methods can be found in [3].

Remark 4.20. Instead of the mollifier method, we could have also used the approach of Fourier multipliers. This means that in (4.4) the coefficients $\hat{\psi}(n)$, that depend only on n, are replaced by the Fourier multipliers $\hat{\psi}(n,k)$ depending on n and k. So, despite being very similar, this approach is slightly more general. However, in the case of this thesis, it would have led to the same results as for the mollifier method. This is, because the shape of the optimal mollifier $\psi_{\tilde{N}}$ only depends on the Sobolev space H^s which can be seen in Lemma 4.15. More precisely, in the definition of the Sobolev norm in (2.40), the Fourier coefficients $\hat{f}(n,k)$ are multiplied by a factor that does only depend on n and not on k.

4.4 Photoacoustic tomography

This section is about an application of the circular average transform in photoacoustic tomography (PAT) (see also [56, 10]). PAT is based on the photoacoustic effect which states that a medium generates sound waves when absorbing electromagnetic waves. PAT is about reconstructing an image of the examined object by detecting the pressure waves outside the object during a time interval. The detected pressure is used to calculate a

3D image of the initial pressure in the object. PAT has several advantages over other technologies. Optical tomography provides high-resolution images but these are limited to a low depth of around 1 mm. Ultrasound imaging has a greater depth than optical tomography while the resolution is lower. PAT combines the high resolution and imaging depth of the former methods. Furthermore, PAT can excite different molecules at different optical wavelengths and therefore shows some information about the chemical composition. Since the 2000s, PAT has found many applications in biomedical imaging including cancer diagnosis and vascular imaging.

There are two main set-ups for the detectors used in PAT: A common way is to use piezoelectric crystals that measure the pressure at some points outside an object, but the resolution of this method is obviously limited to the size of the piezoelectric elements. A different approach makes use of optic pressure sensors that measure the "integral" of the pressure along the sensor. We describe a setting from [59] which uses optical sensors that integrate the pressure along circles. These circles lie on the sphere with their poles on the equator and can be written as $C(\sigma, t)$ like in (3.8). The detector measures the integrated pressure along the circles $C(0, t_m)$, $m = 1, \ldots, M$ at different times τ. This procedure is repeated L times and the detector is rotated around the x_3 axis each time, so we measure the pressure integrated along the circles $C(\sigma_l, t_m)$, $, l = 1, \ldots, L$, $m = 1, \ldots, M$ for different times τ.

As shown in [59], the reconstruction of the initial pressure $f \in C(\mathbb{R}^3)$ can be done in two steps. The first step is basically solving the wave equation. This step gives us the circular average transform $\mathcal{T}_r f(\sigma, t)$ of the initial pressure on spheres $r\mathbb{S}^2$ with radius $r \le r_0$. The second step consists of the reconstruction of the initial pressure f from its circular averages for all spheres with radius $r \le r_0$.

4.5 Practical computation

The algorithm for the computation of the inverse circular average transform is the transpose to that of Section 3.5. Suppose we are given a function

$$g = \mathcal{T} f \in \Pi_N(\mathbb{T} \times [-1, 1]).$$

In the first step, we compute the Fourier coefficients

$$\widehat{\mathcal{T} f}(n, k) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} \int_{0}^{2\pi} g(\sigma, t)\, \mathrm{e}^{-ik\sigma} p_n(t)\, \mathrm{d}\sigma\, \mathrm{d}t.$$

For the numerical computation of $\widehat{\mathcal{T} f}$, we use a quadrature Q from (4.5) with equidistant nodes

$$\sigma_l = \frac{2\pi l}{L}, \quad l = 1, \ldots, L = 2N + 1$$

in the first variable, some arbitrary nodes

$$t_m, \quad m = 1, \ldots, M$$

in the second variable, and the weights

$$w_{l,m} = \frac{2\pi}{L}\omega_m, \quad l = 1, \ldots, L,\ m = 1, \ldots, M$$

and get

$$\widehat{\mathcal{T}f}(n,k) = \frac{1}{\sqrt{2\pi}} \sum_{l=1}^{L} \sum_{m=1}^{M} w_{l,m} g\left(\sigma_l, t_m\right) \mathrm{e}^{-\mathrm{i}k\sigma_l} p_n(t_m) \tag{4.54}$$

for $n = 0, \ldots, N,\ k = -n, \ldots, n$. Like we did in Section 3.5, we denote the inner sum in (4.54) with $A(k,m)$ and see

$$\widehat{\mathtt{Tf}}(n,k) = \frac{2\pi}{L}\sqrt{\frac{2n+1}{4\pi}} \sum_{m=1}^{M} \omega_m P_n(t_m) A(k,m).$$

We can rewrite A using the DFT from (2.27)

$$\begin{aligned} A(k-N,m) &= \sum_{l=0}^{L-1} g(\sigma_l, t_m) \mathrm{e}^{-2\pi \mathrm{i} lk/L} \mathrm{e}^{2\pi \mathrm{i} lN/L} \\ &= \left[\mathrm{DFT}\left(g(\sigma_{(\cdot)}, t_m)\right) \mathrm{e}^{2\pi \mathrm{i}(\cdot)N/L}\right](k), \quad k = 0, \ldots, 2N,\ m = 1, \ldots, M. \end{aligned}$$

For the computation of the DFT, we use the FFT, so the computation of $\mathtt{A}$ needs $\mathcal{O}\left(N^2 \log N\right)$ arithmetic operations.

The next step is the computation of the Legendre transform

$$\widehat{\mathcal{T}f}(n,k) = \frac{2\pi}{L}\sqrt{\frac{2n+1}{4\pi}} \sum_{m=1}^{M} P_n(t_m) \mathtt{A}(k,m).$$

Using the fast Legendre transform algorithms from [43] or [21], the computation of $\widehat{\mathcal{T}f}$ is done in $\mathcal{O}\left(N^2 \left(\log N\right)^2\right)$ or $\mathcal{O}\left(N^2 \left(\log N\right)^2 / \log\log N\right)$ operations, respectively.

In the next step, we multiply the Fourier coefficients $\widehat{\mathcal{T}f}$ with the inverse singular values $\lambda\left(n,k\right)^{-1}$ and the regularization coefficients $\hat{\psi}(n)$. Before we can do this, we have to project $\mathcal{T}f$ onto the range of $\mathcal{T}$ by defining $\widehat{\mathcal{T}f}(n,k) = 0$ for $n+k$ odd or $|k| > n$. Now we set

$$\hat{f}_\psi(n,k) = \lambda\left(n,k\right)^{-1} \hat{\psi}(n) \widehat{\mathcal{T}f}(n,k), \quad n = 0, \ldots, N,\ k = -n, \ldots, n.$$

From the definition,

$$f_\psi = \psi \star \mathcal{T}^\dagger \mathcal{L}_N \mathcal{T} f = \sum_{n=0}^{N} \sum_{k=-n}^{n} \hat{f}_\psi(n,k) Y_{n,k}$$

In the last step, we compute the function values of f_ψ. The computation of $\mathcal{O}(N^2)$ values of f_ψ can be done using the NFSFT with $\mathcal{O}\left(N^2 \left(\log N\right)^2\right)$ operations, see [34]. Summing up these considerations yields to the following algorithm:

Algorithm 4.21. *Input: Function values* $g(\sigma_l, t_m)$ *of the function* $g = \mathcal{T}f \in \Pi_N\,(\mathbb{T} \times [-1,1])$ *at nodes* $\sigma_l = 2\pi l/L$, $l = 1, \ldots, L = 2N+1$ *and* t_m *with the corresponding weights* ω_m, $m = 1, \ldots, M$.

1. *Compute for each* $m = 1, \ldots, M$ *the fast Fourier transform*

$$A(k-N, m) = \left[\mathrm{DFT}\left(g(\sigma_{(\cdot)}, t_m)\right) \mathrm{e}^{2\pi \mathrm{i}(\cdot)N/L}\right](k), \quad k = 0, \ldots, 2N.$$

2. *Compute the fast Legendre transform*

$$\widehat{\mathcal{T}f}(n,k) = \frac{2\pi}{L}\sqrt{\frac{2n+1}{4\pi}} \sum_{m=1}^{M} \omega_m P_n(t_m) A(k,m), \quad n = 0, \ldots, N,\ k = -N, \ldots, N.$$

3. *Multiply the Fourier coefficients with the singular values, compute*

$$\hat{f}_\psi(n,k) = \begin{cases} \lambda\,(n,k)^{-1}\,\hat{\psi}(n)\widehat{\mathcal{T}f}(n,k), & n = 0, \ldots, N,\ k = -n, \ldots, n,\ 2|\,(n+k) \\ 0, & \textit{otherwise} \end{cases}$$

4. *Compute the spherical Fourier transform*

$$f_\psi(\boldsymbol{x}_j) = \sum_{n=0}^{N} \sum_{k=-n}^{n} \hat{f}_\psi(n,k) Y_{n,k}(\boldsymbol{x}_j), \quad j = 1, \ldots, J.$$

Output: Function values $f_\psi\,(\boldsymbol{x}_j)$ *of the function* $\psi \star \mathcal{T}^\dagger \mathcal{L}_N \mathcal{T} f$ *at nodes* $\boldsymbol{x}_j$, $j = 1, \ldots, J$.

5 Conclusion and Outlook

Results

Here we are at the very last chapter of this thesis. We have examined the reconstruction of a function given data sampled from its circular average transform on the sphere. We have chosen the mollifier approach to construct an estimator and found out that it suffices to choose the mollifier from the class $\psi_{\tilde{N}}$ (see (4.37)) for achieving the optimal minimax rate for the error. This means that, for practical applications, we need to consider only the real number $\tilde{N}$ as a regularization parameter. The existence of such minimax estimators like $\psi_{\tilde{N}}$ has been proved before for a more general setting, which differs from ours in the fact the data function is given in terms of its Fourier coefficients, cf. [5].

To obtain the optimality of $\psi_{\tilde{N}}$, we have imposed some conditions on the quadrature formulas that we used, and we have proved that these conditions are fulfilled for the Fejér quadrature. Moreover, we have generalized existing estimates of the hyperinterpolation error for a different domain, namely the side of a cylinder.

Outlook

Since we have proposed an algorithm to compute the circular average transform, computational tests should be done to confirm our theoretical results numerically. This is subject to further research.

Another interesting point is the concept of applicability of quadrature formulas, which was introduced in Definition 4.7. So far, we have only proved the applicability of the constant-weight and the Fejér quadrature. It is possible that this result could be generalized to other quadratures like the Gauss-Legendre rule.

In this thesis, we have shown the optimality of the class $\psi_{\tilde{N}}$ only for the circular average transform. Nevertheless, our setting of sampled noisy data can also be used for other integral transforms on the sphere. It would be interesting to see if (or under which conditions) this class is optimal for other transforms, too.

Bibliography

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 1964.

[2] Lars-Erik Andersson. On the determination of a function from spherical averages. *SIAM Journal on Mathematical Analysis*, 19(1):214–232, January 1988.

[3] Frank Bauer and Mark A. Lukas. Comparing parameter choice methods for regularization of ill-posed problems. *Mathematics and Computers in Simulation*, 81(9):1795–1841, May 2011.

[4] Hubert Berens, Paul Leo Butzer, and Siegfried Pawelke. Limitierungsverfahren von Reihen mehrdimensionaler Kugelfunktionen und deren Saturationsverhalten. *Publications of the Research Institute for Mathematical Sciences*, 4(2):201–268, March 1968.

[5] Laurent Cavalier. Nonparametric statistical inverse problems. *Inverse Problems*, 24(3):034004, June 2008.

[6] James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 19(90):297, April 1965.

[7] Feng Dai and Yuan Xu. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer, New York, NY, 2013.

[8] Susanna Dann. On the Minkowski-Funk transform. March 2010.

[9] Philip J. Davis and Philip Rabinowitz. Some geometrical theorems for abscissas and weights of Gauss type. *Journal of Mathematical Analysis and Applications*, 2(3):428–437, June 1961.

[10] Etienne De Montigny. Photoacoustic tomography : Principles and applications. September 2011.

[11] Mehdi Dehghan, M. Masjed-Jamei, and M. R. Eslahchi. On numerical improvement of the second kind of Gauss–Chebyshev quadrature rules. *Applied Mathematics and Computation*, 168(1):431–446, September 2005.

[12] Paul Erdős and Paul Turán. On interpolation. III. Interpolatory theory of polynomials. *Annals of Mathematics*, 41(3):510–553, July 1940.

[13] M. R. Eslahchi, M. Dehghan, and M. Masjed-Jamei. On numerical improvement of the first kind Gauss–Chebyshev quadrature rules. *Applied Mathematics and Computation*, 165(1):5–21, June 2005.

[14] Klaus-Jürgen Förster and Knut Petras. On estimates for the weights in Gaussian quadrature in the ultraspherical case. *Mathematics of Computation*, 55(191):243–243, September 1990.

[15] Willi Freeden, Theo Gervens, and Michael Schreiner. *Constructive Approximation on the Sphere with Applications to Geomathematics.* Numerical Mathematics and Scientific Computation. Clarendon Press; Oxford University Press, Oxford; New York, 1998.

[16] Paul Funk. Über Flächen mit lauter geschlossenen geodätischen Linien. *Mathematische Annalen*, 74(2):278–300, June 1913.

[17] Richard J. Gardner. *Geometric tomography.* Number 58 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge; New York, second edition, 2006.

[18] Walter Gautschi. Numerical quadrature in the presence of a singularity. *SIAM Journal on Numerical Analysis*, 4(3):357–362, September 1967.

[19] Walter Gautschi. Advances in Chebyshev quadrature. In G. Alistair Watson, editor, *Numerical Analysis*, volume 506 of *Lecture Notes in Mathematics.* Springer Berlin Heidelberg, 1976.

[20] Izrail Solomonovich Gradshteyn, Iosif Moiseevich Ryzhik, Alan Jeffrey, Daniel Zwillinger, and Scripta Technica. *Table of Integrals, Series, and Products*, volume 6. Academic Press New York, 1965.

[21] Nicholas Hale and Alex Townsend. A fast, simple, and stable Chebyshev–Legendre transform using an asymptotic formula. *SIAM Journal on Scientific Computing*, 36(1):A148–A167, February 2014.

[22] Markus Haltmeier. A mollification approach for inverting the spherical mean Radon transform. *SIAM Journal on Applied Mathematics*, 71(5):1637–1652, September 2011.

[23] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4):561–580, December 1992.

[24] Sigurdur Helgason. *The Radon Transform.* Number 5 in Progress in Mathematics. Birkhäuser, Boston, second edition, 1999.

[25] Kerstin Hesse and Ian Hugh Sloan. Hyperinterpolation on the sphere. *Frontiers in Interpolation and Approximation*, page 213, May 2006.

[26] Edwin Hewitt and Robert E. Hewitt. The Gibbs-Wilbraham phenomenon: An episode in Fourier analysis. *Archive for History of Exact Sciences*, 21(2):129–160, 1979.

[27] Charles Austin Hobbs. *Solid Geometry.* Camebridge, Massachusetts, 1921.

[28] Bernd Hofmann. *Mathematik inverser Probleme.* Teubner, Stuttgart, 1999.

[29] Sergey I. Kabanikhin. *Inverse and Ill-posed Problems: Theory and Applications.* De Gruyter, Berlin, December 2011.

[30] Jens Keiner, Stefan Kunis, and Daniel Potts. NFFT, C subroutine library, 2013.

[31] Jens Keiner and Daniel Potts. Fast evaluation of quadrature formulae on the sphere. *Mathematics of Computation*, 77(261):397–419, January 2008.

[32] Arno B. J. Kuijlaars. Chebyshev-type quadrature for Jacobi weight functions. *Journal of Computational and Applied Mathematics*, 57(1-2):171–180, February 1995.

[33] Arno B. J. Kuijlaars and W. Van Assche. The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients. *Journal of Approximation Theory*, 99(1):167–197, July 1999.

[34] Stefan Kunis and Daniel Potts. Fast spherical Fourier algorithms. *Journal of Computational and Applied Mathematics*, 161(1):75–98, December 2003.

[35] Stefan Kunis and Daniel Potts. Time and memory requirements of the nonequispaced FFT. *Sampling Theory in Signal and Image Processing*, (7):77 – 100, January 2008.

[36] Leonid A. Kunyansky. Explicit inversion formulae for the spherical mean Radon transform. *Inverse Problems*, 23(1):373, January 2007.

[37] Martin Kütz. On the asymptotic behaviour of some Cotes numbers. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 66(8):373–375, 1986.

[38] Alfed Karl Louis and Peter Maass. A mollifier method for linear operator equations of the first kind. *Inverse Problems*, 6(3):427–440, June 1990.

[39] August Ferdinand Möbius. Ueber die Bestimmung des Inhaltes eines Polyëders. In Felix Klein, editor, *Gesammelte Werke*, volume 2. Hirzel, Leipzig, 1886.

[40] Claus Müller. *Spherical Harmonics.* Springer-Verlag, Berlin; New York, 1966.

[41] Stefan Nilsson and Lars-Erik Andersson. Application of fast back-projection techniques for some inverse problems of synthetic aperture radar. In Edmund G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery V*, volume 3370 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 62–72. International Society for Optics and Photonics, September 1998.

[42] Viktor Pavlovich Palamodov. *Reconstructive Integral Geometry*, volume 98 of *Monographs in Mathematics Series.* Birkhäuser, Basel, 2004.

[43] Daniel Potts, Gabriele Steidl, and Manfred Tasche. Fast algorithms for discrete polynomial transforms. *Mathematics of Computation*, 67(224):1577–1591, October 1998.

[44] Johann Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig*, April 1917.

[45] Johann Radon. On the determination of functions from their integral values along certain manifolds. *IEEE Transactions on Medical Imaging*, 5(4):170–176, December 1986.

[46] Nicholas J. Redding and Garry N. Newsam. Inverting the circular Radon transform. Research Report DSTO-RR-0211, Defence Science and Technology Organization, Edinburgh, South Australia, Australia, August 2001.

[47] Herbert Robbins. A remark on Stirling's formula. *American Mathematical Monthly*, 62(1):26–29, January 1955.

[48] Walter Rudin. Uniqueness theory for Laplace series. *Transactions of the American Mathematical Society*, 68(2):287–303, March 1950.

[49] Lawrence Alan Shepp and Joseph Bernard Kruskal. Computerized tomography: The new medical X-ray technology. *The American Mathematical Monthly*, 85(6):420–439, June 1978.

[50] Ian Hugh Sloan. Polynomial interpolation and hyperinterpolation over general regions. *Journal of Approximation Theory*, 83(2):238–254, November 1995.

[51] Gábor Szegő. *Orthogonal Polynomials*. Number v. 23 in Colloquium publications - American Mathematical Society. American Mathematical Society, Providence, fourth edition, 1939.

[52] Lloyd N. Trefethen. *Approximation Theory and Approximation Practice*. Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 2013.

[53] Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, New York; London, 2009.

[54] Matti Vuorinen. *Conformal Geometry and Quasiregular Mappings*, volume 1319 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1988.

[55] Jörg Waldvogel. Fast construction of the Fejér and Clenshaw–Curtis quadrature rules. *BIT Numerical Mathematics*, 46(1):195–202, March 2006.

[56] Lihong V. Wang and Song Hu. Photoacoustic tomography: In vivo imaging from organelles to organs. *Science*, 335(6075):1458–1462, March 2012.

[57] Dirk Werner. *Funktionalanalysis*. Springer, Berlin; Heidelberg, seventh edition, 2011.

[58] Can Evren Yarman and Birsen Yazici. Inversion of the circular averages transform using the Funk transform. *Inverse Problems*, 27(6):065001, June 2011.

[59] Gerhard Zangerl and Otmar Scherzer. Exact reconstruction in photoacoustic tomography with circular integrating detectors II: Spherical geometry. *Mathematical Methods in the Applied Sciences*, 33(15):1771–1782, August 2010.

Selbstständigkeitserklärung (Declaration of Authorship)

Ich erkläre gegenüber der Technischen Universität Chemnitz, dass ich die vorliegende Diplomarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe.

Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch bei keinem anderen Prüfer als Prüfungsleistung eingereicht und ist auch noch nicht veröffentlicht. [1]

Michael Quellmalz
Chemnitz, September 2014

[1] Der Text dieser Erklärung ist identisch mit dem der Vorlage für die Selbstständigkeitserklärung des Zentralen Prüfungsamtes der Technischen Universität Chemnitz.