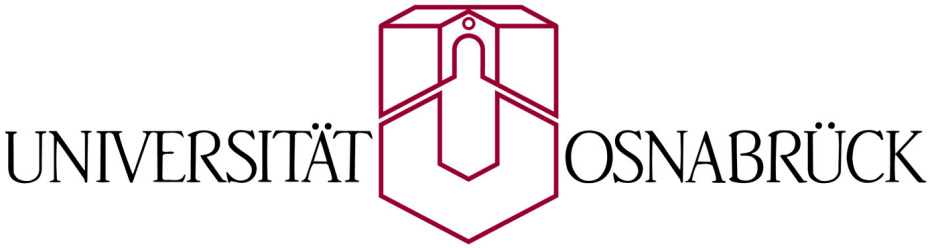


Spherical mean values

Efficient computation by Fourier techniques and regularized reconstructions of function samples from discrete means

zur Erlangung des Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
des Fachbereiches Mathematik/Informatik
der Universität Osnabrück

vorgelegt von
Torsten Görner

Betreuer
Prof. Dr. Stefan Kunis

Osnabrück, Mai 2015

Contents

1 List of publications **1**

2 Introduction **3**
2.1 Outline of the thesis . . . 5
2.2 Acknowledgments . . . 6

3 Basics **7**
3.1 Notations . . . 7
3.2 Spherical coordinates . . . 8
3.3 Fourier analysis . . . 11
3.4 Mean value operators . . . 15

4 Efficient discretization **19**
4.1 Numerical integration . . . 19
4.2 Fourier based approach . . . 24
4.2.1 Regular Cartesian frequency grids . . . 24
4.2.2 Polar frequency grids . . . 34

5 Inverse problem **43**
5.1 Total variation based regularization . . . 44
5.1.1 Primal and dual problems . . . 44
5.1.2 Numerical solution . . . 50
5.2 Shearlet based regularization . . . 52
5.2.1 Directional representation systems . . . 54
5.2.2 Implementations of Shearlet transforms . . . 54
5.2.3 Numerical solution . . . 57
5.2.4 Shearlets on polar frequency grids . . . 58

6 Implementation **65**
6.1 Object-oriented programming in Matlab . . . 67
6.2 Utility and geometry class . . . 68
6.3 Test functions . . . 68
6.4 Forward operators . . . 70
6.5 Reconstructions . . . 70
6.6 Auxiliary classes . . . 70

7 Numerical experiments **71**
7.1 Error analysis 71
7.2 Running times 75
7.3 Reconstructions 76
7.3.1 Surrounding center points 77
7.3.2 Incomplete data 79

8 Conclusion **85**

Bibliography **87**

List of Figures **91**

List of Tables **93**

List of Algorithms **95**

1

List of publications

Parts of this thesis have already been published in the following references. The respective numbers are the same as in the full bibliography at the end of this thesis.

[13] Y. Dong, T. Görner, and S. Kunis. An algorithm for total variation regularized photoacoustic imaging. *Adv. Comput. Math.*, 41(2):423–438, 2015.

[25] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. *Inverse Probl. Imag.*, 6(4):645–661, 2012.

[26] T. Görner and S. Kunis. Effective discretization of the two-dimensional wave equation. *PAMM*, 14(1):947–948, 2014.

[27] T. Görner and S. Kunis. SMV, Matlab toolbox for computing spherical mean values. http://torstengoerner.de/software/. May 2015.

2

Introduction

Different imaging modalities are of great interest in various fields of current research. In the meantime, numerous different techniques were introduced and further improved. While some of them are still under development, many methods are now widely used in a variety of applications. Biomedical imaging represents one meaningful area. Topics like computed tomography (CT), magnetic resonance imaging (MRI), ultrasonography, and thermography play an important role in medical diagnoses and therapies.
A well-known hybrid biomedical imaging modality is photoacoustic imaging (PAI), which is based on the photoacoustic effect. Here, delivered electromagnetic energy, mostly a laser pulse, is absorbed by the tissue and converted into heat and this leads to the formation of ultrasonic pressure waves.
Typically, the concerning object is located inside of a domain Ω and surrounded by several ultrasonic detectors, which are arranged on the boundary $\partial\Omega$ of the domain Ω. Due to limitations by certain applications, the detectors might cover the boundary of the domain only partially. After the stimulation of the tissue by a laser pulse, the arising ultrasonic signals are measured by the detectors at different points in time.
Under some assumptions, this process can mathematically be modeled by a partial differential equation, the so-called Cauchy problem of the wave equation. Especially for a constant speed of sound inside the domain, the measured data can be expressed by spherical mean values

$$\mathcal{M}f(\mathbf{y}, r) = \int_{\mathbb{S}^{d-1}} f\left(\mathbf{y} + r\boldsymbol{\xi}\right) \mathrm{d}\sigma(\boldsymbol{\xi})$$

of the object f. We interpret $\mathbf{y} \in \partial\Omega$ as the positions of the ultrasonic detectors and $r > 0$ as the times of the measurements. This is analog to the concept of x-ray computed tomography, where the measured data can be modeled by the classical Radon transform, which assigns functions their averages over lines. Hence, the evaluation of the above operator $\mathcal{M}$ is sometimes called spherical Radon transform.
Relevant applications ask for a nondestructive examination of internal structures of in vivo entities, for example parts of the body or inner organs. With the previous mathematical modeling of PAI, this conceptual formulation can be translated into the inversion of the spherical mean value operator $\mathcal{M}$, in other words, the recovery of functions f from their mean values $\mathcal{M}f$.

In general, the inversion of the mean value operator $\mathcal{M}$ is a non well-defined task. For arbitrary mean values $\mathcal{M}f$, the existence of an underlying function f is actually not guaranteed. Even in the case of the existence, a unique solution is not ensured. It turns out, that this issue depends essentially on the properties of the boundary $\partial\Omega$ of the domain Ω. For a detailed treatment of this fundamental topic for the two-dimensional case and for specific three-dimensional geometries, we refer to [18, 5, 4, 44, 57] and the references therein.

Different concepts for the inversion of the spherical mean value operator $\mathcal{M}$ have been widely studied. For specific geometries of the domain Ω, various direct reconstruction algorithms are discussed for example in [30, 47, 48, 29, 3, 16, 46, 31, 17]. For general geometries, explicit inversion formulas are not known and do not represent a realistic goal, see also [55, 28, 32].

We focus on iterative reconstruction methods. Typically, we start with a rough approximation as initial guess. Each subsequent iteration of the algorithm produces a reconstruction with an increased quality. This requires the evaluation of the spherical mean value operator $\mathcal{M}$ multiple times per iteration step. Hence, the availability of fast algorithms for the so-called forward problem, the computation of spherical means, is crucial for the development of efficient reconstruction methods.

The implementation and application of the algorithms require to work with finite data. In particular, we have to approximate continuous functions by discrete data sets. This thesis is based on the application of tools from Fourier analysis, a well-known and extremely successful field in science and technology, which enables us to develop fast and accurate discretizations of the considered problems. In particular, we use the fact [42, eq. (1.5)], that the complex exponential function

$$\mathrm{e}_{\boldsymbol{\xi}} \colon \mathbb{R}^d \to \mathbb{C}, \quad \mathrm{e}_{\boldsymbol{\xi}}(\mathbf{x}) = \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\xi}\cdot\mathbf{x}}, \quad \boldsymbol{\xi} \in \mathbb{R}^d,$$

is an eigenfunction of the spherical mean value operator $\mathcal{M}$ for a fixed time r, more precise, the relation

$$\mathcal{M}\mathrm{e}_{\boldsymbol{\xi}}(\mathbf{y}, r) = \frac{\Gamma\left(\frac{d}{2}\right) \mathcal{J}_{\frac{d}{2}-1}\left(2\pi r|\boldsymbol{\xi}|\right)}{\left(\pi r|\boldsymbol{\xi}|\right)^{\frac{d}{2}-1}} \cdot \mathrm{e}_{\boldsymbol{\xi}}(\mathbf{y}) \tag{2.1}$$

is fulfilled. Especially the spherical mean values of trigonometric polynomials can be expressed by nonequispaced Fourier transforms, whose efficient evaluation is given by well-developed algorithms [43, 45].

Since the spherical mean value operator $\mathcal{M}$ is an integral operator, it is also a smoothing operator, which implies, that its inversion is an ill-posed problem. With noisy or incomplete data in addition, this issue causes inappropriate results in most cases. Therefore, the utilization of regularization techniques is a necessary element in reconstruction algorithms. The basic idea is to force the solutions to fulfill some a-priori characteristics. In medical imaging, piecewise smoothness is a suitable assumption on the objects. However, even if the smooth parts can be well reconstructed, the accurate handling of edges and other structures remains a challenging task. To overcome these difficulties, we make use of two different regularizing techniques. In one variant, we force a small total variation of the reconstruction, and in another approach, we prefer results with sparse Shearlet coefficients [52, 24].

2.1 Outline of the thesis

This thesis is organized as follows. To make it self-contained, we depict in Chapter 3 some related basic knowledge. We start with the used notations and list some fundamental concepts for dealing with spherical integrals. Afterwards, we introduce basic tools from Fourier analysis, including the approximation of functions by trigonometric polynomials and associated error estimates. The chapter finishes with the definition of the considered mean value operators and the proof of the fundamental statement in equation (2.1), the application of the mean value operators to the complex exponential function.

Chapter 4 treats different discretizations of the forward problem and thus the efficient computation of mean values. As a first variant, we consider the numerical evaluation of integrals by quadrature rules. While this approach is used for comparison, the thesis focuses on a Fourier based approach, obtained fundamentally from equation (2.1). Here we consider different frequency grids and several algorithms for the evaluation of the appearing Fourier sums. In addition, we analyze also the computational complexities of these concepts.

For cartesian frequency grids, a fast computation of mean values is proven by Theorems 26 and 27. The implementations are realized by using d-dimensional nonequispaced fast Fourier transforms (NFFT, [43]) and are given by Algorithms 4 and 5.

Based on Theorem 28, we formulate with Algorithm 6 an efficient computation of mean values for large problem sizes. In particular the computation of mean values for $\mathcal{O}\left(N^{d-1}\right)$ center points, located on a smooth $(d-1)$-dimensional submanifold, and $\mathcal{O}(N)$ radii can be realized with a $(d+1)$-dimensional butterfly sparse Fourier transform [45] and ends up in an almost linear runtime complexity.

To obtain better rotational invariance, it is more natural to use polar frequency grids. Theorem 34 and Algorithm 7 provide the efficient computation of mean values for this setting. We complete this chapter by proving error bounds and giving stability arguments.

The inverse problem, the reconstruction of function samples from given mean values, is the topic of Chapter 5. We present two approaches, which are suitable for piecewise smooth functions. Firstly, we consider a regularization technique, where we force the reconstructions to have a small total variation. We solve the appearing system of equations with the semismooth Newton method and summarize the implementation in Algorithm 8. Another promising approach is to favor reconstructions with sparse Shearlet coefficients [52, 24]. We derive a realization with the alternating direction method of multipliers (ADMM, [7]) and summarize the whole procedure in Algorithm 9. This chapter concludes with the construction of a Shearlet system on polar frequencies, where we follow the approach from [36, 35].

All derived algorithms are implemented and published as a Matlab toolbox [27]. To make the large amount of code maintainable and extensible for several developers, we use an object-oriented design, illustrated in Figure 6.1. Additionally, some thematically related work is also included in this toolbox [17]. Chapter 6 presents the concept and the handling of the software. We start with a brief introduction to object-oriented programming in Matlab and afterwards, we explain the structure of the code and comment on the features of all the different classes.

Finally, in Chapter 7, we analyze the results of numerous numerical experiments. We

consider the behavior of the running times as well as of the approximation error. The chapter finishes with a combination of all essential parts of this thesis by demonstrating the reconstruction capabilities of our algorithms in Figures 7.6 and 7.8.

2.2 Acknowledgments

First of all, I want to express my great gratitude to my parents and my whole family. They supported me extremely in all situations of my life and made me to the person I am today. This offered me significantly the basis to realize this project.
I thank my doctoral advisor Prof. Stefan Kunis for giving me the opportunity to work on this thesis and for his great guidance throughout the years. I thank also Prof. Daniel Potts for everything I have learned from him during my studies in Chemnitz and for the helpful suggestions in the last stage of this thesis.
In addition, I thank all colleagues for the valuable assistance, suggestions, and discussions, that improved my work substantially. In particular, among all the numerous supporting people, I want to name Christina Brandt, Yiqiu Dong, Sören Häuser, Ralf Hielscher, Ines Melzer, Thomas Peter, and Ruben Seyfried.
Finally, I highly appreciate the magnificent support in all other things in my life besides the work. Doing a lot of sports and music with all my amazing friends gave me the necessary motivation, energy, and recovery to finish this thesis.

3

Basics

This chapter is a collection of some basic knowledge for this thesis. We start with the used notations, list some fundamental concepts for dealing with spherical integrals. Afterwards, we summarize some basic tools from Fourier analysis and finish this chapter with the definition of the mean value operators and the formulation of some useful properties.

3.1 Notations

We denote in the usual manner by $\mathbb{N} := \{1, 2, \dots\}$ the natural numbers, by $\mathbb{Z}$ the integers, by $\mathbb{R}$ the real numbers, and by $\mathbb{C}$ the complex numbers. We use abbreviatory $2\mathbb{N} := \{2, 4, \dots\}$ for the even natural numbers and $\mathbb{I}^d := \left[-\frac{1}{2}, \frac{1}{2}\right]^d$, $d \in \mathbb{N}$, for the d-dimensional unit cube.
Matrices and vectors are labeled with bold symbols, in particular, we use

$$\mathbf{1} := \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{I} := \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

for a matrix with ones at every position and the identity matrix, where the dimensions are apparent from the context. We define the inner product of vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $n \in \mathbb{N}$, by $\mathbf{a} \cdot \mathbf{b} := a_1 b_1 + \dots + a_n b_n$ and the according Euclidean norm by $\|\mathbf{a}\|_2 := |\mathbf{a}| := \sqrt{\mathbf{a} \cdot \mathbf{a}}$. Moreover, the l^p-norm $\|\cdot\|_p \colon \mathbb{C}^n \to [0, \infty)$ of a vector $\mathbf{a} \in \mathbb{C}^n$, $p \in [1, \infty]$, $n \in \mathbb{N}$, is defined by

$$\|\mathbf{a}\|_p := \left(\sum_{k=1}^{n} |a_k|^p \right)^{\frac{1}{p}} \quad \text{for } p \in [1, \infty) \quad \text{and} \quad \|\mathbf{a}\|_\infty := \max_{k \in \{1, \dots, n\}} |a_k|.$$

The summation of a matrix $\mathbf{A}$ with a scalar $\lambda \in \mathbb{R}$ is given as $\mathbf{A} + \lambda := \mathbf{A} + \lambda \mathbf{1}$, where $\mathbf{1}$ has the same dimensions as $\mathbf{A}$. For matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$, $m, n \in \mathbb{N}$, we denote by

$$\mathbf{A} \odot \mathbf{B} := \begin{pmatrix} A_{1,1} B_{1,1} & \dots & A_{1,n} B_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} B_{m,1} & \dots & A_{m,n} B_{m,n} \end{pmatrix} \in \mathbb{C}^{m \times n}$$

the element-wise product and for matrices $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{B} \in \mathbb{C}^{p\times r}$, $m, n, p, r \in \mathbb{N}$, we denote by

$$\mathbf{A} \otimes \mathbf{B} := \begin{pmatrix} A_{1,1}\mathbf{B} & \dots & A_{1,n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ A_{m,1}\mathbf{B} & \dots & A_{m,n}\mathbf{B} \end{pmatrix} \in \mathbb{C}^{mp\times nr}$$

the Kronecker product. For a spatial dimension $d \in \mathbb{N}$ and a discretization parameter $N \in 2\mathbb{N}$, the following sets are frequently used in this thesis,

$$J_N := \left\{ \mathbf{z} = (z_1, \dots, z_d)^\top \in \mathbb{Z}^d : z_1, \dots, z_d \in \left\{ -\frac{N}{2}, \dots, \frac{N}{2} - 1 \right\} \right\}, \tag{3.1}$$

$$Z_N := \left\{ \mathbf{z} = (z_1, \dots, z_d)^\top \in \mathbb{Z}^d : z_1, \dots, z_d \in \left\{ -\frac{N}{2}, \dots, \frac{N}{2} \right\} \right\}, \tag{3.2}$$

and

$$X_N := \left\{ \mathbf{x} = \frac{\mathbf{z}}{N} + \frac{1}{2N} : \mathbf{z} \in J_N \right\} \subset \left(-\frac{1}{2}, \frac{1}{2} \right)^d. \tag{3.3}$$

Let a set $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be given. We use the notation $C(\Omega)$ for the set of all continuous functions $f\colon \Omega \to \mathbb{R}$, and $C(\Omega, \mathbb{C})$ for the set of all continuous functions $f\colon \Omega \to \mathbb{C}$. For a measure space $(\Omega, \mathcal{A}, \mu)$ with a nonempty set $\Omega \subset \mathbb{R}^d$, a σ-algebra $\mathcal{A}$ over Ω and a measure $\mu\colon \mathcal{A} \to [0, \infty]$, the function space $L^p(\Omega, \mu)$, $p \in [1, \infty]$, contains all measurable functions $f\colon \Omega \to \mathbb{R}$ with a finite norm $\|\cdot\|_{L^p(\Omega,\mu)}$,

$$\|f\|_{L^p(\Omega,\mu)} := \left(\int_\Omega |f(\mathbf{x})|^p \mathrm{d}\mu(\mathbf{x}) \right)^{\frac{1}{p}} \quad \text{for } p \in [1, \infty) \quad \text{and} \quad \|f\|_{L^\infty(\Omega)} := \operatorname*{ess\,sup}_{\mathbf{x}\in\Omega} |f(\mathbf{x})|.$$

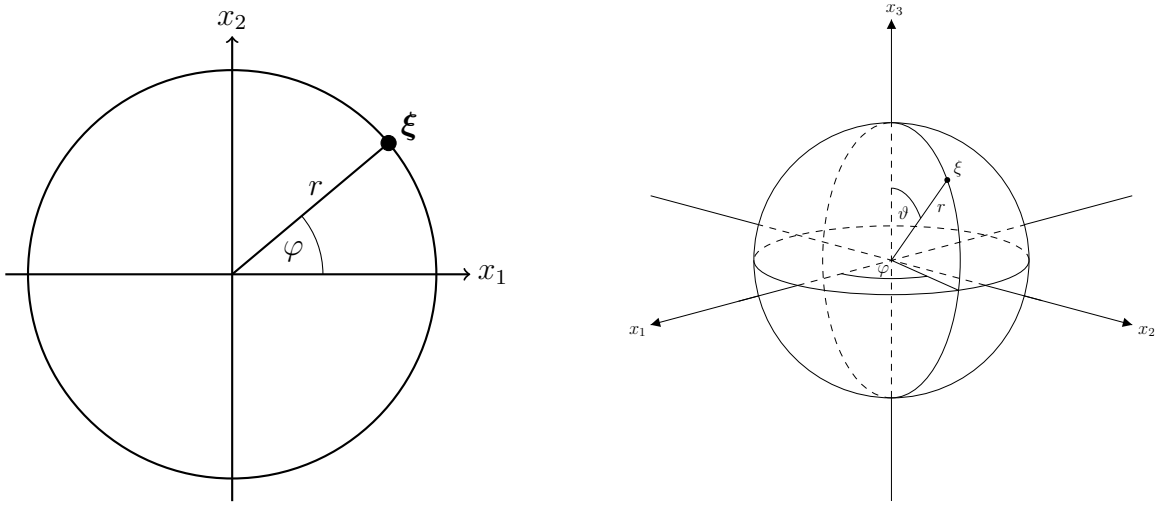
Unless otherwise stated, μ is the Lebesgue measure $\lambda\colon \mathcal{A} \to [0, \infty]$, which is the unique measure on $\mathbb{R}^d$, that assigns d-dimensional cuboids their volumes,

$$\lambda([a_1, b_1] \times [a_d, b_d]) := (b_1 - a_1) \cdots (b_d - a_d).$$

Abbreviatory, we set $L^p(\Omega) := L^p(\Omega, \lambda)$. For sets Ω_1, Ω_2 and Ω_3, we denote the set of all linear functions $f\colon \Omega_1 \to \Omega_2$ by $\mathcal{L}(\Omega_1, \Omega_2)$ and the product $fg\colon \Omega_3 \to \Omega_2$ of functions $f\colon \Omega_1 \to \Omega_2$ and $g\colon \Omega_3 \to \Omega_1$ is defined by the composition $fg(x) := f(g(x))$.

3.2 Spherical coordinates

The elements of $\mathbb{R}^d$ can be represented by different coordinate systems. The well-known Cartesian coordinates are in some cases unqualified for dealing with spherical geometries and hence, polar and spherical coordinates in Figure 3.1 are typically used for two and three dimensions. For higher dimensions, we use a generalization of this concept [33, Sec. 21.2.1].



(a) Polar coordinates. (b) Spherical coordinates.

Figure 3.1: Spherical coordinate systems for spatial dimensions $d = 2$ (left) and $d = 3$ (right).

Definition 1. We define the operator $\mathcal{P}_d$, which maps generalized d-dimensional spherical coordinates to Cartesian coordinates, $d \in \mathbb{N} \setminus \{1\}$, recursively by

$$\mathcal{P}_2 \colon [0,\infty) \times [0,2\pi) \to \mathbb{R}^2, \quad \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix},$$

for $d = 2$ and

$$\mathcal{P}_d \colon [0,\infty) \times [0,2\pi) \times [0,\pi]^{d-2} \to \mathbb{R}^d,$$

$$(r,\varphi,\boldsymbol{\vartheta}) := (r,\varphi,\vartheta_1,\ldots,\vartheta_{d-2})^\top \mapsto \begin{pmatrix} \sin\vartheta_{d-2}\mathcal{P}_{d-1}\left(r,\varphi,\vartheta_1,\ldots,\vartheta_{d-3}\right) \\ r\cos\vartheta_{d-2} \end{pmatrix},$$

for $d \in \mathbb{N}$, $d \geq 3$. Since the restriction of the operator $\mathcal{P}_d$ to $(0,\infty) \times [0,2\pi) \times (0,\pi)^{d-2}$ is bijective, the inverse operator

$$\mathcal{P}^{-1} \colon \mathbb{R}^d \setminus \left(\{\mathbf{0}\} \times \mathbb{R}^{d-2}\right) \to (0,\infty) \times [0,2\pi) \times (0,\pi)^{d-2}$$

is well-defined. □

We proceed with some basic transformations for spherical integrals and start with the substitution between Cartesian and spherical coordinates.

Lemma 2. *Let a spatial dimension $d \in \mathbb{N} \setminus \{1\}$, an open set*

$$\Omega \subset (0,\infty) \times [0,2\pi) \times (0,\pi)^{d-2},$$

and an integrable function $f \in C\left(\mathcal{P}_d\Omega\right)$ be given. With

$$\mathcal{D}_d \colon (0,\infty) \times (0,\pi)^{d-2} \to \mathbb{R}, \quad \mathcal{D}_d(r,\boldsymbol{\vartheta}) := r^{d-1}\sin\vartheta_1\left(\sin\vartheta_2\right)^2 \ldots \left(\sin\vartheta_{d-2}\right)^{d-2}, \tag{3.4}$$

it follows

$$\int_{\mathcal{P}_d\Omega} f(\mathbf{x})\mathrm{d}\mathbf{x} = \int_\Omega f\left(\mathcal{P}_d(r,\varphi,\boldsymbol{\vartheta})\right)\mathcal{D}_d(r,\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta}\mathrm{d}\varphi\mathrm{d}r.$$

Proof. With the notation $\mathbf{w} := (r, \varphi, \boldsymbol{\vartheta})$ and the Jacobian matrix

$$\mathbf{J}_{\mathcal{P}_d} := \left(\frac{\partial \left(\mathcal{P}_d\right)_j}{\partial w_k} \right)_{j,k=1,\ldots,d} (\mathbf{w}),$$

complete induction shows

$$\det \mathbf{J}_{\mathcal{P}_d}(r, \boldsymbol{\vartheta}) = (-1)^d \mathcal{D}_d(r, \boldsymbol{\vartheta}).$$

The substitution formula for integrals finishes the proof. □

Next, we specify the integration over surfaces of spheres.

Definition 3. Let a spatial dimension $d \in \mathbb{N} \setminus \{1\}$ be given. We denote by

$$\mathbb{S}^{d-1} := \left\{ \mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| = 1 \right\} \subset \mathbb{R}^d$$

the $(d-1)$-dimensional unit sphere, use the notation

$$\mathbf{G}_d(w_2, \ldots, w_d) := \left(\frac{\partial \left(\mathcal{P}_d\right)_j}{\partial w_k} \right)_{j,k=1,\ldots,d} (1, w_2, \ldots, w_d),$$

and define the spherical surface measure $\mathrm{d}\sigma$ by

$$\int_{\mathbb{S}^{d-1}} f(\mathbf{x}) \mathrm{d}\sigma(\mathbf{x}) := \int_0^{2\pi} \int_{(0,\pi)^{d-2}} f\left(\mathcal{P}_d(1, \varphi, \boldsymbol{\vartheta})\right) \sqrt{\mathbf{G}_d^\top(\varphi, \boldsymbol{\vartheta}) \mathbf{G}_d(\varphi, \boldsymbol{\vartheta})} \mathrm{d}\boldsymbol{\vartheta} \mathrm{d}\varphi \quad (3.5)$$

and the surface ω_{d-1} of the unit sphere $\mathbb{S}^{d-1}$ by

$$\omega_{d-1} := \int_{\mathbb{S}^{d-1}} 1 \mathrm{d}\sigma(\mathbf{x}).$$

□

We obtain the following useful representations for spherical integrals.

Corollary 4. *Let a spatial dimension $d \in \mathbb{N} \setminus \{1\}$ be given. For an integrable function $f \in C\left(\mathbb{S}^{d-1}\right)$, it follows*

$$\int_{\mathbb{S}^{d-1}} f(\mathbf{x}) \mathrm{d}\sigma(\mathbf{x}) = \int_0^{2\pi} \int_{(0,\pi)^{d-2}} f\left(\mathcal{P}_d(1, \varphi, \boldsymbol{\vartheta})\right) \mathcal{D}_d(1, \boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta} \mathrm{d}\varphi,$$

with $\mathcal{P}_d$ from Definition 1 and $\mathcal{D}_d$ from equation (3.4). For an integrable function $f \in C(\mathbb{R}^d)$ and a radius $R > 0$, it follows

$$\int_{|\mathbf{x}| \leq R} f(\mathbf{x}) \mathrm{d}\mathbf{x} = \int_0^R \int_{\mathbb{S}^{d-1}} f(r\boldsymbol{\xi}) \mathrm{d}\sigma(\boldsymbol{\xi}) r^{d-1} \mathrm{d}r.$$

Proof. Complete induction shows

$$\sqrt{\mathbf{G}_d^\top(\varphi, \boldsymbol{\vartheta}) \mathbf{G}_d(\varphi, \boldsymbol{\vartheta})} = \mathcal{D}_d(1, \boldsymbol{\vartheta})$$

and we obtain the first assertion from the definition in equation (3.5). Furthermore, Lemma 2 and the definition of $\mathcal{D}_d$ in equation (3.4) leads to

$$\begin{aligned}\int_{|\mathbf{x}|\leq R} f(\mathbf{x})\mathrm{d}\mathbf{x} &= \int_0^R \int_0^{2\pi} \int_{(0,\pi)^{d-2}} f\left(\mathcal{P}_d(r\varphi,\boldsymbol{\vartheta})\right)\mathcal{D}_d(r\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta}\mathrm{d}\varphi\mathrm{d}r \\ &= \int_0^R \int_0^{2\pi} \int_{(0,\pi)^{d-2}} f\left(r\mathcal{P}_d(1,\varphi,\boldsymbol{\vartheta})\right)\mathcal{D}_d(1,\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta}\mathrm{d}\varphi r^{d-1}\mathrm{d}r,\end{aligned}$$

and applying of equation (3.5) proves the second assertion. □

With the previous statements, we can state an explicit formula for the surface of the unit sphere.

Corollary 5. *The surface of the unit sphere ω_{d-1} for a spatial dimension $d \in \mathbb{N}\setminus\{1\}$ is given by*

$$\omega_{d-1} = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}.$$

Proof. The proof is given by complete induction, where the inductive step follows from [1, eq. 6.2.1, 6.2.2, 6.1.18],

$$\begin{aligned}\omega_{d-1} &= \int_0^{2\pi}\int_{(0,\pi)^{d-2}} \mathcal{D}_d(1,\boldsymbol{\vartheta})\mathrm{d}\varphi\mathrm{d}\boldsymbol{\vartheta} = 2\omega_{d-2}\int_0^{\frac{\pi}{2}} \left(\sin\vartheta_{d-2}\right)^{d-2}\mathrm{d}\vartheta_{d-2} \\ &= \frac{2\pi^{\frac{d-1}{2}}}{\Gamma\left(\frac{d-1}{2}\right)}\cdot\frac{\Gamma\left(\frac{d-1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}.\end{aligned}$$

□

3.3 Fourier analysis

We define the d-dimensional torus $\mathbb{T}^d := \left(\mathbb{R}^d,+\right)/\mathbb{Z}^d$, $d\in\mathbb{N}$, as quotient group with the common periodical addition and identify its elements by the unique representatives in $\left[-\frac{1}{2},\frac{1}{2}\right)^d$. We denote with $C\left(\mathbb{T}^d\right)$ the space of the continuous and periodic functions $f\colon \mathbb{T}^d\to\mathbb{R}$ and can approximate these functions with their Fourier partial sum.

Definition 6. Let a spatial dimension $d\in\mathbb{N}$ and a function $f\in C\left(\mathbb{T}^d\right)$ be given. We define the Fourier coefficients $c_{\mathbf{z}}(f)\in\mathbb{C}$, $\mathbf{z}\in\mathbb{Z}^d$, of the function f by

$$c_{\mathbf{z}}(f) := \int_{\mathbb{T}^d} f(\mathbf{x})\mathrm{e}^{-2\pi\mathrm{i}\mathbf{z}\cdot\mathbf{x}}\mathrm{d}\mathbf{x}. \tag{3.6}$$

Furthermore, we denote by $\mathcal{S}_N f\colon \mathbb{T}^d\to\mathbb{C}$, $N\in 2\mathbb{N}$,

$$\mathcal{S}_N f(\mathbf{x}) := \sum_{\mathbf{z}\in J_N} c_{\mathbf{z}}(f)\mathrm{e}^{2\pi\mathrm{i}\mathbf{z}\cdot\mathbf{x}},$$

the N-th Fourier partial sum. In the case of convergence, we call the function

$$\tilde{f}\colon \mathbb{T}^d\to\mathbb{C},$$

pointwise defined by the limit

$$\tilde{f}(\mathbf{x}) := \lim_{N\to\infty} \mathcal{S}_N f(\mathbf{x}),$$

the Fourier series of the function f. □

If the Fourier coefficients of a function $f \in C\left(\mathbb{T}^d\right)$ fulfill $\sum_{\mathbf{z}\in\mathbb{Z}^d} |c_{\mathbf{z}}(f)| < \infty$, then the Fourier series converges absolutely and thus uniformly to the function f and it follows $f(\mathbf{x}) = \lim_{N\to\infty} \mathcal{S}_N f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{T}^d$. However, in general it is challenging to evaluate the integral (3.6) accurately. The numerical integration with a trapezoidal rule leads to the definition of the trigonometric interpolant.

Definition 7. Let a spatial dimension $d \in \mathbb{N}$ and a discretization parameter $N \in 2\mathbb{N}$ be given. We define the discrete Fourier coefficients $\hat{f}_{\mathbf{z}} \in \mathbb{C}$, $\mathbf{z} \in J_N$, of a vector $\mathbf{f} \in \mathbb{R}^{N^d}$ by

$$\hat{f}_{\mathbf{z}} := \frac{1}{N^d} \sum_{\mathbf{k}\in J_N} f_{\mathbf{k}} \mathrm{e}^{-2\pi \mathrm{i}\mathbf{z}\cdot\mathbf{x}_{\mathbf{k}}}, \quad \mathbf{z} \in J_N, \quad \mathbf{x}_{\mathbf{k}} := \frac{\mathbf{k}}{N} + \frac{1}{2N} \in X_N, \tag{3.7}$$

and call the map $\mathbf{f} \mapsto \hat{\mathbf{f}}$ the discrete Fourier transform of the vector $\mathbf{f}$. Moreover, we define the (complex) trigonometric interpolant $\mathcal{I}_N f \colon \mathbb{T}^d \to \mathbb{C}$ of a function $f \in C\left(\mathbb{T}^d\right)$ by

$$\mathcal{I}_N f(\mathbf{x}) := \sum_{\mathbf{z}\in J_N} \hat{f}_{\mathbf{z}} \mathrm{e}^{2\pi \mathrm{i}\mathbf{z}\cdot\mathbf{x}}, \tag{3.8}$$

where we denote by $\hat{f}_{\mathbf{z}} \in \mathbb{C}$, $\mathbf{z} \in J_N$, the discrete Fourier coefficients of the vector $\left(f\left(\mathbf{x}_{\mathbf{k}}\right)\right)_{\mathbf{k}\in J_N}$. □

Remark 8. *It is also common to define the discrete Fourier coefficients from equation (3.7) with points $\mathbf{x}_{\mathbf{k}}$ on an asymmetric grid,*

$$\tilde{f}_{\mathbf{z}} := \frac{1}{N^d} \sum_{\mathbf{k}\in J_N} f_{\mathbf{k}} \mathrm{e}^{-2\pi \mathrm{i}\mathbf{z}\cdot\mathbf{x}_{\mathbf{k}}}, \quad \mathbf{z} \in J_N, \quad \mathbf{x}_{\mathbf{k}} := \frac{\mathbf{k}}{N}. \tag{3.9}$$

If we allow indices $\mathbf{z} \in \mathbb{Z}^d$, then the Definitions (3.7) and (3.9) differ slightly in the periodicity of the discrete Fourier coefficients. For spatial dimension $d = 1$, it follows

$$\hat{f}_{z+N} = -\hat{f}_z \quad \textit{and} \quad \tilde{f}_{z+N} = \tilde{f}_z$$

for all $z \in \mathbb{Z}$ and appropriate relations hold for higher dimensions $d \in \mathbb{N}$.

In applications we are interested in real-valued approximations, which can be obtained by a modification of the trigonometric interpolant.

Definition 9. Let a spatial dimension $d \in \mathbb{N}$ and a discretization parameter $N \in 2\mathbb{N}$ be given. We define the real trigonometric interpolant $\mathcal{T}_N f \colon \mathbb{T}^d \to \mathbb{C}$ of a function $f \in C\left(\mathbb{T}^d\right)$ by

$$\mathcal{T}_N f(\mathbf{x}) := \sum_{\mathbf{z}\in Z_N} b_{\mathbf{z}}(f) \mathrm{e}^{2\pi \mathrm{i}\mathbf{z}\cdot\mathbf{x}}$$

with the coefficients

$$(b_{\mathbf{z}}(f))_{\mathbf{z}\in Z_N} := \mathbf{A}_d\hat{\mathbf{f}} \in \mathbb{C}^{(N+1)^d}, \quad \mathbf{A}_d := \mathbf{A}_1 \otimes \mathbf{A}_{d-1} \in \mathbb{C}^{(N+1)^d \times N^d},$$

and

$$\mathbf{A}_1 := \begin{pmatrix} \frac{1}{2} & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \\ -\frac{1}{2} & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{(N+1)\times N}.$$

□

The notation "interpolant" in Definitions 7 and 9 is justified, since basic computations verify for a given function $f \in C\left(\mathbb{T}^d\right)$ and a discretization parameter $N \in 2\mathbb{N}$ the interpolation property

$$f(\mathbf{x}) = \mathcal{I}_N f(\mathbf{x}) = \mathcal{T}_N f(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in X_N$$

on the regular Cartesian grid X_N. A relation between the Fourier coefficients $c_{\mathbf{k}}(f)$ and the discrete Fourier coefficients $\hat{f}_{\mathbf{k}}$ of a function f is given by the aliasing formula

$$\hat{f}_{\mathbf{k}} = \sum_{\mathbf{z}\in\mathbb{Z}^d} (-1)^{z_1+\dots+z_d} c_{\mathbf{k}+\mathbf{z}N}(f) \quad \text{for all} \quad \mathbf{k} \in J_N, \tag{3.10}$$

which implies the following estimates.

Lemma 10. *Let a spatial dimension $d \in \mathbb{N}$, a discretization parameter $N \in 2\mathbb{N}$, and a function $f \in C(\mathbb{T}^d)$ with an absolutely convergent Fourier series*

$$f(\mathbf{x}) = \sum_{\mathbf{z}\in\mathbb{Z}^d} c_{\mathbf{z}}(f) \mathrm{e}^{2\pi \mathrm{i}\mathbf{z}\cdot\mathbf{x}}, \quad c_{\mathbf{z}}(f) \in \mathbb{C},$$

be given. Then the estimates

$$\|f - \mathcal{S}_N f\|_{L^\infty(\mathbb{T}^d)} \le \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |c_{\mathbf{z}}(f)|, \quad \|f - \mathcal{I}_N f\|_{L^\infty(\mathbb{T}^d)} \le 2 \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |c_{\mathbf{z}}(f)|,$$

and

$$\|f - \mathcal{T}_N f\|_{L^\infty(\mathbb{T}^d)} \le 4 \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_{N-2}} |c_{\mathbf{z}}(f)|$$

are valid.

Proof. The first inequality follows from the definition and the other two from the aliasing formula (3.10), see for example [64, Chap. X, Thm. (5.16)]. □

It follows from Lemma 10, that we obtain a high approximation quality by finite Fourier sums for functions with a fast decrease of the Fourier coefficients. We call such functions smooth and measure the degree of the smoothness with the following norms.

Definition 11. For a spatial dimension $d \in \mathbb{N}$, a smoothness parameter $s \geq 0$, and a function $f \in C\left(\mathbb{T}^d\right)$, we define the following norms,

$$\|f\|_{s,1} := \sum_{\mathbf{z}\in\mathbb{Z}^d} (1+|\mathbf{z}|)^s |c_{\mathbf{z}}(f)|, \quad \|f\|_{s,2} := \left(\sum_{\mathbf{z}\in\mathbb{Z}^d} |c_{\mathbf{z}}(f)|^2 (1+|\mathbf{z}|^2)^s\right)^{\frac{1}{2}},$$

and

$$\|f\|_{s,\infty} := \max_{\mathbf{z}\in\mathbb{Z}^d} (1+|\mathbf{z}|)^s |c_{\mathbf{z}}(f)|.$$

□

To prove some error estimates for the finite Fourier sums, we start with the following statement.

Lemma 12. *Let a spatial dimension $d \in \mathbb{N}$ and parameters $c > 0$, $N \in 2\mathbb{N}$, $N \geq 4\sqrt{d}$, and $s > \frac{d}{c}$ be given. Then the estimate*

$$\sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} (1+|\mathbf{z}|^c)^{-s} \leq \frac{2^{2cs-d-1}\omega_{d-1}}{cs-d} \cdot N^{d-cs}$$

is valid.

Proof. We start with

$$(1+|\mathbf{z}|^c)^{-s} \leq |\mathbf{z}|^{-cs}$$

for all $\mathbf{z} \in \mathbb{Z}^d$ and

$$\begin{aligned}\sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |\mathbf{z}|^{-cs} &\leq \int_{|\mathbf{x}|\geq\frac{N}{2}} \left(|\mathbf{x}|-\sqrt{d}\right)^{-cs} \mathrm{d}\mathbf{x} = \int_{\frac{N}{2}}^{\infty} \int_{\mathbb{S}^{d-1}} \left(r-\sqrt{d}\right)^{-cs} r^{d-1} \mathrm{d}\sigma(\boldsymbol{\xi})\mathrm{d}r \\ &= \omega_{d-1} \int_{\frac{N}{2}-\sqrt{d}}^{\infty} r^{-cs}(r+\sqrt{d})^{d-1}\mathrm{d}r \leq 2^{d-1}\omega_{d-1} \int_{\frac{N}{4}}^{\infty} r^{-cs+d-1}\mathrm{d}r\end{aligned}$$

completes the proof. □

We apply Lemma 12 and obtain different error bounds, which explicitly show, that a higher smoothness of the function, given by a larger smoothness parameter s, leads to a faster decrease of the approximation error.

Lemma 13. *Let a spatial dimension $d \in \mathbb{N}$, a discretization parameter $N \in 2\mathbb{N}$, $N \geq 4\sqrt{d}$, and a continuous function $f \in C(\mathbb{T}^d)$ be given. With*

$$C_1(s,d) := \left(\frac{2^{4s-d-1}}{2s-d}\omega_{d-1}\right)^{\frac{1}{2}}, \ s > \frac{d}{2}, \quad \text{and} \quad C_2(s,d) := \frac{2^{2s-d-1}}{s-d}\omega_{d-1}, \ s > d,$$

the following estimates are valid. For $s \geq 0$ and $\|f\|_{s,1} < \infty$, it follows

$$\|f - \mathcal{S}_N f\|_\infty \leq 2^{-s}\|f\|_{s,1} N^{-s} \quad \text{and} \quad \|f - \mathcal{I}_N f\|_\infty \leq 2^{-s+1}\|f\|_{s,1} N^{-s},$$

for $s > \frac{d}{2}$ and $\|f\|_{s,2} < \infty$, it follows

$$\|f - \mathcal{S}_N f\|_\infty \leq C_1(s,d)\|f\|_{s,2} N^{\frac{d}{2}-s} \quad \text{and} \quad \|f - \mathcal{I}_N f\|_\infty \leq 2C_1(s,d)\|f\|_{s,2} N^{\frac{d}{2}-s},$$

and for $s > d$ and $\|f\|_{s,\infty} < \infty$, it follows

$$\|f - \mathcal{S}_N f\|_\infty \leq C_2(s,d)\|f\|_{s,\infty} N^{d-s} \quad \text{and} \quad \|f - \mathcal{I}_N f\|_\infty \leq 2C_2(s,d)\|f\|_{s,\infty} N^{d-s}.$$

Proof. We give the proof only for the Fourier partial sum $\mathcal{S}_N$; the estimates for the trigonometric interpolant can be shown in the same way. Lemma 10 yields

$$\begin{aligned}\|f - \mathcal{S}_N f\|_\infty &\leq \sum_{\mathbf{z}\in\mathbb{Z}^d} |c_{\mathbf{z}}(f)|(1+|\mathbf{z}|)^s(1+|\mathbf{z}|)^{-s} \\ &\leq \left(\max_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N}(1+|\mathbf{z}|)^{-s}\right)\sum_{\mathbf{z}\in\mathbb{Z}^d}|c_{\mathbf{z}}(f)|(1+|\mathbf{z}|)^s \leq \left(\frac{N}{2}\right)^{-s}\|f\|_{s,1}.\end{aligned}$$

Again Lemma 10, the Cauchy-Schwarz inequality, and Lemma 12 lead to

$$\begin{aligned}\|f - \mathcal{S}_N f\|_\infty &\leq \left(\sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N}|c_{\mathbf{z}}(f)|^2(1+|\mathbf{z}|^2)^s\right)^{\frac{1}{2}}\left(\sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N}(1+|\mathbf{z}|^2)^{-s}\right)^{\frac{1}{2}} \\ &\leq C_1(s,d)\|f\|_{s,2}N^{\frac{d}{2}-s}.\end{aligned}$$

Finally, Lemma 10, Lemma 12, and

$$\|f - \mathcal{S}_N f\|_\infty \leq \left(\max_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N}|c_{\mathbf{z}}(f)|(1+|\mathbf{z}|)^s\right)\sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N}(1+|\mathbf{z}|)^{-s}$$

show the last inequality. □

The Fourier coefficients of periodic functions are given by Definition 6. For functions on $\mathbb{R}^d$, a generalization is given in the following way.

Definition 14. Let a spatial dimension $d \in \mathbb{N}$ and an absolute integrable function $f \in C\left(\mathbb{R}^d\right)$ be given. We define the Fourier transform $\hat{f}\colon \mathbb{R}^d \to \mathbb{C}$ of the function f by

$$\hat{f}(\mathbf{v}) := \int_{\mathbb{R}^d} f(\mathbf{x})\mathrm{e}^{-2\pi\mathrm{i}\mathbf{v}\cdot\mathbf{x}}\mathrm{d}\mathbf{x}.$$

□

3.4 Mean value operators

The fundamental subject of this thesis, the spherical mean value operator, is given by the following Definition 15.

Definition 15. Let a spatial dimension $d \in \mathbb{N}\setminus\{1\}$ be given. The spherical mean value operator $\mathcal{M}\colon C\left(\mathbb{R}^d\right) \to C\left(\mathbb{R}^d\times\mathbb{R}\right)$,

$$\mathcal{M}f(\mathbf{y},r) := \frac{1}{\omega_{d-1}}\int_{\mathbb{S}^{d-1}} f(\mathbf{y}+r\boldsymbol{\xi})\mathrm{d}\sigma(\boldsymbol{\xi}),$$

assigns to functions $f \in C\left(\mathbb{R}^d\right)$ mean values over $(d-1)$-dimensional spheres with center points $\mathbf{y} \in \mathbb{R}^d$ and radii $r \in \mathbb{R}$. □

To make sure, that the previous Definition 15 of the spherical mean value operator is correct, it remains to show the continuity of $\mathcal{M}f\colon \mathbb{R}^d\times\mathbb{R} \to \mathbb{R}$.

Lemma 16. *Let a spatial dimension $d \in \mathbb{N}\setminus\{1\}$ and a continuous function $f \in C\left(\mathbb{R}^d\right)$ be given. Then, it follows $\mathcal{M}f \in C\left(\mathbb{R}^d \times \mathbb{R}\right)$.*

Proof. We choose a center point $\mathbf{y} \in \mathbb{R}^d$, a radius $r \in \mathbb{R}$, and a sequence

$$(\mathbf{y}_k, r_k) \in \mathbb{R}^d \times \mathbb{R}, \quad k \in \mathbb{N},$$

with $\lim_{k\to\infty}(\mathbf{y}_k, r_k) = (\mathbf{y}, r)$ arbitrarily and define the set

$$A := \left\{\mathbf{y}_k + r_k\boldsymbol{\xi} : \boldsymbol{\xi} \in \mathbb{S}^{d-1} \text{ and } k \in \mathbb{N}\right\}.$$

The function f is continuous and hence bounded on A and it follows

$$s := \sup_{\mathbf{x}\in A} |f(\mathbf{x})| < \infty.$$

Furthermore, each element f_k of the function sequence

$$f_k \colon \mathbb{S}^{d-1} \to \mathbb{R}, \quad f_k(\boldsymbol{\xi}) := f\left(\mathbf{y}_k + r_k\boldsymbol{\xi}\right), \quad k \in \mathbb{N},$$

is continuous and thus integrable on $\mathbb{S}^{d-1}$. Since

$$|f_k(\boldsymbol{\xi})| \le s \quad \text{for all} \quad k \in \mathbb{N} \quad \text{and} \quad \boldsymbol{\xi} \in \mathbb{S}^{d-1}$$

is fulfilled, we can apply Lebesgue's dominated convergence theorem, obtain

$$\begin{aligned}\lim_{k\to\infty} \mathcal{M}f\left(\mathbf{y}_k, r_k\right) &= \frac{1}{\omega_{d-1}} \lim_{k\to\infty} \int_{\mathbb{S}^{d-1}} f_k(\boldsymbol{\xi}) \mathrm{d}\sigma(\boldsymbol{\xi}) \\ &= \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} \lim_{k\to\infty} f_k(\boldsymbol{\xi}) \mathrm{d}\sigma(\boldsymbol{\xi}) = \mathcal{M}f(\mathbf{y}, r)\end{aligned}$$

and this proves the assertion. □

It is well known [11, §VI.13, eq. 13-15], that the solution of the Cauchy problem for the d-dimensional wave equation, $d \in \mathbb{N} \setminus \{1\}$,

$$\begin{aligned}\partial_r^2 p(\mathbf{y}, r) - \Delta p(\mathbf{y}, r) &= 0 \quad \text{for } (\mathbf{y}, r) \in \mathbb{R}^d \times (0, \infty), \\ p(\mathbf{y}, 0) &= f(\mathbf{y}) \quad \text{for } \mathbf{y} \in \mathbb{R}^d, \\ \partial_r p(\mathbf{y}, 0) &= 0 \quad \text{for } \mathbf{y} \in \mathbb{R}^d,\end{aligned} \tag{3.11}$$

is given as

$$p_d(\mathbf{y}, r) = \begin{cases} \frac{1}{\Gamma\left(\frac{d}{2}\right) 2^{\frac{d-2}{2}}} \frac{\partial}{\partial r} \int_0^r \frac{\rho}{\sqrt{r^2-\rho^2}} \left(\frac{1}{\rho}\frac{\partial}{\partial \rho}\right)^{\frac{d-2}{2}} \rho^{d-2} \mathcal{M}f\left(\mathbf{y}, \rho\right) \mathrm{d}\rho & \text{for even } d, \\ \frac{\sqrt{\pi}}{\Gamma\left(\frac{d}{2}\right) 2^{\frac{d-1}{2}}} \frac{\partial}{\partial r} \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{\frac{d-3}{2}} \left(r^{d-2} \mathcal{M}f(\mathbf{y}, r)\right) & \text{for odd } d. \end{cases}$$

In particular, for spatial dimension $d = 3$, we obtain the representation

$$p_3(\mathbf{y}, r) = \frac{\partial}{\partial r}\left(r\mathcal{M}f(\mathbf{y}, r)\right)$$

and spatial dimension $d = 2$ leads to

$$\begin{aligned}p_2(\mathbf{y}, r) &= \frac{\partial}{\partial r} \int_0^r \frac{\rho}{\sqrt{r^2-\rho^2}} \mathcal{M}f\left(\mathbf{y}, \rho\right) \mathrm{d}\rho = \frac{\partial}{\partial r} r \int_0^1 \mathcal{M}f\left(\mathbf{y}, r\rho\right) \frac{\rho \mathrm{d}\rho}{\sqrt{r^2-\rho^2}} \\ &= \frac{\partial}{\partial r} r \left(\int_0^1 \int_0^{2\pi} f\left(\mathbf{y} + r\rho \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}\right) \mathrm{d}\varphi \frac{\rho \mathrm{d}\rho}{\sqrt{r^2-\rho^2}} \right),\end{aligned}$$

which motivates the introduction of a further mean value operator.

Definition 17. The mean value operator $\mathcal{N}\colon C(\mathbb{R}^2) \to C(\mathbb{R}^2 \times \mathbb{R})$ is defined by

$$\begin{aligned}\mathcal{N}f(\mathbf{y}, r) &:= \frac{1}{2\pi} \int_{|\mathbf{x}|\leq 1} f(\mathbf{y} + r\mathbf{x}) \frac{\mathrm{d}\mathbf{x}}{\sqrt{1 - |\mathbf{x}|^2}} \\ &= \frac{1}{2\pi} \int_0^1 \int_0^{2\pi} f\left(\mathbf{y} + r\rho \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}\right) \frac{\mathrm{d}\varphi\rho\mathrm{d}\rho}{\sqrt{1-\rho^2}}.\end{aligned}$$

□

Remark 18. *To be complete, we have to show the continuity of $\mathcal{N}f\colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ for a continuous function $f \in C(\mathbb{R}^2)$. The substitution $\rho = \sin\alpha$ leads to*

$$\mathcal{N}f(\mathbf{y}, r) = \frac{1}{2\pi} \int_0^{\frac{\pi}{2}} \int_0^{2\pi} f\left(\mathbf{y} + r\sin\alpha \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}\right) \sin\alpha \mathrm{d}\varphi \mathrm{d}\alpha$$

and this motivates the definition of the function sequence

$$f_k : [0, 2\pi] \times \left[0, \frac{\pi}{2}\right] \to \mathbb{R}, \quad f_k(\varphi, \alpha) := f\left(\mathbf{y}_k + r_k \sin\alpha \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}\right) \sin\alpha,$$

for a given sequence $(\mathbf{y}_k, r_k) \in \mathbb{R}^2 \times \mathbb{R}$, $k \in \mathbb{N}$. The arguments from the proof of Lemma 16 show the continuity of $\mathcal{N}f$.

This thesis focuses on Fourier based methods, where we have to deal with complex exponential functions. Spherical integrals over this basic functions lead to Bessel functions.

Definition 19. For an order $\nu \geq 0$, we define the Bessel functions of the first kind $\mathcal{J}_\nu\colon \mathbb{R} \to \mathbb{R}$ by the integral representation, see [62, Sec. 3.3, eq. (6)],

$$\mathcal{J}_\nu(x) := \frac{x^\nu}{\Gamma\left(\nu + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) 2^\nu} \int_0^\pi \mathrm{e}^{\mathrm{i}x\cos\xi} (\sin\xi)^{2\nu} \mathrm{d}\xi. \tag{3.12}$$

□

We continue with the computation of mean values for the complex exponential function. This fundamental statement builds the basis of the Fourier based algorithms in this thesis.

Theorem 20 ([42, eq. (1.5)])**.** *Let a spatial dimension $d \in \mathbb{N}\setminus\{1\}$ and a vector $\mathbf{v} \in \mathbb{R}^d$ be given. The mean values of the function*

$$\mathrm{e}_\mathbf{v}\colon \mathbb{R}^d \to \mathbb{C}, \quad \mathrm{e}_\mathbf{v}(\mathbf{x}) := \mathrm{e}^{2\pi\mathrm{i}\mathbf{v}\cdot\mathbf{x}},$$

according to a center point $\mathbf{y} \in \mathbb{R}^d$ and a radius $r \geq 0$ can be computed by

$$\mathcal{N}\mathrm{e}_\mathbf{v}(\mathbf{y}, r) = \begin{cases} \mathrm{e}_\mathbf{v}(\mathbf{y}) & \text{for } r = 0 \text{ or } \mathbf{v} = \mathbf{0}, \\ \frac{\sin(2\pi r|\mathbf{v}|)}{2\pi r|\mathbf{v}|} \cdot \mathrm{e}_\mathbf{v}(\mathbf{y}) & \text{otherwise}, \end{cases}$$

for $d = 2$, and

$$\mathcal{M}\mathrm{e}_\mathbf{v}(\mathbf{y}, r) = \begin{cases} \mathrm{e}_\mathbf{v}(\mathbf{y}) & \text{for } r = 0 \text{ or } \mathbf{v} = \mathbf{0}, \\ \frac{\Gamma\left(\frac{d}{2}\right)\mathcal{J}_{\frac{d}{2}-1}(2\pi r|\mathbf{v}|)}{(\pi r|\mathbf{v}|)^{\frac{d}{2}-1}} \cdot \mathrm{e}_\mathbf{v}(\mathbf{y}) & \text{otherwise}, \end{cases}$$

for general dimension $d \in \mathbb{N}\setminus\{1\}$.

Proof. The assertions for $r = 0$ and $\mathbf{v} = \mathbf{0}$ follow directly from the definition of the mean value operators. Otherwise, let $\mathbf{R} \in \mathbb{R}^{d\times d}$ be a rotation around the origin with

$$\mathbf{R}\mathbf{v} = |\mathbf{v}|\mathbf{e}_d, \quad \mathbf{e}_d := \big(0,\dots,0,1\big)^\top \in \mathbb{R}^d.$$

It follows

$$\begin{aligned}
\int_{|\mathbf{x}|\le 1} \mathrm{e}^{2\pi \mathrm{i} r\mathbf{v}\cdot\mathbf{x}} \frac{\mathrm{d}\mathbf{x}}{\sqrt{1-|\mathbf{x}|^2}} &= \int_{|\mathbf{R}^{-1}\mathbf{x}|\le 1} \mathrm{e}^{2\pi \mathrm{i} r\mathbf{v}\cdot\mathbf{R}^{-1}\mathbf{x}} \frac{\mathrm{d}\mathbf{x}}{\sqrt{1-|\mathbf{R}^{-1}\mathbf{x}|^2}} \\
&= \int_0^1 \int_0^{2\pi} \mathrm{e}^{2\pi \mathrm{i} r|\mathbf{v}|\rho\sin\varphi} \mathrm{d}\varphi \frac{\rho \mathrm{d}\rho}{\sqrt{1-\rho^2}} \\
&= 2\int_0^1 \int_0^{\pi} \mathrm{e}^{2\pi \mathrm{i} r|\mathbf{v}|\rho\cos\varphi} \mathrm{d}\varphi \frac{\rho \mathrm{d}\rho}{\sqrt{1-\rho^2}} \\
&= 2\pi \int_0^1 \mathcal{J}_0(2\pi r|\mathbf{v}|\rho) \frac{\rho \mathrm{d}\rho}{\sqrt{1-\rho^2}} \\
&= 2\pi \int_0^{\frac{\pi}{2}} \mathcal{J}_0(2\pi r|\mathbf{v}|\sin\varphi)\sin\varphi \mathrm{d}\varphi
\end{aligned}$$

and Sonine's first finite integral [62, Sec. 12.11, eq. (1)] yields

$$\int_{|\mathbf{x}|\le 1} \mathrm{e}^{2\pi \mathrm{i} r\mathbf{v}\cdot\mathbf{x}} \frac{\mathrm{d}\mathbf{x}}{\sqrt{1-|\mathbf{x}|^2}} = \frac{\pi}{(r|\mathbf{v}|)^{\frac{1}{2}}} \mathcal{J}_{\frac{1}{2}}(2\pi r|\mathbf{v}|) = 2\pi \cdot \frac{\sin(2\pi r|\mathbf{v}|)}{2\pi r|\mathbf{v}|}.$$

Similar computations show

$$\begin{aligned}
\int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi \mathrm{i} r\mathbf{v}\cdot\boldsymbol{\xi}} \mathrm{d}\sigma(\boldsymbol{\xi}) &= \int_{\mathbf{R}\mathbb{S}^{d-1}} \mathrm{e}^{2\pi \mathrm{i} r\mathbf{v}\cdot\mathbf{R}^{-1}\boldsymbol{\xi}} \mathrm{d}\sigma(\boldsymbol{\xi}) = \int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi \mathrm{i} r|\mathbf{v}|\xi_d} \mathrm{d}\sigma(\boldsymbol{\xi}) \\
&= \int_0^{2\pi} \int_{[0,\pi]^{d-2}} \mathrm{e}^{2\pi \mathrm{i} r|\mathbf{v}|\cos\vartheta_{d-2}} \mathcal{D}_d(1,\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta} \mathrm{d}\varphi \\
&= \omega_{d-2} \int_0^{\pi} (\sin\vartheta_{d-2})^{d-2} \mathrm{e}^{2\pi \mathrm{i} r|\mathbf{v}|\cos\vartheta_{d-2}} \mathrm{d}\vartheta_{d-2} \\
&= \omega_{d-2} \frac{\Gamma\left(\frac{d-1}{2}\right)\Gamma\left(\frac{1}{2}\right) 2^{\frac{d}{2}-1}}{(2\pi r|\mathbf{v}|)^{\frac{d}{2}-1}} \mathcal{J}_{\frac{d}{2}-1}(2\pi r|\mathbf{v}|) \\
&= \omega_{d-1} \frac{\Gamma\left(\frac{d}{2}\right) \mathcal{J}_{\frac{d}{2}-1}(2\pi r|\mathbf{v}|)}{(\pi r|\mathbf{v}|)^{\frac{d}{2}-1}}.
\end{aligned}$$

Finally,

$$\mathcal{N}\mathrm{e}_{\mathbf{v}}(\mathbf{y},r) = \frac{1}{2\pi} \int_{|\boldsymbol{\xi}|\le 1} \mathrm{e}^{2\pi \mathrm{i}\mathbf{v}\cdot(\mathbf{y}+r\boldsymbol{\xi})} \frac{\mathrm{d}\sigma(\boldsymbol{\xi})}{\sqrt{1-|\boldsymbol{\xi}|^2}} = \frac{1}{2\pi} \int_{|\boldsymbol{\xi}|\le 1} \mathrm{e}^{2\pi r\mathrm{i}\mathbf{v}\cdot\boldsymbol{\xi}} \frac{\mathrm{d}\sigma(\boldsymbol{\xi})}{\sqrt{1-|\boldsymbol{\xi}|^2}} \cdot \mathrm{e}_{\mathbf{v}}(\mathbf{y})$$

and

$$\mathcal{M}\mathrm{e}_{\mathbf{v}}(\mathbf{y},r) = \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi \mathrm{i}\mathbf{v}\cdot(\mathbf{y}+r\boldsymbol{\xi})} \mathrm{d}\sigma(\boldsymbol{\xi}) = \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} \mathrm{e}^{2\pi r\mathrm{i}\mathbf{v}\cdot\boldsymbol{\xi}} \mathrm{d}\sigma(\boldsymbol{\xi}) \cdot \mathrm{e}_{\mathbf{v}}(\mathbf{y})$$

complete the proof. □

4

Efficient discretization

Definitions 15 and 17 introduced mean values for continuous functions. In common applications, the data is given by discrete values, which motivates the following problem. Let a spatial dimension $d \in \mathbb{N} \setminus \{1\}$, samples $\mathbf{f}$ of a continuous function $f \in C\left(\mathbb{R}^d\right)$, $\operatorname{supp} f \subset \mathbb{I}^d$, on a regular grid X_N,

$$\mathbf{f} = (f(\mathbf{x}))_{\mathbf{x} \in X_N} \in \mathbb{R}^{N^d}, \quad N \in 2\mathbb{N}, \tag{4.1}$$

and sets of center points Y and radii R,

$$Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_{M_1}\} \subset \mathbb{R}^d \quad \text{and} \quad R = \{r_1, \ldots, r_{M_2}\} \subset (0, \infty), \quad M_1, M_2 \in \mathbb{N},$$

be given. The task is to find an efficient approximation of the mean values

$$\mathcal{M}f\left(\mathbf{y}_j, r_k\right), \quad j = 1, \ldots, M_1, \quad k = 1, \ldots, M_2. \tag{4.2}$$

In the following, we consider different discretizations of the mean value operators. Excluding Section 4.2.2.1, this part of the thesis has already been published in [25, 26]. For comparison, we start with approximations of integrals by a simple rectangular rule, which can also be generalized to quadratures of higher orders. The main part of this thesis is a Fourier based approach, which is based on accurate computation of mean values for finite Fourier sums. For frequencies on a rectangular Cartesian grid, this leads to evaluations of nonequispaced discrete Fourier sums. Furthermore, we present a fast discretization of the two-dimensional mean value operator $\mathcal{N}$ for functions, which are given by its Fourier coefficients on polar frequencies.

4.1 Numerical integration

A standard approach for the discretization of an integral operator is the numerical integration, also known as quadrature. The main idea is to replace the integrand by a piecewise polynomial, which is easy to integrate. A simple class of functions for this purpose are piecewise constant functions, labeled as step functions, which assign the function value of the nearest grid point.

Definition 21 (*Approximating step function*)**.** Let a spatial dimension $d \in \mathbb{N}$, a continuous function $f \in C\left(\mathbb{R}^d\right)$, $\operatorname{supp} f \subset \mathbb{I}^d$, and a discretization parameter $N \in 2\mathbb{N}$ be given. The approximating step function $f_Q \colon \mathbb{R}^d \to \mathbb{R}$ of f is defined by

$$f_Q(\mathbf{x}) := f\left(\frac{\left[N \cdot \mathbf{x} - \frac{1}{2}\right] + \frac{1}{2}}{N}\right),$$

where $[\mathbf{x}]$ denotes the componentwise rounding of the vector $\mathbf{x}$. □

We approximate the mean values by replacing the function by its approximating step function and using a rectangular rule for the resulting integral afterwards. For spatial dimension $d = 2$, a fixed center point $\mathbf{y} \in \mathbb{R}^2$, and a fixed radius $r \in (0, \infty)$, this leads to

$$\begin{aligned}\mathcal{M}f(\mathbf{y}, r) \approx \mathcal{M}f_Q(\mathbf{y}, r) &= \frac{1}{2\pi} \int_0^{2\pi} f_Q\left(\mathbf{y} + r\begin{pmatrix}\cos\varphi \\ \sin\varphi\end{pmatrix}\right) \mathrm{d}\varphi \\ &\approx \frac{1}{n} \sum_{l=0}^{n-1} \cdot f_Q\left(\mathbf{y} + r\begin{pmatrix}\cos\varphi_l \\ \sin\varphi_l\end{pmatrix}\right),\end{aligned}$$

where $n \in \mathbb{N}$ denotes the number of nodes for the rectangular rule and

$$\varphi_l := \frac{2\pi l}{n}, \quad l \in \{0, \ldots, n-1\}.$$

A suitable choice of this parameter requires some theory about quadratures. For our purposes it is sufficient to use the heuristic argument, that the distance of two neighboring nodes $\mathbf{x}_l$ and $\mathbf{x}_{l+1}$ should be less than the distance of two neighboring grid points, so for $l \in \{0, \ldots, n-2\}$ we demand

$$|\mathbf{x}_{l+1} - \mathbf{x}_l| \leq r|\varphi_{l+1} - \varphi_l| \leq \frac{1}{N} \Rightarrow \frac{2\pi r}{n} \leq \frac{1}{N} \Leftrightarrow n \geq 2\pi r N.$$

The whole procedure is summarized in Algorithm 1 with the numerical complexity of $\mathcal{O}\left(M_1 M_2 N\right)$ operations.
An appropriate computation for spatial dimension $d = 3$ leads to

$$\begin{aligned}\mathcal{M}f(\mathbf{y}, r) &= \frac{1}{4\pi} \int_0^{\pi} \int_0^{2\pi} f_Q\left(\mathbf{y} + r\begin{pmatrix}\sin\vartheta\cos\varphi \\ \sin\vartheta\sin\varphi \\ \cos\vartheta\end{pmatrix}\right) \mathrm{d}\varphi \sin\theta \mathrm{d}\vartheta \\ &\approx \frac{\pi}{2n} \sum_{l=0}^{n-1} \frac{\sin\vartheta_l}{n_l} \sum_{m=0}^{n_l-1} f_Q\left(\mathbf{y} + r\begin{pmatrix}\sin\vartheta_l\cos\varphi_{l,m} \\ \sin\vartheta_l\sin\varphi_{l,m} \\ \cos\vartheta_l\end{pmatrix}\right)\end{aligned}$$

with

$$n := \lceil \pi r N \rceil, \quad \vartheta_l := \frac{\pi\left(l + \frac{1}{2}\right)}{n}, \quad n_l := \lceil 2\pi N \sin\vartheta_l \rceil, \quad \text{and} \quad \varphi_{l,m} := \frac{2\pi m}{n_l},$$

and for $d = 2$ the mean value operator $\mathcal{N}$ can be discretized as

$$(\mathcal{N}f)(\mathbf{y},r) = \frac{1}{2\pi}\int_0^1\int_0^{2\pi} f\left(\mathbf{y} + r\begin{pmatrix}\rho\cos\varphi\\ \rho\sin\varphi\end{pmatrix}\right)\mathrm{d}\varphi\frac{\rho}{\sqrt{1-\rho^2}}\mathrm{d}\rho$$
$$\approx \frac{1}{n}\sum_{l=0}^{n-1}\frac{1}{n_l}\sum_{m=0}^{n_l-1} f\left(\mathbf{y} + r\begin{pmatrix}\rho_l\cos\varphi_{l,m}\\ \rho_l\sin\varphi_{l,m}\end{pmatrix}\right)\frac{\rho_l}{\sqrt{1-\rho_l^2}}$$

with

$$n := \lceil rN\rceil, \quad \rho_l := \frac{l}{n} + \frac{1}{2n}, \quad n_l := \lceil 2\pi r\rho_l N\rceil, \quad \text{and} \quad \varphi_{l,m} := \frac{2\pi m}{n_l}.$$

Algorithms 2 and 3 with numerical complexity of $\mathcal{O}(M_1M_2N^2)$ operations realize both results.

Remark 22. *The previous approach can be improved by using polynomials and quadratures of higher order, which leads to a higher accuracy of the numerical integration. However, the complexity of the algorithms has still the same order. The concept of Algorithms 1 and 2 can also be generalized to higher dimensions, which produces a numerical complexity of* $\mathcal{O}\left(M_1M_2N^{d-1}\right)$ *operations.*

Algorithm 1 Discrete mean value operator $\mathcal{M}$, using quadrature, $d = 2$.

Input

$d \in \mathbb{N}$ ▷ spatial dimension
$N \in 2\mathbb{N}$, $M_1 \in \mathbb{N}$, $M_2 \in \mathbb{N}$ ▷ discretization parameters
$\mathbf{f} \in \mathbb{R}^{N^d}$ ▷ samples
$\mathbf{y}_j \in \mathbb{R}^d : j = 1,\dots,M_1$ ▷ center points
$r_k \in (0,\infty), k = 1,\dots,M_2$ ▷ radii

Output

$\mathbf{g} \in \mathbb{R}^{M_1M_2}$ ▷ mean values

for $k = 1,...,M_2$ **do**
  $n := \lceil 2\pi r_k N\rceil$ ▷ number of nodes
  for $j = 1,...,M_1$ **do**
    $s := 0$
    for $l = 0,...,n-1$ **do**
      $\varphi_l := \frac{2\pi l}{n}$ ▷ discretisation of the angle
      $\mathbf{x}_l := \mathbf{y}_j + r_k\begin{pmatrix}\cos\varphi_l\\ \sin\varphi_l\end{pmatrix}$ ▷ node
      $\widetilde{\mathbf{x}}_l := \frac{\left\lfloor N\cdot\mathbf{x}_l - \frac{1}{2}\right\rceil + \frac{1}{2}}{N}$ ▷ nearest grid point
      $s := s + f(\widetilde{\mathbf{x}}_l)$ ▷ summation
    end for
    $g_{j,k} := \frac{1}{n}s$
  end for
end for

Algorithm 2 Discrete mean value operator $\mathcal{M}$, using quadrature, $d = 3$.

Input and **Output** as in Algorithm 1

for $k = 1, ..., M_2$ **do**
  $n := \lceil \pi r_k N \rceil$ ▷ number of nodes for the polar angle
  for $j = 1, ..., M_1$ **do**
    $s_1 := 0$
    for $l = 0, ..., n-1$ **do**
      $\vartheta_l := \frac{\pi\left(l+\frac{1}{2}\right)}{n}$ ▷ discretization of the polar angle
      $n_l := \lceil 2\pi N \sin\vartheta_l \rceil$ ▷ number of nodes for the azimuthal angle
      $s_2 := 0$
      for $m = 0, ..., n_l - 1$ **do**
        $\varphi_{l,m} := \frac{2\pi m}{n_l}$ ▷ discretization of the azimuthal angle
        $\mathbf{x}_{l,m} := \mathbf{y}_j + r_k \begin{pmatrix} \sin\vartheta_l \cos\varphi_{l,m} \\ \sin\vartheta_l \sin\varphi_{l,m} \\ \cos\vartheta_l \end{pmatrix}$ ▷ node
        $\widetilde{\mathbf{x}}_{l,m} := \frac{\left[N\cdot\mathbf{x}_{l,m}-\frac{1}{2}\right]+\frac{1}{2}}{N}$ ▷ nearest grid point
        $s_2 := s_2 + f\left(\widetilde{\mathbf{x}}_{l,m}\right)$ ▷ summation
      end for
      $s_1 := s_1 + \frac{\sin\vartheta_l}{n_l} \cdot s_2$
    end for
    $g_{j,k} := \frac{\pi}{2n} \cdot s_1$
  end for
end for

Algorithm 3 Discrete mean value operator $\mathcal{N}$, using quadrature, $d = 2$.

Input and **Output** as in Algorithm 1

for $k = 1, ..., M_2$ **do**
  $n := \lceil r_k N \rceil$ ▷ number of nodes for the radius
  for $j = 1, ..., M_1$ **do**
    $s_1 := 0$
    for $l = 0, ..., n-1$ **do**
      $\rho_l := \frac{l}{n} + \frac{1}{2n}$ ▷ discretization of the radius
      $n_l := \lceil 2\pi r_k \rho_l N \rceil$ ▷ number of nodes for the angle
      $s_2 := 0$
      for $m = 0, ..., n_l - 1$ **do**
        $\varphi_{l,m} := \frac{2\pi m}{n_l}$ ▷ discretization of the angle
        $\mathbf{x}_{l,m} := \mathbf{y}_j + r_k \rho_l \begin{pmatrix} \cos \varphi_{l,m} \\ \sin \varphi_{l,m} \end{pmatrix}$ ▷ node
        $\widetilde{\mathbf{x}}_{l,m} := \frac{2\left[N \cdot \mathbf{x}_{l,m} + \frac{N}{2} - \frac{1}{2}\right] + 1 - N}{2N}$ ▷ nearest grid point
        $s_2 := s_2 + f\left(\widetilde{\mathbf{x}}_{l,m}\right) \frac{\rho_l}{\sqrt{1-\rho_l^2}}$ ▷ summation
      end for
      $s_1 := s_1 + \frac{s_2}{n_l}$
    end for
    $g_{j,k} := \frac{s_1}{n}$
  end for
end for

4.2 Fourier based approach

The following approach for disretizing the mean value operators is a fundamental part of this thesis. With the intention to use tools from Fourier analysis, we approximate functions by trigonometric functions $p\colon \mathbb{R}^d \to \mathbb{C}$,

$$p(\mathbf{x}) = \sum_{k=1}^{M} \hat{p}_k \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\xi}_k \cdot \mathbf{x}}, \quad M \in \mathbb{N}, \tag{4.3}$$

with coefficients $\hat{p}_k \in \mathbb{C}$ and frequencies $\boldsymbol{\xi}_k \in \mathbb{R}^d$, $k \in \{1, \dots, M\}$. The mean values of such functions are given by the application of Theorem 20.

Corollary 23. *Let a spatial dimension $d \in \mathbb{N} \setminus \{1\}$, a center point $\mathbf{y} \in \mathbb{R}^d$, a radius $r > 0$, and a trigonometric function $p\colon \mathbb{R}^d \to \mathbb{C}$ as in equation (4.3) with $\hat{p}_k \in \mathbb{C}$ and $\boldsymbol{\xi}_k \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, $k \in \{1, \dots, M\}$, $M \in \mathbb{N}$, be given. Then it follows*

$$\mathcal{M}p(\mathbf{y}, r) = \sum_{k=1}^{M} \tilde{h}_k \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\xi}_k \cdot \mathbf{y}} \quad \text{for } d \in \mathbb{N} \setminus \{1\} \tag{4.4}$$

and

$$\mathcal{N}p(\mathbf{y}, r) = \sum_{k=1}^{M} \tilde{g}_k \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\xi}_k \cdot \mathbf{y}} \quad \text{for } d = 2 \tag{4.5}$$

with

$$\tilde{h}_k := \hat{p_k} \cdot \frac{\Gamma\left(\frac{d}{2}\right) \mathcal{J}_{\frac{d-2}{2}}\left(2\pi r |\boldsymbol{\xi}_k|\right)}{\left(\pi r |\boldsymbol{\xi}_k|\right)^{\frac{d-2}{2}}} \quad \text{and} \quad \tilde{g}_k := \hat{p_k} \cdot \frac{\sin\left(2\pi r |\boldsymbol{\xi}_k|\right)}{2\pi r |\boldsymbol{\xi}_k|}, \quad k \in \{1, \dots, M\}.$$

Proof. The assertion follows from Theorem 20 and the linearity of the operators. □

In the following sections we discuss different methods for evaluating the sums in equations (4.4) and (4.5) with suitable choices for the frequencies $\boldsymbol{\xi}_k$.

4.2.1 Regular Cartesian frequency grids

A standard approximation of functions $f \in C\left(\mathbb{T}^d\right)$ by trigonometric polynomials is given by the trigonometric interpolant $\mathcal{I}_N f\colon \mathbb{T}^d \to \mathbb{C}$, see equation (3.8), with frequencies on a regular Cartesian grid J_N, see equation (3.1) and Figure 4.1. The appropriate Fourier coefficients are provided by the discrete Fourier transform, see equation (3.7). Finally, we approximate the mean values of a function f by

$$\mathcal{M}f(\mathbf{y}, r) \approx \mathcal{M}\mathcal{I}_N f(\mathbf{y}, r) \quad \text{and} \quad \mathcal{N}f(\mathbf{y}, r) \approx \mathcal{N}\mathcal{I}_N f(\mathbf{y}, r),$$

where the right sides are given by Corollary 23.

Remark 24. *While the trigonometric interpolant is restricted to functions on the torus $\mathbb{T}^d$, we can also handle nonperiodic functions, if some assumptions are fulfilled. For example, let a function $f \in C(\mathbb{R}^d)$ with bounded support,*

$$\operatorname{supp} f \subset \left\{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| \leq a\right\}, \quad 0 < a < \frac{1}{2},$$

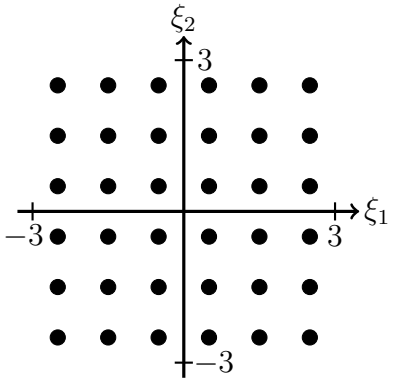


Figure 4.1: Regular Cartesian frequency grid for discretization parameter $N = 6$.

a center point $\mathbf{y} \in \mathbb{R}^d$*, and a radius* $r > 0$ *be given. Under the condition*

$$r + |\mathbf{y}| < 1 - a,$$

the associated periodic function $\tilde{f} : \mathbb{T}^d \to \mathbb{R}$

$$\tilde{f}(\mathbf{x}) := \sum_{\mathbf{z} \in \mathbb{Z}^d} f(\mathbf{x} + \mathbf{z})$$

and the nonperiodic function f *have the same mean values,*

$$\mathcal{M}f(\mathbf{y}, r) = \mathcal{M}\tilde{f}(\mathbf{y}, r) \quad \textit{and} \quad \mathcal{N}f(\mathbf{y}, r) = \mathcal{N}\tilde{f}(\mathbf{y}, r).$$

Remark 25. *The proposed approach is based on computing means of the trigonometric interpolant. For real-valued functions, the computed polynomials and means are nevertheless complex-valued in general. This matter can be avoided by applying the following procedure to the real trigonometric interpolant, see Definition 9. For reasons of simplicity, we restrict our further considerations to the complex-valued interpolation.*

4.2.1.1 Nonequispaced FFT based algorithm

For a regular frequency grid, a set of center points $\{\mathbf{y}_1, \dots, \mathbf{y}_{M_1}\}$, and a fixed radius r, the sums in equations (4.4) and (4.5) are nonequispaced discrete Fourier transforms, which can be straightforwardly evaluated.

Theorem 26. *Let* $\mathbf{f}$ *be the samples of a function* $f \colon \mathbb{T}^d \to \mathbb{R}$ *at the regular grid* $X_N \subset \mathbb{T}^d$ *with discretization parameter* $N \in 2\mathbb{N}$ *and spatial dimension* $d \in \mathbb{N} \setminus \{1\}$*. Then the spherical mean values*

$$g_{j,k} = \mathcal{M}\mathcal{I}_N f(\mathbf{y}_j, r_k)$$

at center points $\mathbf{y}_1, \dots, \mathbf{y}_{M_1} \in \mathbb{T}^d$*,* $M_1 \in \mathbb{N}$*, and radii* $r_1, \dots, r_{M_2} \in (0, 1]$*,* $M_2 \in \mathbb{N}$*, can be computed by Algorithm 4. If we evaluate the sum in line 16 approximately with a fixed accuracy, then it exists a fast realization of Algorithm 4 with the numerical complexity* $\mathcal{O}\left(M_2(N^d \log N + M_1)\right)$*.*

Proof. We start the algorithm by computing the discrete Fourier coefficients in line 8 by a fast Fourier transform with the numerical complexity $\mathcal{O}\left(N^d \log N\right)$, [10]. In line 12 we continue for each radius $r_1, \ldots, r_{M_2}$ with a multiplication of vectors with length N^d. The numerical complexity of this step is $\mathcal{O}\left(M_2 N^d\right)$. Finally, for each radius the evaluation of the nonequispaced discrete Fourier transform (NDFT) in line 16 can be approximated with fixed accuracy by a nonequispaced FFT (NFFT), which produces the leading numerical complexity $\mathcal{O}\left(M_2\left(N^d \log N + M_1\right)\right)$, see [43]. Corollary 23 completes the proof. □

Algorithm 4 Discrete mean value operator $\mathcal{M}$, using NFFT [43].

Input

1: $d \in \mathbb{N} \setminus \{1\}$ ▷ spatial dimension
2: $N \in 2\mathbb{N}$, $M_1 \in \mathbb{N}$, $M_2 \in \mathbb{N}$ ▷ discretization parameter
3: $\mathbf{f} \in \mathbb{R}^{N^d}$ ▷ samples
4: $\mathbf{y}_j \in \mathbb{T}^d : j = 1, \ldots, M_1$ ▷ center points
5: $r_k \in (0, 1], k = 1, \ldots, M_2$ ▷ radii

Output

6: $\mathbf{g} \in \mathbb{C}^{M_1 M_2}$ ▷ mean values

7: **for** $\mathbf{z} \in J_N$ **do**
8: $\quad \hat{p}_{\mathbf{z}} := \frac{1}{N^d} \sum_{\mathbf{x} \in X_N} f(\mathbf{x}) \mathrm{e}^{-2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{x}}$ ▷ fast computation by a FFT
9: **end for**
10: **for** $k = 1, ..., M_2$ **do**
11: $\quad$ **for** $\mathbf{z} \in J_N \setminus \{\mathbf{0}\}$ **do**
12: $\quad\quad \tilde{h}_{\mathbf{z},k} := \hat{p}_{\mathbf{z}} \cdot \frac{\Gamma\left(\frac{d}{2}\right) \mathcal{J}_{\frac{d-2}{2}}(2\pi |\mathbf{z}| r_k)}{(\pi |\mathbf{z}| r_k)^{\frac{d-2}{2}}}$ ▷ multiplier
13: $\quad$ **end for**
14: $\quad \tilde{h}_{\mathbf{0},k} := \hat{f}_{\mathbf{0}}$
15: $\quad$ **for** $j = 1, \ldots, M_1$ **do**
16: $\quad\quad g_{j,k} := \sum_{\mathbf{z} \in J_N} \tilde{h}_{\mathbf{z},k} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{y}_j}$ ▷ fast approximation by NFFTs [43]
17: $\quad$ **end for**
18: **end for**

Theorem 27. *Let* $\mathbf{f}$ *be the samples of a function* $f \colon \mathbb{T}^2 \to \mathbb{R}$ *at the regular grid* $X_N \subset \mathbb{T}^2$ *with discretization parameter* $N \in 2\mathbb{N}$. *Then the spherical mean values*

$$g_{j,k} = \mathcal{N}\mathcal{I}_N f(\mathbf{y}_j, r_k)$$

at center points $\mathbf{y}_1, \ldots, \mathbf{y}_{M_1} \in \mathbb{T}^2$, $M_1 \in \mathbb{N}$, *and radii* $r_1, \ldots, r_{M_2} \in (0, 1]$, $M_2 \in \mathbb{N}$, *can be computed by Algorithm 5. As in Theorem 26, if we approximate the sum in line 10 with fixed accuracy, then it exists a fast realization of Algorithm 5 with the numerical complexity* $\mathcal{O}\left(M_2\left(N^2 \log N + M_1\right)\right)$.

Proof. Similar steps as in the proof of Theorem 26 show the assertion. □

Algorithm 5 Discrete mean value operator $\mathcal{N}$, using NFFT [43], $d = 2$.

Input and **Output** as in Algorithm 4

for $\mathbf{z} \in J_N$ **do**
  $\hat{p}_{\mathbf{z}} := \frac{1}{N^2} \sum_{\mathbf{x} \in X} f(\mathbf{x}) \mathrm{e}^{-2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{x}}$ ▷ fast computation by a FFT
end for
for $k = 1, \ldots, M_2$ **do**
  for $\mathbf{z} \in J \setminus \{\mathbf{0}\}$ **do**
    $\tilde{h}_{\mathbf{z},k} := \hat{p}_{\mathbf{z}} \cdot \frac{\sin 2\pi r_k |\mathbf{z}|}{2\pi r_k |\mathbf{z}|}$ ▷ multiplier
  end for
  $\tilde{h}_{\mathbf{0},k} := \hat{p}_{\mathbf{0}}$
  for $j = 1, \ldots, M_1$ **do**
    $g_{j,k} := \sum_{\mathbf{z} \in J_N} \tilde{h}_{\mathbf{z},k} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{y}_j}$ ▷ fast approximation by NFFTs [43]
  end for
end for

In order to derive fast algorithms for the transpose operators, we set

$$\omega_{\mathbf{z},k} := \begin{cases} \frac{\Gamma\left(\frac{d}{2}\right) \mathcal{J}_{\frac{d-2}{2}}(2\pi |\mathbf{z}| r_k)}{(\pi |\mathbf{z}| r_k)^{\frac{d-2}{2}}} & \text{for} \quad \mathbf{z} \in J_N \setminus \{\mathbf{0}\}, \\ 1 & \text{for} \quad \mathbf{z} = \mathbf{0}, \end{cases}$$

for the operator $\mathcal{M}$ and

$$\omega_{\mathbf{z},k} := \begin{cases} \frac{\sin 2\pi r_k |\mathbf{z}|}{2\pi r_k |\mathbf{z}|} & \text{for } \mathbf{z} \in J_N \setminus \{\mathbf{0}\}, \\ 1 & \text{for } \mathbf{z} = 0, \end{cases}$$

for the operator $\mathcal{N}$, $k \in \{1, \ldots, M_2\}$, and obtain with the notation from Algorithm 4 and 5 the matrix-vector multiplication

$$g_{j,k} = \sum_{\mathbf{z} \in J_N} \tilde{h}_{\mathbf{z},k} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{y}_j} = \sum_{\mathbf{z} \in J_N} \hat{p}_{\mathbf{z}} \omega_{\mathbf{z},k} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{y}_j} = (\mathbf{A}\mathbf{f})_{(j,k)}$$

with

$$\mathbf{A} \in \mathbb{C}^{M_1 M_2 \times N^d}, \quad A_{(j,k),\mathbf{n}} := \sum_{\mathbf{z} \in J_N} \frac{\omega_{\mathbf{z},k}}{N^d} \mathrm{e}^{-2\pi \mathrm{i} \mathbf{z} \frac{2\mathbf{n}+1-N}{2N}} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{y}_j},$$

$(j,k) \in \{1, \ldots, M_1\} \times \{1, \ldots, M_2\}$, $\mathbf{n} \in J_N$. It follows

$$\left(\mathbf{M}^\top \mathbf{g}\right)_{\mathbf{n}} = \sum_{j,k=1}^{M_1, M_2} M_{\mathbf{n},(j,k)} g_{j,k} = \frac{1}{N^d} \sum_{\mathbf{z} \in J_N} \left(\sum_{k=1}^{M_2} \omega_{\mathbf{z},k} \left(\sum_{j=1}^{M_1} g_{j,k} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{y}_j} \right) \right) \mathrm{e}^{-2\pi \mathrm{i} \mathbf{z} \frac{2\mathbf{n}+1-N}{2N}}.$$

Similar to Algorithms 4 and 5, this formula can be implemented in three steps. We start with adjoint NFFTs for each radius to evaluate the sums over j, do multiplications and summations and close with one FFT for the computation of the outer sum over $\mathbf{z}$. All in all, this discretization of the transpose operators has the numerical complexity $\mathcal{O}\left(M_2 \left(N^d \log N + M_1\right)\right)$.

4.2.1.2 Butterfly sparse FFT based algorithm

The previous method is based on applying a d-dimensional NFFT for each radius. For certain settings we develop another variant for evaluating the sums in equation (4.4) and (4.5), where we use overall one $(d+1)$-dimensional sparse Fourier transform to compute the mean values for all given center points and radii. To be more precise, we define for a spatial dimension $d \in \mathbb{N}$ the following discrete subset

$$\tilde{J}_N := \left\{ (\mathbf{z}, \zeta) \in J_N \times \left[-\frac{\sqrt{d}N}{2}, \frac{\sqrt{d}N}{2} \right] : |\zeta| = |\mathbf{z}| \right\} \setminus \{\mathbf{0}\}$$

of a d-dimensional double cone in $\mathbb{R}^{d+1}$, see Figure 4.2 for dimension $d = 2$.

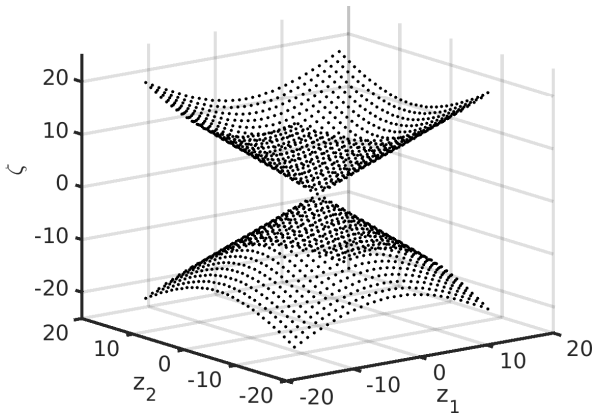


Figure 4.2: Distribution of the frequencies on a double cone.

Theorem 28. *Let a spatial dimension $d \in \{2, 3\}$, a discretization parameter $N \in 2\mathbb{N}$, a center point $\mathbf{y} \in \mathbb{T}^d$, a radius $r \in (0, 1]$, and a trigonometric polynomial*

$$p\colon \mathbb{T}^d \to \mathbb{C}, \quad p(\mathbf{x}) = \sum_{\mathbf{z} \in J_N} \hat{p}_{\mathbf{z}} \mathrm{e}^{2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{x}}, \quad \hat{p}_{\mathbf{z}} \in \mathbb{C},$$

be given. With the operator

$$\mathcal{A} := \begin{cases} \mathcal{N} & \text{for } d = 2, \quad \text{see Definition 17,} \\ \mathcal{M} & \text{for } d = 3, \quad \text{see Definition 15,} \end{cases} \tag{4.6}$$

the mean values of the polynomial p have the representation

$$\mathcal{A}p(\mathbf{y}, r) := \hat{p}_{\mathbf{0}} - \frac{\mathrm{i}}{4\pi r} \sum_{(\mathbf{z}, \zeta) \in \tilde{J}_N} \frac{\hat{p}_{\mathbf{z}}}{\zeta} \mathrm{e}^{2\pi \mathrm{i} (\mathbf{z}, \zeta) \cdot (\mathbf{y}, r)}. \tag{4.7}$$

Proof. First of all we observe that for all $\mathbf{z} \in J_N \setminus \{\mathbf{0}\}$ and $r > 0$ the identity

$$\frac{\Gamma\left(\frac{3}{2}\right) \mathcal{J}_{\frac{1}{2}}(2\pi|\mathbf{z}|r)}{\sqrt{\pi|\mathbf{z}|r}} = \frac{\sin 2\pi|\mathbf{z}|r}{2\pi|\mathbf{z}|r} = \frac{-\mathrm{i}}{4\pi|\mathbf{z}|r} \left(\mathrm{e}^{2\pi \mathrm{i}|\mathbf{z}|r} - \mathrm{e}^{-2\pi \mathrm{i}|\mathbf{z}|r} \right)$$

holds true. Hence, we have by Theorem 20

$$\begin{aligned}\mathcal{A}p(\mathbf{y},r) &= \sum_{\mathbf{z}\in J_N} \hat{p}_{\mathbf{z}} \cdot \left(\mathcal{A}\mathrm{e}^{2\pi\mathrm{i}\mathbf{z}\cdot(\cdot)}\right)(\mathbf{y},r) \\ &= \hat{p}_{\mathbf{0}} + \sum_{\mathbf{z}\in J_N\setminus\{\mathbf{0}\}} \hat{p}_{\mathbf{z}} \cdot \mathrm{e}^{2\pi\mathrm{i}\mathbf{z}\cdot\mathbf{y}} \cdot \frac{\Gamma\left(\frac{3}{2}\right)\mathcal{J}_{\frac{1}{2}}(2\pi|\mathbf{z}|r)}{\sqrt{\pi|\mathbf{z}|r}} \\ &= \hat{p}_{\mathbf{0}} + \frac{-\mathrm{i}}{4\pi r} \cdot \sum_{\mathbf{z}\in J_N\setminus\{\mathbf{0}\}} \frac{\hat{p}_{\mathbf{z}}}{|\mathbf{z}|} \cdot \mathrm{e}^{2\pi\mathrm{i}\mathbf{z}\cdot\mathbf{y}} \cdot (\mathrm{e}^{2\pi\mathrm{i}|\mathbf{z}|r} - \mathrm{e}^{-2\pi\mathrm{i}|\mathbf{z}|r}).\end{aligned}$$

□

For the purpose to derive a fast discretization based on Theorem 28, we consider center points $\mathbf{y}_j \in \mathbb{R}^d$, $j = 1,\dots,M_1$, on smooth $(d-1)$-dimensional submanifolds of $\mathbb{R}^d$, which is a common assumption in relevant applications. For example an arrangement on a cylinder or something similar is for practical reasons an adequate setting. For arbitrary radii $r_k \in (0,1]$, $k = 1,\dots,M_2$, the set of nodes

$$(\mathbf{y}_j, r_k) \in \mathbb{R}^{d+1}, \quad j = 1,\dots,M_1, \quad k = 1,\dots,M_2,$$

lies on a smooth d-dimensional submanifold of $\mathbb{R}^{d+1}$. Hence, equation (4.7) describes a $(d+1)$-dimensional Fourier transform with Fourier coefficients as well as evaluation nodes on smooth $(d+1)$-dimensional submanifolds. An efficient computation of such a sparse Fourier transform is given by the results in [63, 45].
The main idea consists of two ingredients. At first, we employ a low rank approximation

$$\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\nu}\cdot\boldsymbol{\tau}} \approx \sum_{l=1}^{R} a_l(\boldsymbol{\nu}) b_l(\boldsymbol{\tau}), \quad a_l(\boldsymbol{\nu}), b_l(\boldsymbol{\tau}) \in \mathbb{C},$$

of the Fourier kernel $\mathrm{e}^{2\pi\mathrm{i}(\cdot)\cdot(\cdot)} \colon \Lambda \times \Omega \to \mathbb{C}$ for restricted sets $\Lambda, \Omega \subset \mathbb{R}^{d+1}$, where the rank fulfills $R \approx \log^{d+1}\left(\frac{N}{\varepsilon}\right)$ under certain conditions on the widths of the spatial and frequency domains. In order to meet these conditions, the second ingredient is a dyadic subdivision of both domains, resulting in a particular divide and conquer strategy, the so-called butterfly scheme. The following Corollary 29 summarizes its application to the computation of mean values and an implementation for spatial dimension $d = 3$ is given by Algorithm 6.

Corollary 29. *Let* $\mathbf{f}$ *be the samples of the function* $f \in C\left(\mathbb{T}^d\right)$ *at the regular grid* $X_N \subset \mathbb{T}^2$ *with discretization parameter* $N \in 2\mathbb{N}$ *and spatial dimension* $d \in \{2,3\}$*, and* $\mathcal{A}$ *the operator from equation (4.6). Then, for fixed accuracy, the spherical mean values*

$$g_{j,k} = \mathcal{A}\mathcal{I}_N f(\mathbf{y}_j, r_k)$$

at center points $\mathbf{y}_1,\dots,\mathbf{y}_{M_1} \in \mathbb{T}^d$*,* $M_1 \in \mathbb{N}$*, located at a smooth* $(d-1)$*-dimensional submanifold of* $\mathbb{T}^d$*, and radii* $r_1,\dots,r_{M_2} \in (0,1]$*,* $M_1 \in \mathbb{N}$*, can be approximately computed with the numerical complexity* $\mathcal{O}\left((\log N)^{d+1} M_1 M_2 + (\log N)^{d+3} N^d\right)$*. In particular for* $M_1 = \mathcal{O}\left(N^{d-1}\right)$ *and* $M_2 = \mathcal{O}(N)$ *this leads to the almost linear complexity* $\mathcal{O}\left((\log N)^{d+3} N^d\right)$.

Proof. As in Theorem 26 we start with the computation of the discrete Fourier coefficients by a FFT. The essential complexity is given by the Fourier sum in equation (4.7). A fast computation is given by the algorithm described in [45], which results in the claimed complexity. Theorem 28 completes the proof. □

Algorithm 6 Discrete mean value operator $\mathcal{M}$, using sparse FFT [45], $d = 3$.

Input and **Output** as in Algorithm 4

for $\mathbf{z} \in J_N$ **do**
  $\hat{p}_{\mathbf{z}} := \frac{1}{N^3} \sum_{\mathbf{x} \in X_N} f(\mathbf{x}) \mathrm{e}^{-2\pi \mathrm{i} \mathbf{z} \cdot \mathbf{x}}$ ▷ fast computation by a FFT
end for
for $(\mathbf{z}, \zeta) \in \tilde{J}_N$ **do**
  $\hat{h}_{\mathbf{z},\zeta} := \frac{\hat{p}_{\mathbf{z}}}{\zeta}$ ▷ coefficients for sparse FFT
end for
for $j = 1, \ldots, M_1$ **do**
  for $k = 1, \ldots, M_2$ **do**
    $g_{j,k} := \sum_{(\mathbf{z},\zeta) \in \tilde{J}_N} \hat{h}_{\mathbf{z},\zeta} \cdot \mathrm{e}^{2\pi \mathrm{i} (\mathbf{z},\zeta) \cdot (\mathbf{y}_j, r_k)}$ ▷ fast approx. by a sparse FFT [45]
  end for
end for
$\mathbf{g} := \hat{p}_{\mathbf{0}} - \frac{\mathrm{i}}{4\pi r} \mathbf{g}$

4.2.1.3 Error estimates

The presented algorithms in the current Section 4.2.1 consist mainly of two steps. We replace the function by the trigonometric interpolant and utilize fast approximation algorithms for the evaluation of the nonequispaced Fourier transforms afterwards. In the following, we assume a sufficient high accuracy of the second step, such that the approximation error from applying the NFFT and butterfly sparse FFT is negligible compared to the interpolation error. For this reason, we continue with analyzing the convergence rates of the discretization error, which is caused by the approximation of the function by the trigonometric interpolant. We start with a bound for the operator norm of $\mathcal{M}$.

Lemma 30. *For a given spatial dimension $d \in \mathbb{N} \setminus \{1\}$, the spherical mean value operator $\mathcal{M}$ satisfies*

$$\|\mathcal{M} f\|_{L^\infty(\mathbb{T}^d \times [0,1])} \leq \|f\|_{L^\infty(\mathbb{T}^d)}$$

for $f \in L^\infty\left(\mathbb{T}^d\right)$ and

$$\|\mathcal{M} f\|_{L^p(\mathbb{T}^d \times [0,1], \mathrm{d}\mathbf{y} r^{d-1} \mathrm{d}r)} \leq d^{-\frac{1}{p}} \|f\|_{L^p(\mathbb{T}^d)}$$

for $1 \leq p < \infty$ and $f \in L^p\left(\mathbb{T}^d\right)$.

Proof. We use abbreviatory $L^p(\mathbb{T}^d \times [0,1]) := L^p(\mathbb{T}^d \times [0,1], \mathrm{d}\mathbf{y} r^{d-1}\mathrm{d}r)$, $p \in [1,\infty)$. For $p = \infty$ we have

$$\begin{aligned}\|\mathcal{M}f\|_{L^\infty(\mathbb{T}^d\times[0,1])} &= \sup_{r\in[0,1]} \sup_{\mathbf{y}\in\mathbb{T}^d} \left| \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} f(\mathbf{y}+r\boldsymbol{\xi}) \mathrm{d}\sigma(\boldsymbol{\xi}) \right| \\ &\leq \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma(\boldsymbol{\xi}) \sup_{\mathbf{y}\in\mathbb{T}^d} |f(\mathbf{y})|,\end{aligned}$$

for $p = 1$ the estimate

$$\begin{aligned}\|\mathcal{M}f\|_{L^1(\mathbb{T}^d\times[0,1])} &= \int_0^1 \int_{\mathbb{T}^d} \frac{1}{\omega_{d-1}} \left| \int_{\mathbb{S}^{d-1}} f(\mathbf{y}+r\boldsymbol{\xi}) \mathrm{d}\sigma(\boldsymbol{\xi}) \right| \mathrm{d}\mathbf{y} r^{d-1}\mathrm{d}r \\ &\leq \int_0^1 \int_{\mathbb{T}^d} \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} |f(\mathbf{y}+r\boldsymbol{\xi})| \, \mathrm{d}\sigma(\boldsymbol{\xi}) \mathrm{d}\mathbf{y} r^{d-1}\mathrm{d}r \\ &\leq \frac{1}{\omega_{d-1}} \int_0^1 r^{d-1}\mathrm{d}r \int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma(\boldsymbol{\xi}) \sup_{\mathbf{y}\in\mathbb{T}^d} |f(\mathbf{y})| = \frac{1}{d}\|f\|_{L^1(\mathbb{T}^d)}\end{aligned}$$

is valid, and for $1 < p < \infty$ and $q := \left(1 - \frac{1}{p}\right)^{-1}$, the Hölder inequality yields

$$\begin{aligned}\|\mathcal{M}f\|^p_{L^p(\mathbb{T}^d\times[0,1])} &= \int_0^1 \int_{\mathbb{T}^d} \frac{1}{\omega_{d-1}^p} \left| \int_{\mathbb{S}^{d-1}} f(\mathbf{y}+r\boldsymbol{\xi}) \mathrm{d}\sigma(\boldsymbol{\xi}) \right|^p \mathrm{d}\mathbf{y} r^{d-1}\mathrm{d}r \\ &\leq \int_0^1 \int_{\mathbb{T}^d} \frac{1}{\omega_{d-1}^p} \left(\int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma(\boldsymbol{\xi}) \right)^{\frac{p}{q}} \int_{\mathbb{S}^{d-1}} |f(\mathbf{y}+r\boldsymbol{\xi})|^p \mathrm{d}\sigma(\boldsymbol{\xi}) \mathrm{d}\mathbf{y} r^{d-1}\mathrm{d}r \\ &= \int_0^1 \omega_{d-1}^{\frac{p}{q}-p} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{T}^d} |f(\mathbf{y}+r\boldsymbol{\xi})|^p \mathrm{d}\mathbf{y} \mathrm{d}\sigma(\boldsymbol{\xi}) r^{d-1}\mathrm{d}r \\ &= \|f\|^p_{L^p(\mathbb{T}^d)} \int_0^1 r^{d-1}\mathrm{d}r = \frac{1}{d}\|f\|^p_{L^p(\mathbb{T}^d)}.\end{aligned}$$

□

An appropriate version of Lemma 30 for the mean value operator $\mathcal{N}$ is given by the following Lemma 31.

Lemma 31. *The mean value operator $\mathcal{N}$ satisfies*

$$\|\mathcal{N}f\|_{L^\infty(\mathbb{T}^2\times[0,1])} \leq \|f\|_{L^\infty(\mathbb{T}^2)},$$

for $f \in L^\infty(\mathbb{T}^2)$ and

$$\|\mathcal{N}f\|_{L^p(\mathbb{T}^2\times[0,1])} \leq 2^{-\frac{1}{p}}\|f\|_{L^p(\mathbb{T}^2)}$$

for $1 \leq p < \infty$ and $f \in L^p(\mathbb{T}^2)$.

Proof. The assertion is proven similar to Lemma 30. Again, we use abbreviatory $L^p(\mathbb{T}^2 \times [0,1]) := L^p(\mathbb{T}^2 \times [0,1], \mathrm{d}\mathbf{y} r \mathrm{d}r)$, $p \in [1,\infty)$. For $p = 1$ and $p = \infty$ the estimate

follows from the definition, and for $1 < p < \infty$, we apply the Minkowski inequality instead of the Hölder inequality and appropriate computations lead to

$$\begin{aligned}\|\mathcal{N}f\|^p_{L^p(\mathbb{T}^2\times[0,1])} &= \left(\int_{\mathbb{T}^2}\int_0^1 \left|\frac{1}{2\pi}\int_0^1\int_0^{2\pi} f\left(\mathbf{y}+r\rho\begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix}\right)\frac{\mathrm{d}\varphi\rho\mathrm{d}\rho}{\sqrt{1-\rho^2}}\right|^p \mathrm{d}\mathbf{y}r\mathrm{d}r\right)^{\frac{1}{p}}\\ &\leq \frac{1}{2\pi}\int_0^1\int_0^{2\pi}\left(\int_{\mathbb{T}^2}\int_0^1\left|f\left(\mathbf{y}+r\rho\begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix}\right)\right|^p \mathrm{d}\mathbf{y}r\mathrm{d}r\right)^{\frac{1}{p}}\frac{\mathrm{d}\varphi\rho\mathrm{d}\rho}{\sqrt{1-\rho^2}}\\ &= \frac{1}{2\pi}\|f\|_{L^p}\left(\int_0^1 r\mathrm{d}r\right)^{\frac{1}{p}}\left(\int_0^1\frac{\rho\mathrm{d}\rho}{\sqrt{1-\rho^2}}\right)\left(\int_0^{2\pi}\mathrm{d}\varphi\right) = 2^{-\frac{1}{p}}\|f\|_{L^p(\mathbb{T}^2)}.\end{aligned}$$

□

Now we can bound the discretization errors of the mean value operators by the interpolation errors.

Theorem 32. *Let a spatial dimension $d \in \mathbb{N}\setminus\{1\}$, a discretiation parameter $N \in 2\mathbb{N}$, $N \geq 4\sqrt{d}$, and a continuous function $f \in C(\mathbb{T}^d)$ be given. With*

$$C_1(s,d) := \left(\frac{2^{4s-d+1}}{2s-d}\omega_{d-1}\right)^{\frac{1}{2}},\ s > \frac{d}{2}, \quad C_2(s,d) := \frac{2^{2s-d}}{s-d}\omega_{d-1},\ s > d,$$

$\mathcal{A} := \mathcal{M}$ for $d \in \mathbb{N}\setminus\{2\}$ and $\mathcal{A} \in \{\mathcal{M},\mathcal{N}\}$ for $d = 2$, the following estimates are valid. For $s \geq 0$ and $\|f\|_{s,1} < \infty$

$$\|\mathcal{A}f(\mathbf{y},r) - \mathcal{A}\mathcal{I}_N f(\mathbf{y},r)\|_\infty \leq 2^{-s+1}\|f\|_{s,1}N^{-s},$$

for $s > \frac{d}{2}$ and $\|f\|_{s,2} < \infty$

$$\|\mathcal{A}f(\mathbf{y},r) - \mathcal{A}\mathcal{I}_N f(\mathbf{y},r)\|_\infty \leq C_1(s,d)\|f\|_{s,2}N^{\frac{d}{2}-s},$$

and for $s > d$ and $\|f\|_{s,\infty} < \infty$

$$\|\mathcal{A}f(\mathbf{y},r) - \mathcal{A}\mathcal{I}_N f(\mathbf{y},r)\|_\infty \leq C_2(s,d)\|f\|_{s,\infty}N^{d-s}.$$

Proof. For fixed $\mathbf{y} \in \mathbb{T}^d$ and $r \in [0,1]$, it follows

$$\begin{aligned}|\mathcal{M}f(\mathbf{y},r) - \mathcal{M}\mathcal{I}_N f(\mathbf{y},r)| &= \frac{1}{\omega_{d-1}}\left|\int_{\mathbb{S}^{d-1}} f(\mathbf{y}+r\boldsymbol{\xi}) - \mathcal{I}_N f(\mathbf{y}+r\boldsymbol{\xi})\mathrm{d}\sigma(\boldsymbol{\xi})\right|\\ &= \frac{\|f-\mathcal{I}_N f\|_\infty}{\omega_{d-1}}\int_{\mathbb{S}^{d-1}}\mathrm{d}\sigma(\boldsymbol{\xi}) = \|f-\mathcal{I}_N f\|_\infty\end{aligned}$$

and for $d = 2$ also

$$|\mathcal{N}f(\mathbf{y},r) - \mathcal{N}\mathcal{I}_N f(\mathbf{y},r)| \leq \|f-\mathcal{I}_N f\|_\infty.$$

Lemma 13 completes the proof. □

Due to the smoothing effect of integral operators, we expect, that the proven convergence rates in Theorem 32 are not optimal. If we replace the function by the Fourier partial sum instead of the trigonometric interpolant, then we can show a faster decrease of the approximation error. For this reason, we will also use this variant in the numerical experiments for a more precise evaluation of the results and state an appropriate error estimate. However, this approximation is not applicable in general, because it is hard to compute exactly the Fourier coefficients of an arbitrary function.

Theorem 33. *Let a spatial dimension $d \in \mathbb{N}\setminus\{1\}$, a discretization parameter $N \in 2\mathbb{N}$, $N \geq 4\sqrt{d}$, a center point $\mathbf{y} \in \mathbb{T}^d$, a radius $r \in (0,1]$, and a continuous function $f \in C(\mathbb{T}^d)$ be given. If we define*

$$C_1(s,d) := \frac{\Gamma\left(\frac{d}{2}\right) 2^{s+\frac{d}{2}-\frac{1}{2}}}{\pi^{\frac{d}{2}}}, \ s \geq 0, \quad C_2(s,d) := \frac{\Gamma\left(\frac{d}{2}\right) 2^{2s+\frac{d}{2}-\frac{3}{2}}\sqrt{\omega_{d-1}}}{\pi^{\frac{d}{2}}\sqrt{2s-1}}, \ s > \frac{1}{2},$$

and

$$C_3(s,d) := \frac{\Gamma\left(\frac{d}{2}\right) 2^{2s-2}\omega_{d-1}}{\pi^{\frac{d}{2}}\left(s-\frac{d}{2}-\frac{1}{2}\right)}, \ s > \frac{d}{2}+\frac{1}{2},$$

then there exist constants $C_4(f,s,d,r), C_5(f,s,d,r), C_6(f,s,d,r) > 0$ such that the following estimates are valid. For $s \geq 0$ and $\|f\|_{s,1} < \infty$

$$|\mathcal{M}f(\mathbf{y},r) - \mathcal{M}\mathcal{S}_N f(\mathbf{y},r)| \leq C_1(s,d) r^{-\frac{d}{2}+\frac{1}{2}} \|f\|_{s,1} N^{-s-\frac{d}{2}+\frac{1}{2}} + C_4(f,s,d,r) N^{-s-\frac{d}{2}-\frac{1}{2}},$$

for $s > \frac{1}{2}$ and $\|f\|_{s,2} < \infty$

$$|\mathcal{M}f(\mathbf{y},r) - \mathcal{M}\mathcal{S}_N f(\mathbf{y},r)| \leq C_2(s,d) r^{-\frac{d}{2}+\frac{1}{2}} \|f\|_{s,2} N^{-s+\frac{1}{2}} + C_5(f,s,d,r) N^{-s-\frac{1}{2}},$$

and for $s > \frac{d}{2}+\frac{1}{2}$ and $\|f\|_{s,\infty} < \infty$

$$|\mathcal{M}f(\mathbf{y},r) - \mathcal{M}\mathcal{S}_N f(\mathbf{y},r)| \leq C_3(s,d) r^{-\frac{d}{2}+\frac{1}{2}} \|f\|_{s,\infty} N^{-s+\frac{d}{2}+\frac{1}{2}} + C_6(f,s,d,r) N^{-s+\frac{d}{2}-\frac{1}{2}}.$$

Proof. Let $x_0 > 0$ be fixed. From the asymptotic expansion of the Bessel functions [1, p. 9.2.1] we know that there is a constant $C_7(d) > 0$ such that for all $x \geq x_0$ we have

$$\left| \mathcal{J}_{\frac{d-2}{2}}(x) - \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{d-3}{4}\pi\right) \right| \leq C_7(d) x^{-\frac{3}{2}}, \quad C_7(d) > 0$$

and with

$$x_0 := \min_{\mathbf{z}\in J_N} 2\pi|\mathbf{z}|r = 2\pi\frac{N}{2}r = \pi N r > 0,$$

it follows

$$
\begin{aligned}
|\mathcal{M}f(\mathbf{y},r) - \mathcal{M}\mathcal{S}_N f(\mathbf{y},r)| &\leq \left| \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} c_{\mathbf{z}}(f)\mathcal{M}\mathrm{e}^{2\pi\mathrm{i}\mathbf{z}\cdot(\cdot)}(\mathbf{y},r) \right| \\
&= \left| \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} c_{\mathbf{z}}(f) \frac{\Gamma\left(\frac{d}{2}\right) \mathcal{J}_{\frac{d-2}{2}}(2\pi|\mathbf{z}|r)}{(\pi|\mathbf{z}|r)^{\frac{d-2}{2}}} \right| \\
&\leq \frac{\Gamma\left(\frac{d}{2}\right)}{(\pi r)^{\frac{d-2}{2}}} \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} \frac{|c_{\mathbf{z}}(f)|}{|\mathbf{z}|^{\frac{d-2}{2}}} \left(\sqrt{\frac{1}{\pi^2|\mathbf{z}|r}} + \frac{C_7(d)}{(2\pi|\mathbf{z}|r)^{\frac{3}{2}}} \right) \\
&\leq \frac{\Gamma\left(\frac{d}{2}\right)}{\pi^{\frac{d}{2}} r^{\frac{d-1}{2}}} \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |c_{\mathbf{z}}(f)| \left(|\mathbf{z}|^{\frac{1-d}{2}} + C_8(d,r)|\mathbf{z}|^{-\frac{d+1}{2}} \right)
\end{aligned}
$$

with a constant $C_8(d,r) > 0$. With the methods from the proof of Lemma 13 we can show upper bounds of the sums. For example the estimate

$$
\begin{aligned}
\sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |c_{\mathbf{z}}(f)||\mathbf{z}|^{\frac{1-d}{2}} &= \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |c_{\mathbf{z}}(f)| \cdot |\mathbf{z}|^s \cdot |\mathbf{z}|^{\frac{1-d}{2}-s} \\
&\leq \max_{z\in\mathbb{Z}^d\setminus J_N} |\mathbf{z}|^{\frac{1-d}{2}-s} \sum_{\mathbf{z}\in\mathbb{Z}^d\setminus J_N} |c_{\mathbf{z}}(f)| \cdot (1+|\mathbf{z}|)^s \\
&= \left(\frac{N}{2}\right)^{\frac{1-d}{2}-s} \|f\|_{s,1}
\end{aligned}
$$

is valid. □

We summarize the previous results for the convergence rates in Table 4.1.

Assumptions	$\lvert\mathcal{A}f(\mathbf{y},r) - \mathcal{A}\mathcal{I}_N f(\mathbf{y},r)\rvert$	$\lvert\mathcal{M}f(\mathbf{y},r) - \mathcal{M}\mathcal{S}_N f(\mathbf{y},r)\rvert$
$\|f\|_{s,1} < \infty, \quad s \geq 0$	$\mathcal{O}\left(N^{-s}\right)$	$\mathcal{O}\left(N^{-s-\frac{d}{2}+\frac{1}{2}}\right)$
$\|f\|_{s,2} < \infty, \quad s \geq \frac{d}{2}$	$\mathcal{O}\left(N^{-s+\frac{d}{2}}\right)$	$\mathcal{O}\left(N^{-s+\frac{1}{2}}\right)$
$\|f\|_{s,\infty} < \infty, \quad s \geq d$	$\mathcal{O}\left(N^{-s+d}\right)$	$\mathcal{O}\left(N^{-s+\frac{d}{2}+\frac{1}{2}}\right)$

Table 4.1: Convergence rates from Theorems 32 and 33 for fixed center point $\mathbf{y} \in \mathbb{T}^d$ and radius $r > 0$. The operator $\mathcal{A}$ denotes $\mathcal{A} \in \{\mathcal{M}, \mathcal{N}\}$ for spatial dimension $d = 2$ and $\mathcal{A} := \mathcal{M}$ for $d \in \mathbb{N} \setminus \{1, 2\}$.

4.2.2 Polar frequency grids

For spatial dimension $d = 2$, we restrict to the special case when the center points

$$
\mathbf{y}_j := \frac{1}{2}\left(\cos\frac{2\pi j}{M_1}, \sin\frac{2\pi j}{M_1}\right)^\top, \quad j = 0, \dots M_1 - 1, \quad M_1 \in \mathbb{N}, \tag{4.8}
$$

discretize a circle and the radii

$$r_k := \frac{k}{M_2}, \quad k = 0, \dots, M_2 - 1, \quad M_2 \in \mathbb{N}, \tag{4.9}$$

are equally spaced. Moreover, the function $f\colon \mathbb{R}^2 \to \mathbb{C}$ is assumed to be given by the coefficients $\hat{f}_{s,l} \in \mathbb{C}$ in the nonharmonic Fourier series

$$f(\mathbf{x}) = \sum_{s=0}^{M_1-1} \sum_{l=0}^{\lceil M_3\sqrt{2}\rceil-1} \hat{f}_{s,l} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{\xi}_{s,l} \cdot \mathbf{x}} \tag{4.10}$$

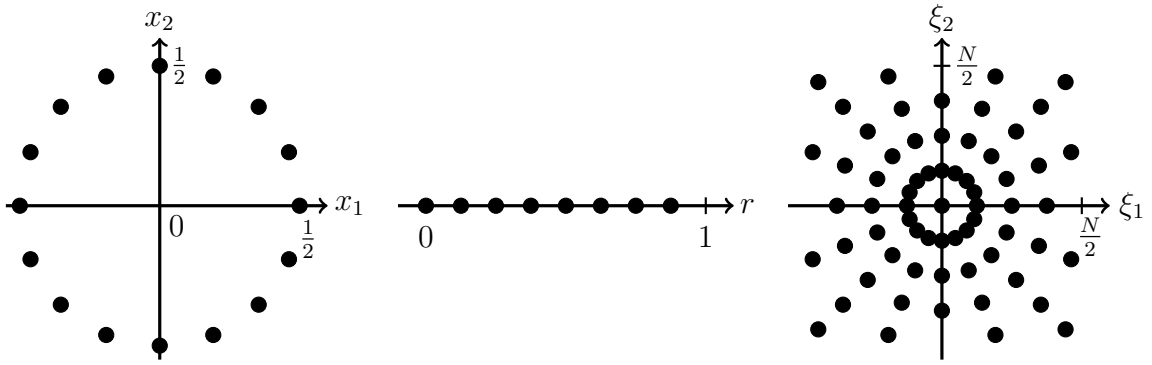
with frequencies on the polar grid

$$J'_N := \left\{ \boldsymbol{\xi}_{s,l} := \frac{lN}{2M_3} \left(\cos \frac{2\pi s}{M_1}, \sin \frac{2\pi s}{M_1} \right)^\top : \begin{matrix} s = 0, \dots, M_1 - 1 \\ l = 0, \dots, \lceil M_3\sqrt{2} \rceil - 1 \end{matrix} \right\} \tag{4.11}$$

and

$$\hat{f}_{s,l} := 0 \quad \text{for} \quad \boldsymbol{\xi}_{s,l} \notin \left(-\frac{N}{2}, \frac{N}{2} \right)^2,$$

see also Figure 4.3. The mean values for the operator $\mathcal{N}$ can be computed efficiently in the following way.



(a) Center points arranged on a circle, $M_1 = 16$.

(b) Equispaced radii, $M_2 = 8$.

(c) Frequencies on a polar grid, $M_1 = 16$, $M_3 = 4$.

Figure 4.3: Setting for a fast computation of mean values of functions, which are given by Fourier sums with frequencies on a polar grid.

Theorem 34. *Let parameters $M_1, M_2, M_3 \in \mathbb{N}$ and coefficients*

$$\hat{f}_{s,l} \in \mathbb{C}, \quad s = 0, \dots, M_1 - 1, \quad l = 0, \dots, \left\lceil M_3\sqrt{2} \right\rceil - 1,$$

of the nonharmonic Fourier series from a function $f\colon \mathbb{R}^2 \to \mathbb{C}$, see equation (4.10), with frequencies on a polar grid J'_N, see equation (4.11), be given. Using the notations from equations (4.8) and (4.9), the spherical means

$$g_{j,k} = \mathcal{N} f(\mathbf{y}_j, r_k), \quad j = 0, \dots, M_1 - 1, \quad k = 0, \dots, M_2 - 1,$$

can be computed by Algorithm 7. For an evaluation of the sum in line 14 with a fixed accuracy, the algorithm can be realized with the numerical complexity

$$\mathcal{O}\left(M_1\left(M_3 \log M_1 + M_2 \log M_2\right)\right).$$

In particular for $M_1 = \mathcal{O}(N)$, $M_2 = \mathcal{O}(N)$, and $M_3 = \mathcal{O}(N)$, such a realization has the almost linear complexity $\mathcal{O}(N^2 \log N)$.

Proof. With $M_4 := \lceil M_3\sqrt{2} \rceil$ and Corollary 23, for $k \neq 0$ it follows

$$\mathcal{N}f(\mathbf{y}_j, r_k) = \sum_{s=0}^{M_1-1} \hat{f}_{s,0} + \frac{M_2 M_3}{\pi k N} \sum_{l=1}^{M_4-1} \left(\sum_{s=0}^{M_1-1} \frac{\hat{f}_{s,l}}{l} \mathrm{e}^{\frac{\pi i l N}{2M_3}} \cos\tfrac{2\pi(s-j)}{M_1} \right) \sin \frac{\pi l k N}{M_2 M_3}.$$

For each $l = 1, \ldots, M_4 - 1$, the inner sum is a one-dimensional cyclic convolution of length M_1, which can be realized by FFTs, multiplications of the discrete Fourier coefficients and inverse FFTs. This first step ends in the total complexity

$$\mathcal{O}\left(M_4 M_1 \log M_1\right). \tag{4.12}$$

For each $j = 0, \ldots, M_1 - 1$, the outer sums are one-dimensional discrete sine transforms, which can for instance be computed by adjoint nonequispaced fast Fourier transforms with bandwith M_2 and $2M_4$ nodes. Denoting by $\varepsilon > 0$ the desired accuracy of the NFFT, this second step can be computed with the total complexity

$$\mathcal{O}\left(M_1\left(M_2 \log M_2 + 2|\log \varepsilon| M_4\right)\right). \tag{4.13}$$

Combining equations (4.12) and (4.13) finishes the proof. □

Remark 35. *The original task was the computation of spherical means from given function samples $\mathbf{f}$, see equation (4.2). In contrast, Algorithm 7 requires Fourier coefficients $\hat{\mathbf{f}}$ as input. For the utilizing of this approach in iterative reconstruction methods, we propose to iterate in the Fourier domain and to compute once the sampels from the reconstructed Fourier coefficients at the end.*
However, even in the discrete case, it is a challenging task to give conditions to the number of angles M_1 and radii M_3, which guarantees a representation as in equation (4.10). Numerical experiments indicate at least approximations with arbitrary precisions for sufficiently large numbers M_1 and M_3. This fits to typical applications, where we have to deal with noisy data and other circumstances, such that exact reconstructions are no realistic goal.

4.2.2.1 Frame properties

Remark 35 motivates the use of polar frequencies with a numerical argument. Another promising fact is given by Theorem 40, which states stability in a certain sense, if the frequency grid has a sufficient high density.

Definition 36 (δ-*dense*)**.** For $\delta > 0$, a given set $\Omega = \{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_M\} \subset \mathbb{T}^2$, $M \in \mathbb{N}$, is called δ-dense in $\mathbb{T}^2$, if

$$\sup_{\mathbf{x} \in \mathbb{T}^2} \inf_{\boldsymbol{\zeta} \in \Omega} |\mathbf{x} - \boldsymbol{\zeta}| \leq \delta$$

is fulfilled. □

Algorithm 7 Discrete mean value operator $\mathcal{N}$, using a polar frequency grid, $d = 2$.

Input

$M_1, M_2, M_3 \in \mathbb{N}$ ▷ parameters

$\hat{f}_{s,l} \in \mathbb{C}^{M_1 M_4}$ ▷ coefficients

Output

$\mathbf{g} \in \mathbb{C}^{M_1 M_2}$ ▷ spherical means

$M_4 := \lceil M_3 \sqrt{2} \rceil$

$\mathbf{s} := (0, \dots, M_1 - 1)^\top$

for $l = 1, \dots, M_4 - 1$ **do**

  $\mathbf{a} := \frac{\hat{f}_{\mathbf{s},l}}{l}$

  $\mathbf{b} := \mathrm{e}^{\frac{\pi \mathrm{i} l N}{2M_3}} \cos \frac{2\pi \mathbf{s}}{M_1}$

  $h_{l,\mathbf{s}} := \mathtt{iFFT}\left(\mathtt{FFT}(\mathbf{a}) \odot \mathtt{FFT}(\mathbf{b})\right)^\top$ ▷ convolution

end for

$g_{\mathbf{s},0} := \mathbf{0}$

for $j = 0, \dots, M_1 - 1$ **do**

  for $k = 1, \dots, M_2 - 1$ **do**

    $g_{j,k} := \frac{M_2 M_3}{\pi k N} \sum_{l=1}^{M_4 - 1} h_{l,j} \sin \frac{\pi l N k}{M_2 M_3}$ ▷ discrete sine transform

  end for

end for

$\mathbf{g} := \mathbf{g} + \sum_{s=0}^{M_1 - 1} \hat{f}_{s,0}$

Definition 37 (*Voronoi regions and weights*)**.** Let a set $\Omega = \{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_M\} \subset \mathbb{T}^2$, $M \in \mathbb{N}$, be given. For $s \in \{1, \dots, M\}$, the Voronoi regions V_s of $\boldsymbol{\zeta}_s \in \Omega$ and the associated weights $w_s \geq 0$ are defined by

$$V_s := \left\{\mathbf{x} \in \mathbb{T}^2 : |\mathbf{x} - \boldsymbol{\zeta}_s| \leq |\mathbf{x} - \boldsymbol{\eta}| \text{ for all } \boldsymbol{\eta} \in \Omega \setminus \{\boldsymbol{\zeta}_s\}\right\}$$

and

$$w_s := \int_{\mathbb{T}^2} \chi_{V_s}(\mathbf{x}) \mathrm{d}\mathbf{x}.$$

We will denote by $\mathbf{W}$ the collection of the weights to all elements $\boldsymbol{\zeta}_1 \dots, \boldsymbol{\zeta}_M \in \Omega$ in the diagonal matrix

$$\mathbf{W} := \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix} \in \mathbb{R}^{M \times M}.$$

□

Remark 38. *For reasons of simplicity we use the Voronoi weights from Definition 37 for the following procedure. Similar results can also be obtained by weights constructed from another partition $V_1, \dots, V_M \subset \mathbb{T}^2$ of $\mathbb{T}^2$ with $\bigcup_{s=1}^M V_s = \mathbb{T}^2$ and $\boldsymbol{\zeta}_s \in V_s$ for all $s \in \{1, \dots, M\}$.*

Definition 39 (*Weighted norm*)**.** The weighted norm $\|\cdot\|_{\mathbf{W}} \colon \mathbb{C}^M \to [0, \infty]$ for given weights $w_1, \dots w_M > 0$, $M \in \mathbb{N}$, is defined by

$$\|\mathbf{f}\|_{\mathbf{W}} := \left(\sum_{s=1}^{M} w_s |f_s|^2\right)^{\frac{1}{2}}.$$

□

For a sufficiently large number of frequencies, the samples and related polar Fourier coefficients are norm-equivalent in the following sense.

Theorem 40 ([23, Sec. 4.3],[2, Sec. 2.2]). *Let $\delta > 0$, a vector $\mathbf{f} = (f_{\mathbf{z}})_{\mathbf{z}\in J_N} \in \mathbb{C}^{N^2}$, $N \in 2\mathbb{N}$, and a δ-dense set $\Omega = \{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_M\} \subset \mathbb{T}^2$ with associated Voronoi weights $w_1, \dots, w_M \geq 0$, $M \in \mathbb{N}$, be given. With*

$$\tilde{f}\colon \mathbb{T}^2 \to \mathbb{C}, \quad \tilde{f}(\mathbf{x}) := \sum_{\mathbf{z}\in J_N} f_{\mathbf{z}} \mathrm{e}^{-2\pi \mathrm{i}\mathbf{x}\cdot\mathbf{z}}, \quad \textit{and} \quad \tilde{\mathbf{f}} := \left(\tilde{f}\left(\boldsymbol{\zeta}_s\right)\right)_{s\in\{1,\dots,M\}},$$

it follows that

$$\left| \|\mathbf{f}\|_2 - \left\|\tilde{\mathbf{f}}\right\|_{\mathbf{w}} \right| \leq \left(\mathrm{e}^{\pi N\delta\sqrt{2}} - 1\right) \|\mathbf{f}\|_2.$$

Proof. The main steps of this proof are given by [23, Sec. 4.3] and [2, Sec. 2.2], where general dimensions $d \in \mathbb{N}$ are considered. We adapt this ideas to our two-dimensional situation, but remark, that we use nonessential simplifications concerning the special case $d = 2$.
For abbreviation, we define

$$\psi_s\colon \mathbb{T}^2 \to \{1, 0\}, \quad \psi_s := \chi_{V_s}, \quad s \in \{1, \dots, M\},$$

and with

$$\begin{aligned}
\left\|\tilde{\mathbf{f}}\right\|_{\mathbf{w}}^2 &= \sum_{s=1}^{M} w_s \left|\tilde{f}_s\right|^2 = \sum_{s=1}^{M} \left|\tilde{f}_s\right|^2 \int_{\mathbb{T}^2} \psi_s(\mathbf{x})\mathrm{d}\mathbf{x} = \int_{\mathbb{T}^2} \sum_{s=1}^{M} \left|\tilde{f}_s\psi_s(\mathbf{x})\right|^2 \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{T}^2} \sum_{s=1}^{M} \left|\tilde{f}_s\psi_s(\mathbf{x}) \sum_{l=1}^{M} \psi_l(\mathbf{x})\right|^2 \mathrm{d}\mathbf{x} = \int_{\mathbb{T}^2} \left|\sum_{l=1}^{M} \tilde{f}_l\psi_l(\mathbf{x})\right|^2 \mathrm{d}\mathbf{x} \\
&= \left\|\sum_{s=1}^{M} \tilde{f}_s\psi_s\right\|_{L^2(\mathbb{T}^2)}^2,
\end{aligned}$$

the Parseval identity, and the reverse triangle inequality, it follows

$$\left| \|\mathbf{f}\|_2 - \left\|\tilde{\mathbf{f}}\right\|_{\mathbf{w}} \right| \leq \left\|\tilde{f} - \sum_{s=1}^{M} \tilde{f}_s\psi_s\right\|_{L^2(\mathbb{T}^2)}. \tag{4.14}$$

Fixing $s \in \{1, \dots, M\}$ and $\mathbf{x} \in V_s$, using the Taylor series around $\mathbf{x}$, and applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\left|\tilde{f}(\mathbf{x}) - \tilde{f}\left(\boldsymbol{\zeta}_s\right)\right|^2 &= \left| \sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^2\setminus\{\mathbf{0}\}} \frac{(\mathbf{x}-\boldsymbol{\zeta}_s)^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}}\tilde{f}(\mathbf{x}) \right|^2 \\
&\leq \left(\sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^2\setminus\{\mathbf{0}\}} \frac{\sqrt{c^{\|\boldsymbol{\alpha}\|_1}}\,|(\mathbf{x}-\boldsymbol{\zeta}_s)^{\boldsymbol{\alpha}}|}{\sqrt{\boldsymbol{\alpha}!}} \cdot \frac{\sqrt{c^{-\|\boldsymbol{\alpha}\|_1}}|D^{\boldsymbol{\alpha}}\tilde{f}(\mathbf{x})|}{\sqrt{\boldsymbol{\alpha}!}} \right)^2 \\
&\leq \left(\sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^2\setminus\{\mathbf{0}\}} \frac{c^{\|\boldsymbol{\alpha}\|_1}\,(\mathbf{x}-\boldsymbol{\zeta}_s)^{2\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!} \right) \left(\sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^2\setminus\{\mathbf{0}\}} \frac{c^{-\|\boldsymbol{\alpha}\|_1}|D^{\boldsymbol{\alpha}}\tilde{f}(\mathbf{x})|^2}{\boldsymbol{\alpha}!} \right)
\end{aligned} \tag{4.15}$$

with $c := \frac{\pi N\sqrt{2}}{\delta}$. We continue with separate estimates of the two sums in equation (4.15). The binomial theorem yields

$$\sum_{\|\boldsymbol{\alpha}\|_1=k} \frac{k!}{\boldsymbol{\alpha}!}(\mathbf{x}-\boldsymbol{\zeta}_s)^{2\boldsymbol{\alpha}} = |\mathbf{x}-\boldsymbol{\zeta}_s|^{2k}$$

and it follows

$$\begin{aligned}\sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^2\setminus\{\mathbf{0}\}} \frac{c^{\|\boldsymbol{\alpha}\|_1}(\mathbf{x}-\boldsymbol{\zeta}_s)^{2\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!} &= -1+\sum_{k=0}^{\infty}\frac{c^k}{k!}\sum_{\|\boldsymbol{\alpha}\|_1=k}\frac{k!}{\boldsymbol{\alpha}!}(\mathbf{x}-\boldsymbol{\zeta}_s)^{2\boldsymbol{\alpha}} \\ &\le -1+\sum_{k=0}^{\infty}\frac{(\delta^2 c)^k}{k!} = \mathrm{e}^{\delta^2 c}-1 \end{aligned} \tag{4.16}$$

and furthermore, the Parseval identity and again the binomial theorem lead to

$$\begin{aligned}\int_{\mathbb{T}^2}\sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^2\setminus\{\mathbf{0}\}}\frac{c^{-\|\boldsymbol{\alpha}\|_1}\left|D^{\boldsymbol{\alpha}}\tilde{f}(\mathbf{x})\right|^2}{\boldsymbol{\alpha}!}\mathrm{d}\mathbf{x} &= \sum_{k=1}^{\infty}\frac{c^{-k}}{k!}\sum_{\|\boldsymbol{\alpha}\|_1=k}\frac{k!}{\boldsymbol{\alpha}!}\left\|D^{\boldsymbol{\alpha}}\tilde{f}\right\|^2_{L^2(\mathbb{T}^2)} \\ &= \sum_{k=1}^{\infty}\frac{c^{-k}}{k!}\sum_{\mathbf{z}\in J_N}|f_{\mathbf{z}}|^2\sum_{\|\boldsymbol{\alpha}\|_1=k}\frac{k!}{\boldsymbol{\alpha}!}(2\pi\mathbf{z})^{2\boldsymbol{\alpha}} \\ &= \sum_{k=1}^{\infty}\frac{c^{-k}}{k!}\sum_{\mathbf{z}\in J_N}|f_{\mathbf{z}}|^2|2\pi\mathbf{z}|^{2k} \\ &\le \left\|\tilde{f}\right\|^2_{L^2(\mathbb{T}^2)}\sum_{k=1}^{\infty}\frac{c^{-k}\left(\pi N\sqrt{2}\right)^{2k}}{k!} \\ &= \left(\mathrm{e}^{\frac{(\pi N\sqrt{2})^2}{c}}-1\right)\left\|\tilde{f}\right\|^2_{L^2(\mathbb{T}^2)}. \end{aligned} \tag{4.17}$$

Finally, combining equations (4.15), (4.16) and (4.17) results in

$$\begin{aligned}\left\|\tilde{f}-\sum_{s=1}^{M}\tilde{f}_s\psi_s\right\|^2_{L^2(\mathbb{T}^2)} &\le \sum_{s=1}^{M}\int_{\mathbb{T}^2}\left|\tilde{f}(\mathbf{x})-\tilde{f}\left(\boldsymbol{\zeta}_s\right)\right|^2\psi_s(\mathbf{x})\mathrm{d}\mathbf{x} \\ &\le \left(\mathrm{e}^{\delta^2 c}-1\right)\left(\mathrm{e}^{\frac{(\pi N)^2}{c}}-1\right)\left\|\tilde{f}\right\|^2_{L^2(\mathbb{T}^2)} \\ &= \left(\mathrm{e}^{\pi N\delta\sqrt{2}}-1\right)^2\left\|\tilde{f}\right\|^2_{L^2(\mathbb{T}^2)}\end{aligned}$$

and equation (4.14) completes the proof. □

Let $\mathbf{A}$ be the matrix, whose adjoint $\mathbf{A}^*$ assigns samples $\mathbf{f}$ to polar frequencies $\hat{\mathbf{f}}$, $\mathbf{f} = \mathbf{A}^*\hat{\mathbf{f}}$. Then, Theorem 40 implies the following statement about the eigenvalues of $\mathbf{A}^*\mathbf{W}\mathbf{A}$.

Corollary 41 (Eigenvalue version)**.** *Let $\delta > 0$ and a δ-dense set*

$$\Omega = \{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_M\} \subset \mathbb{T}^2$$

with the associated Voronoi weights $w_1, \dots, w_M \geq 0$, $M \in \mathbb{N}$, *be given. With*

$$\mathbf{A} := \left(\mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\zeta}_s \cdot \mathbf{z}}\right)_{s \in \{1,\dots,M\}, \mathbf{z} \in J_N} \in \mathbb{C}^{M \times N^2}$$

the eigenvalues $\lambda_1, \dots, \lambda_{N^2}$ *of the matrix* $\mathbf{A}^*\mathbf{W}\mathbf{A}$ *fulfill*

$$\left(2 - \mathrm{e}^{\pi N \delta \sqrt{2}}\right)^2 \leq \lambda_1, \dots, \lambda_{N^2} \leq \mathrm{e}^{2\pi N \delta \sqrt{2}}$$

in particular for $\delta < \frac{\log 2}{\pi N \sqrt{2}}$, *the matrix* $\mathbf{A}^*\mathbf{W}\mathbf{A}$ *has full rank* N^2.

Proof. We choose $\mathbf{f} \in \mathbb{C}^{N^2} \setminus \{\mathbf{0}\}$ arbitrarily, and set $\tilde{\mathbf{f}} := \mathbf{A}\mathbf{f}$ and $\varepsilon := \mathrm{e}^{\pi N \delta \sqrt{2}} - 1$. From Theorem 40 it follows

$$\left| \|\mathbf{f}\|_2 - \left\|\tilde{\mathbf{f}}\right\|_{\boldsymbol{W}} \right| \leq \varepsilon \|\mathbf{f}\|_2,$$

which is equivalent to

$$(1-\varepsilon)^2 \leq \frac{\left\|\tilde{\mathbf{f}}\right\|_{\boldsymbol{W}}^2}{\|\mathbf{f}\|_2^2} \leq (1+\varepsilon)^2. \tag{4.18}$$

Writing

$$\left\|\tilde{\mathbf{f}}\right\|_{\mathbf{W}}^2 = \left\|\mathbf{W}^{\frac{1}{2}}\tilde{\mathbf{f}}\right\|_2^2 = \tilde{\mathbf{f}}^*\mathbf{W}^{\frac{1}{2}}\mathbf{W}^{\frac{1}{2}}\tilde{\mathbf{f}} = \tilde{\mathbf{f}}^*\mathbf{W}\tilde{\mathbf{f}} = \mathbf{f}^*\mathbf{A}^*\mathbf{W}\mathbf{A}\mathbf{f}$$

yields

$$(1-\varepsilon)^2 \leq \frac{\mathbf{f}^*\mathbf{A}^*\mathbf{W}\mathbf{A}\mathbf{f}}{\mathbf{f}^*\mathbf{f}} \leq (1+\varepsilon)^2$$

and hence, the matrix $\mathbf{A}^*\mathbf{W}\mathbf{A}$ has eigenvalues

$$\lambda_1, \dots, \lambda_M \in \left[(1-\varepsilon)^2, (1+\varepsilon)^2\right].$$

With the assumption $\delta < \frac{\log 2}{\pi N \sqrt{2}}$, we obtain positive eigenvalues $\lambda_1, \dots, \lambda_N > 0$ and consequently, the matrix $\mathbf{A}^*\mathbf{W}\mathbf{A}$ has full rank in this case. □

Theorem 40 can also be used to interpret the rows of the matrix $\mathbf{A}$ as elements of a frame.

Definition 42 (*Weighted frame*)**.** A set of vectors $\left\{\mathbf{b}_s \in \mathbb{C}^N\right\}_{s \in \{1,\dots,M\}}$, $M, N \in \mathbb{N}$, is said to be a weighted frame for $\mathbb{C}^N$ with weights $w_1, \dots, w_M > 0$, if there exist constants $A, B > 0$ such that for all $\mathbf{f} \in \mathbb{C}^N$ the relation

$$A\|\mathbf{f}\|_2^2 \leq \sum_{s=1}^{M} w_s \left|\langle \mathbf{f}, \mathbf{b}_s \rangle\right|^2 \leq B\|\mathbf{f}\|_2^2$$

is valid. The constants A and B are called upper and lower frame bounds, respectively. □

The results from Corollary 41 provides directly frame bounds for the rows of the matrix $\mathbf{A}$.

Corollary 43 (Frame version)**.** *Let $N \in \mathbb{N}$, $0 < \delta < \frac{\log 2}{\pi N \sqrt{2}}$, and a δ-dense set $\Omega = \{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_M\} \subset \mathbb{T}^2$ with the associated Voronoi weights $w_1, \dots, w_M \geq 0$, $M \in \mathbb{N}$, be given. With*

$$\mathbf{A} := \left(\mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\zeta}_s \cdot \mathbf{z}}\right)_{s \in \{1,\dots,M\}, \mathbf{z} \in J_N} \in \mathbb{C}^{M \times N^2}$$

the set of rows $\left\{\mathbf{a}_s := \left(\mathbf{A}_{s,J_N}\right)^\top\right\}_{s \in \{1,\dots,M\}}$ of the matrix $\mathbf{A}$ is a weighted frame for $\mathbb{C}^{N^2}$ with the frame bounds

$$A := \left(2 - \mathrm{e}^{\pi N \delta \sqrt{2}}\right)^2 \quad \text{and} \quad B := \mathrm{e}^{2\pi N \delta \sqrt{2}}.$$

Proof. We choose $\mathbf{f} \in \mathbb{C}^{N^2}$ arbitrarily, and set $\tilde{\mathbf{f}} := \mathbf{A}\mathbf{f}$ and $\varepsilon := \mathrm{e}^{\pi N \delta \sqrt{2}} - 1$. Analog to Equation (4.18), we obtain

$$(1-\varepsilon)^2 \|\mathbf{f}\|_2^2 \leq \left\|\tilde{\mathbf{f}}\right\|_{\mathbf{w}}^2 \leq (1+\varepsilon)^2 \|\mathbf{f}\|_2^2.$$

Writing

$$\left\|\tilde{\mathbf{f}}\right\|_{\mathbf{w}}^2 = \sum_{s=1}^{M} w_s \left|\tilde{\mathbf{f}}_s\right|^2 = \sum_{s=1}^{M} w_s \left|\langle \mathbf{f}, \mathbf{a}_s \rangle\right|^2$$

yields the stated frame bounds $A = (1-\varepsilon)^2$ and $B = (1+\varepsilon)^2$. □

5

Inverse problem

Several applications in photoacoustic imaging require an inversion of the mean value operator. More precise, let mean values $\mathbf{g} \in \mathbb{R}^{M_1 M_2}$ for center points $\mathbf{y}_1, \dots \mathbf{y}_{M_1} \in \mathbb{R}^2$, $M_1 \in \mathbb{N}$, and radii $r_1, \dots, r_{M_2} > 0$, $M_2 \in \mathbb{N}$, be given. The task is to find a function $f \colon \mathbb{R}^2 \to \mathbb{R}$, such that

$$\mathcal{M} f\,(\mathbf{y}_j, r_k) = \mathbf{g} \quad \text{for all} \quad j \in \{1, \dots, M_1\} \quad \text{and} \quad k \in \{1, \dots, M_2\} \tag{5.1}$$

is fulfilled. Since we are restricted to finite data, we consider a discretization of equation (5.1). Let a discretization parameter $N \in 2\mathbb{N}$ be given, we denote by $\mathbf{M} \in \mathbb{R}^{M_1 M_2 \times N^2}$ one of the discretizations of the mean value operator from Section 4 with the modification for real data, see Remark 25. Typically, the function f is discretized as the collection of samples on a regular Cartesian grid X_N in a matrix $\mathfrak{f} \in \mathbb{R}^{N \times N}$,

$$\mathfrak{f} \approx (f(\mathbf{x}))_{\mathbf{x} \in X_N}\,.$$

For algorithmic reasons, we will denote by

$$\mathbf{f} := \begin{pmatrix} \mathfrak{f}_{:,1} \\ \vdots \\ \mathfrak{f}_{:,N} \end{pmatrix} \in \mathbb{R}^{N^2}$$

the concatenation of the columns of the matrix $\mathfrak{f} \in \mathbb{R}^{N \times N}$ into a vector of length N^2. With this notation, equation (5.1) can be modified to the task of finding a vector $\mathbf{f} \in \mathbb{R}^{N^2}$, such that $\mathbf{M}\mathbf{f} = \mathbf{g}$ is fulfilled.
However, we have to consider inaccuracies for different reasons. For example, there are simplifications in the modelling of the whole process. Furthermore, instead of continuous functions, we work with discrete data. Another fact is given by noisy measurements, which are inevitable in applications. All in all, an existence of an unique solution is not guaranteed and even if the system of equations has a solution, the results are usually imprecise. To eliminate this issue, it is common to add a convex regularizer $\mathcal{R} \colon \mathbb{R}^{N^2} \to [0, \infty)$ to favor solutions with some special characteristics, given

by a-priori information from the specific application. This leads to the minimization problem

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2 + \mathcal{R}(\mathbf{f}). \tag{5.2}$$

In medical imaging, a suitable assumption to the function f is smoothness away from smooth curves. With a Fourier based approximation, we can expect a good processing of the smooth areas, but for reconstruction of the edges, we have to utilize regularization. In the following, we consider two variants, a total variation and a Shearlet based approach.

5.1 Total variation based regularization

The total variation (TV) seminorm, having great success in image restoration [58], is defined by mixed ℓ^1-ℓ^2-norms of discrete derivatives. Reducing the TV seminorm leads to denoising, while edges are preserved. For the solution of the minimization problem (5.2) with TV based regularization, we proceed with a detailed description of an approach, whose main ideas are published in [13]. We start with a precise formulation of the problem and present a numerical solution with the semismooth Newton method afterwards.

5.1.1 Primal and dual problems

For a stable solution of minimization problems, it is advantageous to analyze the primal and the dual problem, whose derivation requires some notations.

Definition 44. For a discretization parameter $N \in \mathbb{N}$, we define the following discrete gradient operators,

$$\boldsymbol{\nabla}_x \in \mathbb{R}^{N^2\times N^2}, \quad (\boldsymbol{\nabla}_x\mathbf{f})_l := \begin{cases} 0 & \text{for} \quad \mod(l,N)=0, \\ f_{l+1}-f_l & \text{else}, \end{cases}$$

$$\boldsymbol{\nabla}_y \in \mathbb{R}^{N^2\times N^2}, \quad (\boldsymbol{\nabla}_y\mathbf{f})_l := \begin{cases} 0 & \text{for } l\in\{N^2-N+1,\dots,N^2\}, \\ f_{N+l}-f_l & \text{for } l\in\{1,\dots,N^2-N+1\}, \end{cases}$$

$$\boldsymbol{\nabla}_l \in \mathbb{R}^{2\times N^2}, \quad \boldsymbol{\nabla}_l\mathbf{f} := \begin{pmatrix} (\boldsymbol{\nabla}_x\mathbf{f})_l \\ (\boldsymbol{\nabla}_y\mathbf{f})_l \end{pmatrix}, \quad l=1,\dots,N^2,$$

$$\boldsymbol{\nabla} \in \mathbb{R}^{2N^2\times N^2}, \quad \boldsymbol{\nabla}\mathbf{f} := \begin{pmatrix} \boldsymbol{\nabla}_1\mathbf{f} \\ \vdots \\ \boldsymbol{\nabla}_{N^2}\mathbf{f} \end{pmatrix}.$$

Since these operators are linear, we interprete them also as matrices, so that in particular the transposes are well defined. We denote by

$$\mathrm{div} := \boldsymbol{\nabla}^\top \in \mathbb{R}^{N^2\times 2N^2}$$

the discrete divergence operator. □

Definition 45. The total variation (TV) seminorm $\|\cdot\|_{\mathrm{TV}}\colon \mathbb{R}^{N^2} \to \mathbb{R}$ for a discretization parameter $N \in \mathbb{N}$ is given by

$$\|\mathbf{f}\|_{\mathrm{TV}} := \sum_{l=1}^{N^2} \|\boldsymbol{\nabla}_l \mathbf{f}\|_2.$$

□

In addition to the natural Definition 45 of the TV seminorm we will also use the dual representation.

Lemma 46. *Let $\alpha > 0$, $N \in \mathbb{N}$, and $\mathbf{f} \in \mathbb{R}^{N^2}$ be given. Then, for*

$$P_\alpha := \left\{ \mathbf{p} \in \mathbb{R}^{2N^2} : \|\mathbf{p}_l\|_2 \leq \alpha \quad \text{for all} \quad l \in \{1, \dots, N^2\}, \quad \mathbf{p}_l := \begin{pmatrix} p_{2l-1} & p_{2l} \end{pmatrix}^\top \in \mathbb{R}^2 \right\}$$

it follows

$$\alpha \|\mathbf{f}\|_{\mathrm{TV}} = \max_{\mathbf{p} \in P_\alpha} \mathbf{f}^\top \operatorname{div} \mathbf{p}.$$

Proof. Let $\mathbf{f} \in \mathbb{R}^{N^2}$ be arbitrarily. As a consequence from the Cauchy-Schwarz inequality we know

$$\|\mathbf{x}\|_2 = \max_{\mathbf{y} \in \mathbb{R}^2, \|\mathbf{y}\|_2 \leq 1} \mathbf{x}^\top \mathbf{y} \tag{5.3}$$

and it follows

$$\alpha \|\mathbf{f}\|_{\mathrm{TV}} = \alpha \sum_{l=1}^{N^2} \|\boldsymbol{\nabla}_l \mathbf{f}\|_2 = \sum_{l=1}^{N^2} \max_{\mathbf{p}_l \in \mathbb{R}^2, \|\mathbf{p}_l\|_2 \leq \alpha} \mathbf{f}^\top \boldsymbol{\nabla}_l^\top \mathbf{p}_l = \max_{\mathbf{p} \in P_\alpha} \mathbf{f}^\top \operatorname{div} \mathbf{p}.$$

□

For a better readability we introduce a matrix related inner product and the induced norm.

Definition 47. Let a matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$, $M, N \in \mathbb{N}$, be given, such that $\mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{N \times N}$ is invertible. We define the inner product $\langle \cdot, \cdot \rangle_{\mathbf{A}} : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ and the norm $\|\cdot\|_{\mathbf{A}} : \mathbb{R}^N \to \mathbb{R}$ by

$$\langle \mathbf{f}, \mathbf{g} \rangle_{\mathbf{A}} := \mathbf{f}^\top \left(\mathbf{A}^\top \mathbf{A} \right)^{-1} \mathbf{g} \quad \text{and} \quad \|\mathbf{f}\|_{\mathbf{A}} := \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle_{\mathbf{A}}}.$$

□

Remark 48. *For arbitrary center points and radii, and the associated discretization $\mathbf{M}$ of the mean value operator, the invertibility of the matrix $\mathbf{M}^\top \mathbf{M}$ is not given in general. We assume that $\mathbf{M}^\top \mathbf{M}$ is invertible and will add a small multiple of the identity matrix if this is not the case.*

Combining the TV regularization with an ℓ^2-data-fitting term, we reconstruct the image $\mathbf{f}$ by solving the minimization problem

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \mathcal{J}_0(\mathbf{f}), \quad \mathcal{J}_0\colon \mathbb{R}^{N^2} \to \mathbb{R}, \quad \mathcal{J}_0(\mathbf{f}) := \frac{1}{2}\|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2 + \alpha\|\mathbf{f}\|_{\mathrm{TV}}, \tag{5.4}$$

where $\alpha > 0$ is the regularization parameter to control the trade-off between a good fitness to the data $\mathbf{g}$ and the smoothness from the TV term. However, since the TV term is not differentiable, standard methods for minimizing of smooth functionals cannot be applied. Using Lemma 46 and [39, Chapter VII.4] leads to

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \mathcal{J}_0(\mathbf{f}) = \max_{\mathbf{p}\in P_\alpha} \min_{\mathbf{f}\in\mathbb{R}^{N^2}} \frac{1}{2}\|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2 + \mathbf{f}^\top \operatorname{div}\mathbf{p}. \tag{5.5}$$

A solution $\overline{\mathbf{f}} \in \mathbb{R}^{N^2}$ of the inner minimum satisfies

$$0 = \mathbf{M}^\top\left(\mathbf{M}\overline{\mathbf{f}}-\mathbf{g}\right) + \operatorname{div}\mathbf{p} \Leftrightarrow \overline{\mathbf{f}} = (\mathbf{M}^\top\mathbf{M})^{-1}\left(\mathbf{M}^\top\mathbf{g} - \operatorname{div}\mathbf{p}\right).$$

Substituting this solution back to equation (5.5) results in

$$\begin{aligned}
\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \mathcal{J}_0(\mathbf{f}) &= \max_{\mathbf{p}\in P_\alpha} \frac{1}{2}\left\|\mathbf{M}\overline{\mathbf{f}}\right\|_2^2 - \left\langle \mathbf{M}\overline{\mathbf{f}}, \mathbf{g}\right\rangle + \frac{1}{2}\|\mathbf{g}\|_2^2 + (\operatorname{div}\mathbf{p})^\top\overline{\mathbf{f}} \\
&= \max_{\mathbf{p}\in P_\alpha} \frac{1}{2}\left\|\mathbf{M}(\mathbf{M}^\top\mathbf{M})^{-1}\left(\mathbf{M}^\top\mathbf{g}-\operatorname{div}\mathbf{p}\right)\right\|_2^2 \\
&\qquad - \left\langle \mathbf{M}(\mathbf{M}^\top\mathbf{M})^{-1}\left(\mathbf{M}^\top\mathbf{g}-\operatorname{div}\mathbf{p}\right), \mathbf{g}\right\rangle \\
&\qquad + \frac{1}{2}\|\mathbf{g}\|_2^2 + (\operatorname{div}\mathbf{p})^\top(\mathbf{M}^\top\mathbf{M})^{-1}\left(\mathbf{M}^\top\mathbf{g}-\operatorname{div}\mathbf{p}\right) \\
&= \max_{\mathbf{p}\in P_\alpha} \frac{1}{2}\left\|\mathbf{M}^\top\mathbf{g}\right\|_{\mathbf{M}}^2 - \left\langle \mathbf{M}^\top\mathbf{g}, \operatorname{div}\mathbf{p}\right\rangle_{\mathbf{M}} + \frac{1}{2}\|\operatorname{div}\mathbf{p}\|_{\mathbf{M}}^2 - \left\|\mathbf{M}^\top\mathbf{g}\right\|_{\mathbf{M}}^2 \\
&\qquad + \left\langle \mathbf{M}^\top\mathbf{g}, \operatorname{div}\mathbf{p}\right\rangle_{\mathbf{M}} + \frac{1}{2}\|\mathbf{g}\|_2^2 + \left\langle \mathbf{M}^\top\mathbf{g}, \operatorname{div}\mathbf{p}\right\rangle_{\mathbf{M}} - \|\operatorname{div}\mathbf{p}\|_{\mathbf{M}}^2 \\
&= \max_{\mathbf{p}\in P_\alpha} -\frac{1}{2}\left\|\mathbf{M}^\top\mathbf{g}\right\|_{\mathbf{M}}^2 + \left\langle \mathbf{M}^\top\mathbf{g}, \operatorname{div}\mathbf{p}\right\rangle_{\mathbf{M}} - \frac{1}{2}\|\operatorname{div}\mathbf{p}\|_{\mathbf{M}}^2 + \frac{1}{2}\|\mathbf{g}\|_2^2
\end{aligned}$$

and we obtain the dual maximization problem

$$\max_{\mathbf{p}\in P_\alpha} -\frac{1}{2}\left\|\mathbf{M}^\top\mathbf{g} - \operatorname{div}\mathbf{p}\right\|_{\mathbf{M}}^2 + \frac{1}{2}\|\mathbf{g}\|_2^2, \quad \alpha > 0. \tag{5.6}$$

The kernel of the divergence operator $\operatorname{div} \in \mathbb{R}^{N^2\times 2N^2}$ is nontrivial and this holds also for the restricted domain P_α. Consequently, the solution of the dual maximization problem (5.6) is not unique, which poses an algorithmic challenge. To overcome the nondifferentiability of the TV term in the primal problem (5.4) as well as the nonuniqueness of the solution of the dual problem (5.6), we replace similarly to [38] the TV term locally by a quadratic regularization, controlled by a parameter $\gamma > 0$. This acts also as a dual regularization, more precisely, we solve

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \mathcal{J}(\mathbf{f}), \quad \mathcal{J}\colon \mathbb{R}^{N^2} \to \mathbb{R}, \quad \mathcal{J}(\mathbf{f}) := \frac{1}{2}\|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2 + \sum_{l=1}^{N^2} \Phi_{\alpha,\gamma,l}(\boldsymbol{\nabla}\mathbf{f}), \quad \alpha,\gamma > 0, \tag{5.7}$$

with the following definition of the Huber function [41].

Definition 49. For a discretization parameter $N \in \mathbb{N}$ and regularization parameters $\alpha, \gamma > 0$, the Huber function $\Phi_{\alpha,\gamma,l} \colon \mathbb{R}^{2N^2} \to \mathbb{R}$, $l \in \{1, \dots, N^2\}$, is defined by

$$\Phi_{\alpha,\gamma,l}(\mathbf{x}) := \begin{cases} \frac{\alpha}{2\gamma} \|\mathbf{x}_l\|_2^2 & \text{for } \|\mathbf{x}_l\|_2 < \gamma, \\ \alpha \left(\|\mathbf{x}_l\|_2 - \frac{\gamma}{2} \right) & \text{for } \|\mathbf{x}_l\|_2 \geq \gamma \end{cases}$$

with the notation $\mathbf{x}_l = \begin{pmatrix} x_{2l-1} & x_{2l} \end{pmatrix}^\top \in \mathbb{R}^2$ from Definition 45. □

To motivate this construction of the Huber function, we give some additional remarks.

Remark 50. *Definition 49 of the Huber function is essentially based on the norm*

$$h \colon \mathbb{R} \to [0, \infty), \quad h(x) := \alpha \|x\|_2 = \alpha |x|, \quad \alpha > 0,$$

which is used in the Definition 45 of the TV seminorm. To avoid the nondifferentiability in $x = 0$, we replace this norm by the quadratic map

$$h_1 \colon \mathbb{R} \to [0, \infty), \quad h_1(x) := c_1 \|x\|_2^2 = c_1 x^2, \quad \gamma, c_1 > 0,$$

in a neighbourhood $(-\gamma, \gamma)$ of zero. Outside of this region in $(-\infty, \gamma] \cup [\gamma, \infty)$, the Huber function is defined with the original norm h, shifted by $c_2 \in \mathbb{R}$ to obtain a differentiable function. With

$$h_2 \colon \mathbb{R} \to \mathbb{R}, \quad h_2(x) := h(x) + c_2 = \alpha \|x\|_2 + c_2 = \alpha |x| + c_2, \quad \alpha > 0, \ c_2 \in \mathbb{R},$$

and the conditions

$$h_1(\pm\gamma) = h_2(\pm\gamma) \quad \text{and} \quad h_1'(\pm\gamma) = h_2'(\pm\gamma)$$

we arrive at

$$c_1 := \frac{\alpha}{2\gamma} \quad \text{and} \quad c_2 := -\frac{\alpha\gamma}{2}.$$

The resulting functions are illustrated in Figure 5.1.

Similar to Lemma 46 we continue with a dual representation of the regularization term in (5.7).

Lemma 51. *Let a discretization parameter $N \in \mathbb{N}$ and regularization parameters $\alpha, \gamma > 0$ be given. With the notations P_α and $\mathbf{p}_l$ from Lemma 46 it follows for $\mathbf{f} \in \mathbb{R}^{N^2}$ the relation*

$$\sum_{l=1}^{N^2} \Phi_{\alpha,\gamma,l}(\nabla \mathbf{f}) = \max_{\mathbf{p} \in P_\alpha} \mathbf{f}^\top \operatorname{div} \mathbf{p} - \frac{\gamma}{2\alpha} \sum_{l=1}^{N^2} \|\mathbf{p}_l\|_2^2.$$

Proof. The convex conjugate

$$\varphi^* \colon P \to \mathbb{R}, \quad P := \left\{ \mathbf{p} \in \mathbb{R}^2 : \sup_{\mathbf{x} \in \mathbb{R}^2} \mathbf{x}^\top \mathbf{p} - \varphi(\mathbf{x}) < \infty \right\},$$

of a convex function $\varphi \colon \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$\varphi^*(\mathbf{p}) := \sup_{\mathbf{x} \in \mathbb{R}^2} \mathbf{x}^\top \mathbf{p} - \varphi(\mathbf{x}).$$

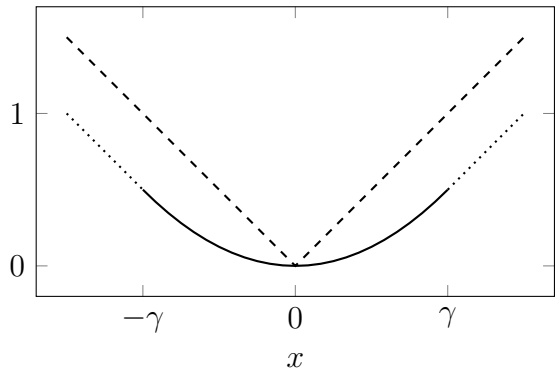


Figure 5.1: The maps from Definition 50, which are the basis for the TV seminorm, $h(x)$ (dashed), and for the Huber function, $h_1(x)$ (solid) and $h_2(x)$ (dotted), with $\alpha = \gamma = 1$.

We consider the function

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}, \quad \varphi(\mathbf{x}) := \Phi_{\alpha,\gamma,1}\left(\begin{pmatrix}\mathbf{x}\\ \mathbf{0}\end{pmatrix}\right) = \begin{cases} \frac{\alpha}{2\gamma}\|\mathbf{x}\|_2^2 & \text{for } \|\mathbf{x}\|_2 < \gamma, \\ \alpha\left(\|\mathbf{x}\|_2 - \frac{\gamma}{2}\right) & \text{for } \|\mathbf{x}\|_2 \geq \gamma. \end{cases}$$

For $\mathbf{p} \in \mathbb{R}^2$, $\|\mathbf{p}\|_2 > \alpha$, we obtain

$$\sup_{\mathbf{x}\in\mathbb{R}^2} \mathbf{x}^\top\mathbf{p} - \varphi(\mathbf{x}) \geq \lim_{n\to\infty} (n\mathbf{p})^\top\mathbf{p} - \varphi(n\mathbf{p}) = \lim_{n\to\infty} n\|\mathbf{p}\|_2\left(\|\mathbf{p}\|_2 - \alpha\right) + \frac{\alpha\gamma}{2} = \infty,$$

and therefore, we assume $\mathbf{p} \in \mathbb{R}^2$, $\|\mathbf{p}\|_2 \leq \alpha$, in the following steps. A maximizer $\overline{\mathbf{x}} \in \mathbb{R}^2$ of

$$\max_{\mathbf{x}\in\mathbb{R}^2} \mathbf{x}^\top\mathbf{p} - \varphi(\mathbf{x}) \tag{5.8}$$

fulfills

$$\mathbf{0} = \begin{cases} \mathbf{p} - \frac{\alpha}{\gamma}\overline{\mathbf{x}} & \text{for } \|\overline{\mathbf{x}}\|_2 < \gamma, \\ \mathbf{p} - \frac{\alpha}{\|\overline{\mathbf{x}}\|_2}\overline{\mathbf{x}} & \text{for } \|\overline{\mathbf{x}}\|_2 \geq \gamma, \end{cases} \tag{5.9}$$

and consequently, it exists a maximum of equation (5.8) in $\overline{\mathbf{x}} = \frac{\gamma}{\alpha}\mathbf{p}$ with the function value $\frac{\gamma}{2\alpha}\|\mathbf{p}\|_2^2$. From the Cauchy-Schwarz inequality it follows

$$\begin{aligned}\sup_{\mathbf{x}\in\mathbb{R}^2,\|\mathbf{x}\|_2\geq\gamma} \mathbf{x}^\top\mathbf{p} - \alpha\left(\|\mathbf{x}\|_2 - \frac{\gamma}{2}\right) &\leq \sup_{\mathbf{x}\in\mathbb{R}^2,\|\mathbf{x}\|_2\geq\frac{\gamma(\|\mathbf{p}\|_2+\alpha)}{2\alpha}} \|\mathbf{x}\|_2\left(\|\mathbf{p}\|_2 - \alpha\right) + \frac{\alpha\gamma}{2} \\ &= \frac{\gamma\left(\|\mathbf{p}\|_2 + \alpha\right)}{2\alpha}\left(\|\mathbf{p}\|_2 - \alpha\right) + \frac{\alpha\gamma}{2} = \frac{\gamma}{2\alpha}\|\mathbf{p}\|_2^2,\end{aligned}$$

which leads to

$$\sup_{\mathbf{x}\in\mathbb{R}^2} \mathbf{x}^\top\mathbf{p} - \varphi(\mathbf{x}) = \max_{\mathbf{x}\in\mathbb{R}^2} \mathbf{x}^\top\mathbf{p} - \varphi(\mathbf{x}),$$

and we obtain the convex conjugate

$$\varphi^*\colon \{\mathbf{p} \in \mathbb{R}^2 : \|\mathbf{p}\|_2 \leq \alpha\} \to \mathbb{R}, \quad \varphi^*(\mathbf{p}) = \frac{\gamma}{2\alpha}\|\mathbf{p}\|_2^2.$$

With [40, Chapter X, Corollary 1.3.6], for $\mathbf{x} \in \mathbb{R}^2$ it follows

$$\varphi(\mathbf{x}) = \varphi^{**}(\mathbf{x}) = \sup_{\mathbf{p}\in\mathbb{R}^2, \|\mathbf{p}\|_2 \leq \alpha} \mathbf{x}^\top \mathbf{p} - \frac{\gamma}{2\alpha}\|\mathbf{p}\|_2^2 = \max_{\mathbf{p}\in\mathbb{R}^2, \|\mathbf{p}\|_2 \leq \alpha} \mathbf{x}^\top \mathbf{p} - \frac{\gamma}{2\alpha}\|\mathbf{p}\|_2^2,$$

which leads to

$$\sum_{l=1}^{N^2} \Phi_{\alpha,\gamma,l}(\boldsymbol{\nabla}\mathbf{f}) = \sum_{l=1}^{N^2} \max_{\mathbf{p}_l \in \mathbb{R}^2 : \|\mathbf{p}_l\|_2 \leq \alpha} \mathbf{f}^\top \boldsymbol{\nabla}^\top \mathbf{p}_l - \frac{\gamma}{2\alpha}\|\mathbf{p}_l\|_2^2$$

and this proves the assertion. □

With Lemma 51 and [39, Chapter VII.4], we obtain

$$\begin{aligned}\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \mathcal{J}(\mathbf{f}) &= \min_{\mathbf{f}\in\mathbb{R}^{N^2}} \frac{1}{2}\|\mathbf{M}\mathbf{f} - \mathbf{g}\|_2^2 + \max_{\mathbf{p}\in P_\alpha} \mathbf{f}^\top \operatorname{div} \mathbf{p} - \frac{\gamma}{2\alpha}\sum_{l=1}^{N^2} \|\mathbf{p}_l\|_2^2 \\ &= \max_{\mathbf{p}\in P_\alpha} \min_{\mathbf{f}\in\mathbb{R}^{N^2}} \frac{1}{2}\|\mathbf{M}\mathbf{f} - \mathbf{g}\|_2^2 + \mathbf{f}^\top \operatorname{div} \mathbf{p} - \frac{\gamma}{2\alpha}\sum_{l=1}^{N^2} \|\mathbf{p}_l\|_2^2.\end{aligned}$$

Since the sum is independent of $\mathbf{f}$, we can use the results from the derivation of equation (5.6) and conclude, that the primal minimization problem (5.7) is equivalent to the dual maximization problem

$$\max_{\mathbf{p}\in P_\alpha} -\frac{1}{2}\left\|\mathbf{M}^\top \mathbf{g} - \operatorname{div} \mathbf{p}\right\|_{\mathbf{M}}^2 + \frac{1}{2}\|\mathbf{g}\|_2^2 - \frac{\gamma}{2\alpha}\sum_{l=1}^{N^2} \|\mathbf{p}_l\|_2^2, \quad \alpha, \gamma > 0. \tag{5.10}$$

For a better understanding of the workflow, we summarize the previous optimization problems.

Remark 52. *We start with the primal problem*

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \frac{1}{2}\|\mathbf{M}\mathbf{f} - \mathbf{g}\|_2^2 + \alpha\|\mathbf{f}\|_{\mathrm{TV}}, \tag{5.11}$$

which is equivalent to the dual problem

$$\max_{\mathbf{p}\in P_\alpha} -\frac{1}{2}\left\|\mathbf{M}^\top \mathbf{g} - \operatorname{div} \mathbf{p}\right\|_{\mathbf{M}}^2 + \frac{1}{2}\|\mathbf{g}\|_2^2, \tag{5.12}$$

where the regularization parameter $\alpha > 0$ controls the trade-off between a good fitness to the data $\mathbf{g}$ and the smoothness from the TV term. To overcome the nondifferentiablity of the primal problem (5.4) and the nonuniqueness of the dual problem (5.6), we consider the modified primal problem

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2}} \frac{1}{2}\|\mathbf{M}\mathbf{f} - \mathbf{g}\|_2^2 + \sum_{l=1}^{N^2} \Phi_{\alpha,\gamma,l}(\boldsymbol{\nabla}\mathbf{f}) \tag{5.13}$$

with the associated dual problem

$$\max_{\mathbf{p}\in P_\alpha} -\frac{1}{2}\left\|\mathbf{M}^\top \mathbf{g} - \operatorname{div} \mathbf{p}\right\|_{\mathbf{M}}^2 + \frac{1}{2}\|\mathbf{g}\|_2^2 - \frac{\gamma}{2\alpha}\sum_{l=1}^{N^2} \|\mathbf{p}_l\|_2^2, \tag{5.14}$$

where the parameter $\gamma > 0$ regulates the balance between quadratic regularization and the TV term. Roughly speaking, large γ leads to smooths edges and benefits the numerical computation. On the other hand, small γ leads to good edge preservation due to the nondifferentiability of the TV term.

To establish a connection between the original optimization problem (5.4) and the modified problem (5.7), we refer to a result from [38].

Lemma 53. *Let $\bar{\mathbf{f}}$ be a solution of (5.4) and $\bar{\mathbf{f}}_\gamma$ be a solution of (5.7). Then, $\bar{\mathbf{f}}_\gamma$ converges to $\bar{\mathbf{f}}$ for $\gamma \to 0$.*

Proof. See [38, Theorem 2.2]. □

5.1.2 Numerical solution

Let $\bar{\mathbf{f}} \in \mathbb{R}^{N^2}$ be a solution of the primal problem (5.7). Based on equation (5.9),

$$\bar{\mathbf{p}} := \begin{pmatrix} \bar{\mathbf{p}}_1 \\ \vdots \\ \bar{\mathbf{p}}_{N^2} \end{pmatrix}, \quad \bar{\mathbf{p}}_l := \begin{cases} \frac{\alpha}{\gamma} \boldsymbol{\nabla}_l \bar{\mathbf{f}} \in \mathbb{R}^{N^2} & \text{for } \left\| \boldsymbol{\nabla}_l \bar{\mathbf{f}} \right\|_2 < \gamma, \\ \frac{\alpha}{\left\| \boldsymbol{\nabla}_l \bar{\mathbf{f}} \right\|_2} \boldsymbol{\nabla}_l \bar{\mathbf{f}} & \text{for } \left\| \boldsymbol{\nabla}_l \bar{\mathbf{f}} \right\|_2 \geq \gamma, \end{cases}$$

is a solution of the dual problem (5.10). From the first order optimality condition for equation (5.7), it follows

$$\mathbf{0} = \mathbf{M}^\top \left(\mathbf{M}\bar{\mathbf{f}} - \mathbf{g} \right) + \sum_{l=1}^{N^2} \boldsymbol{\nabla}_l^\top \bar{\mathbf{p}}_l = \mathbf{M}^\top \left(\mathbf{M}\bar{\mathbf{f}} - \mathbf{g} \right) + \operatorname{div} \bar{\mathbf{p}},$$

which is equivalent to the system of equations

$$\begin{pmatrix} F_1(\mathbf{f}, \mathbf{p}) \\ F_2(\mathbf{f}, \mathbf{p}) \end{pmatrix} = \mathbf{0} \tag{5.15}$$

where the functions $F_1 \colon \mathbb{R}^{N^2} \times \mathbb{R}^{2N^2} \to \mathbb{R}^{N^2}$, $F_2 \colon \mathbb{R}^{N^2} \times \mathbb{R}^{2N^2} \to \mathbb{R}^{2N^2}$, and

$$F_{2,l} \colon \mathbb{R}^{N^2} \times \mathbb{R}^{2N^2} \to \mathbb{R}^2, \; l \in \{1, \ldots, N^2\},$$

are given by

$$F_1(\mathbf{f}, \mathbf{p}) := \mathbf{M}^\top \left(\mathbf{M}\mathbf{f} - \mathbf{g} \right) + \operatorname{div} \mathbf{p}, \quad F_2 := \begin{pmatrix} F_{2,1} \\ \vdots \\ F_{2,N^2} \end{pmatrix}, \tag{5.16}$$

and

$$\begin{aligned} F_{2,l}(\mathbf{f}, \mathbf{p}) &:= \alpha \boldsymbol{\nabla}_l \mathbf{f} - \max\left\{ \gamma, \left\| \boldsymbol{\nabla}_l \mathbf{f} \right\|_2 \right\} \mathbf{p}_l \\ &= \begin{cases} \alpha \boldsymbol{\nabla}_l \mathbf{f} - \gamma \mathbf{p}_l & \text{for } \left\| \boldsymbol{\nabla}_l \mathbf{f} \right\|_2 < \gamma, \\ \alpha \boldsymbol{\nabla}_l \mathbf{f} - \left\| \boldsymbol{\nabla}_l \mathbf{f} \right\|_2 \mathbf{p}_l & \text{for } \left\| \boldsymbol{\nabla}_l \mathbf{f} \right\|_2 \geq \gamma. \end{cases} \end{aligned} \tag{5.17}$$

A well-known method for solving such a system numerically, is the Newton method, but the presence of the max-operator with the nondifferentiability as consequence prevents us to use the standard formulation. Similar to [22, 37], we apply the semismooth or generalized Newton method, which requires the introduction of a generalized derivative.

Definition 54. A function $f\colon \mathbb{R}^{N_1} \to \mathbb{R}^{N_2}$, $N_1, N_2 \in \mathbb{N}$, is called Newton (or slantly) differentiable on $\mathbb{R}^{N_1}$, if there exists a mapping $\mathcal{G}\colon \mathbb{R}^{N_1} \to \mathcal{L}\left(\mathbb{R}^{N_1}, \mathbb{R}^{N_2}\right)$ with

$$\lim_{\substack{\mathbf{h}\to\mathbf{0}\\ \mathbf{h}\neq\mathbf{0}}} \frac{\|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - (\mathcal{G}(\mathbf{x}+\mathbf{h}))(\mathbf{h})\|_2}{\|\mathbf{h}\|_2} = 0 \quad \text{for all} \quad \mathbf{x} \in \mathbb{R}^{N_1}.$$

The function $\mathcal{G}$ is called a generalized derivative (or slanting function) for f. □

Remark 55. *If the function f is differentiable in $\mathbf{x}$, then there exists a generalized derivative $\mathcal{G}$ for f in $\mathbf{x}$ with $\mathcal{G}(\mathbf{x}) = \mathbf{J}$, where $\mathbf{J}$ denotes the Jacobian matrix of f in $\mathbf{x}$.*

For a system of equations $F(\mathbf{x}) = \mathbf{0}$, a iteration step of the semismooth Newton method is given by

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left(\mathcal{G}\left(\mathbf{x}^{(k)}\right)\right)^{-1}\left(F\left(\mathbf{x}^{(k)}\right)\right), \quad k \in \mathbb{N}\cup\{0\}.$$

With the methods from [22], we obtain, that there are general derivatives

$$\mathcal{G}_1\colon \mathbb{R}^{N^2} \times \mathbb{R}^{2N^2} \to \mathcal{L}\left(\mathbb{R}^{N^2} \times \mathbb{R}^{2N^2}, \mathbb{R}^{N^2}\right)$$

and

$$\mathcal{G}_2\colon \mathbb{R}^{N^2} \times \mathbb{R}^{2N^2} \to \mathcal{L}\left(\mathbb{R}^{N^2} \times \mathbb{R}^{2N^2}, \mathbb{R}^{2N^2}\right)$$

for the functions F_1 and F_2, see equations (5.16) and (5.17), with

$$\mathbf{G}_1^{(k)} := \mathcal{G}_1\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) = \begin{pmatrix}\mathbf{M}^\top\mathbf{M} & \boldsymbol{\nabla}^\top\end{pmatrix} \in \mathbb{R}^{N^2\times 3N^2}$$

and

$$\mathbf{G}_2^{(k)} := \mathcal{G}_2\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) = \begin{pmatrix}\mathbf{A}^{(k)} & \mathbf{D}^{(k)}\end{pmatrix} \in \mathbb{R}^{2N^2\times 3N^2}, \quad k \in \mathbb{N}\cup\{0\}.$$

Here, we use for $l \in \{1, \ldots, N^2\}$ and $k \in \mathbb{N}\cup\{0\}$ the following notations,

$$\mathbf{D}^{(k)} := \begin{pmatrix}\mathbf{D}_1^{(k)} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{D}_{N^2}^{(k)}\end{pmatrix} \in \mathbb{R}^{2N^2\times 2N^2},$$

$$\mathbf{D}_l^{(k)} := -\max\left\{\gamma, \left\|\boldsymbol{\nabla}_l\mathbf{f}^{(k)}\right\|_2\right\}\mathbf{I}_{2\times 2} \in \mathbb{R}^{2\times 2},$$

and

$$\mathbf{A}^{(k)} := \alpha\boldsymbol{\nabla} - \boldsymbol{\chi}^{(k)}\mathbf{P}^{(k)}\mathbf{B}^{(k)}\boldsymbol{\nabla} \in \mathbb{R}^{2N^2\times N^2},$$

$$\boldsymbol{\chi}^{(k)} := \begin{pmatrix}\chi_1^{(k)}\mathbf{I}_{2\times 2} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \chi_{N^2}^{(k)}\mathbf{I}_{2\times 2}\end{pmatrix} \in \mathbb{R}^{2N^2\times 2N^2}, \quad \chi_l^{(k)} := \begin{cases}0 & \text{for } \left\|\boldsymbol{\nabla}_l\mathbf{f}^{(k)}\right\|_2 < \gamma,\\ 1 & \text{for } \left\|\boldsymbol{\nabla}_l\mathbf{f}^{(k)}\right\|_2 \geq \gamma,\end{cases}$$

$$\mathbf{P}^{(k)} := \operatorname{diag}\mathbf{p}^{(k)} = \begin{pmatrix}p_1^{(k)} & & 0\\ & \ddots & \\ 0 & & p_{2N^2}^{(k)}\end{pmatrix} \in \mathbb{R}^{2N^2\times 2N^2},$$

$$\mathbf{B}^{(k)} := \begin{pmatrix} \frac{\left(\boldsymbol{\nabla}_1 \mathbf{f}^{(k)} \quad \boldsymbol{\nabla}_1 \mathbf{f}^{(k)}\right)^\top}{\left\|\boldsymbol{\nabla}_1 \mathbf{f}^{(k)}\right\|_2} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \frac{\left(\boldsymbol{\nabla}_{N^2} \mathbf{f}^{(k)} \quad \boldsymbol{\nabla}_{N^2} \mathbf{f}^{(k)}\right)^\top}{\left\|\boldsymbol{\nabla}_{N^2} \mathbf{f}^{(k)}\right\|_2} \end{pmatrix} \in \mathbb{R}^{2N^2 \times 2N^2}.$$

Each Newton step for solving the system (5.15) is given by

$$\begin{pmatrix} \mathbf{f}^{(k+1)} \\ \mathbf{p}^{(k+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{f}^{(k)} \\ \mathbf{p}^{(k)} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} \\ \boldsymbol{\delta}_{\mathbf{p}}^{(k)} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{M}^\top \mathbf{M} & \boldsymbol{\nabla}^\top \\ \mathbf{A}^{(k)} & \mathbf{D}^{(k)} \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} \\ \boldsymbol{\delta}_{\mathbf{p}}^{(k)} \end{pmatrix} = - \begin{pmatrix} F_1\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) \\ F_2\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) \end{pmatrix}$$

and from the lower block it follows

$$\boldsymbol{\delta}_{\mathbf{p}}^{(k)} = - \left(\mathbf{D}^{(k)}\right)^{-1} \left(\mathbf{A}^{(k)} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} + F_2\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right)\right). \tag{5.18}$$

Together with the first block, this leads to

$$\begin{aligned} & \mathbf{M}^\top \mathbf{M} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} - \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} \left(\mathbf{A}^{(k)} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} + F_2\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right)\right) = -F_1\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) \\ \Leftrightarrow & \left(\mathbf{M}^\top \mathbf{M} - \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} \mathbf{A}^{(k)}\right) \boldsymbol{\delta}_{\mathbf{f}}^{(k)} = -F_1\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) + \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} F_2\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right) \\ \Leftrightarrow & \left(\mathbf{M}^\top \mathbf{M} - \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} \mathbf{A}^{(k)}\right) \boldsymbol{\delta}_{\mathbf{f}}^{(k)} = -\mathbf{M}^\top \left(\mathbf{M}\mathbf{f}^{(k)} - \mathbf{g}\right) + \alpha \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} \boldsymbol{\nabla} \mathbf{f}^{(k)} \end{aligned}$$

and again with equation (5.18), we obtain

$$\begin{aligned} \mathbf{p}^{(k+1)} &= \mathbf{p}^{(k)} + \boldsymbol{\delta}_{\mathbf{p}}^{(k)} = \mathbf{p}^{(k)} - \left(\mathbf{D}^{(k)}\right)^{-1} \left(\mathbf{A}^{(k)} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} + F_2\left(\mathbf{f}^{(k)}, \mathbf{p}^{(k)}\right)\right) \\ &= - \left(\mathbf{D}^{(k)}\right)^{-1} \left(\mathbf{A}^{(k)} \boldsymbol{\delta}_{\mathbf{f}}^{(k)} + \alpha \boldsymbol{\nabla} \mathbf{f}^{(k)}\right) \\ &= - \left(\mathbf{D}^{(k)}\right)^{-1} \left(\alpha \boldsymbol{\nabla} \mathbf{f}^{(k+1)} - \boldsymbol{\chi}^{(k)} \mathbf{P}^{(k)} \mathbf{B}^{(k)} \boldsymbol{\nabla} \boldsymbol{\delta}_{\mathbf{f}}^{(k)}\right). \end{aligned}$$

Algorithm 8 summarizes the results of the previous computations.

5.2 Shearlet based regularization

Shearlets are a representation system, generated by translations, scalings, and shearings of a so-called mother Shearlet. Roughly speaking, the shearing operation is something like a rotation, which allows to capture directional features like edges. Moreover, Shearlets provide a sparse decomposition of images which are smooth away from smooth curves. These features motivate to regularize the reconstruction by searching for solutions with sparse Shearlet decompositions. Since a small ℓ^1-norm provokes sparsity, we consider the minimization problem

$$\min_{\mathbf{f} \in \mathbb{R}^{N^2}} \|\mathbf{M}\mathbf{f} - \mathbf{g}\|_2^2 + \lambda \|\mathbf{S}\mathbf{f}\|_1, \quad \lambda > 0, \tag{5.19}$$

where $\mathbf{S}$ is the discrete operator, which assigns a vector its Shearlet coefficients. We continue with a short introduction to Shearlets and present an approach for solving the minimizing problem in equation (5.19) with the alternating direction method of multipliers. Finally, we construct a modification of the fast finite Shearlet transform, which fits to the polar frequency grid from Section 4.2.2.

Algorithm 8 Inversion of the discrete mean value operator, using a generalized Newton method and total variation based regularization.

Input

$N \in 2\mathbb{N}$, $M_1 \in \mathbb{N}$, $M_2 \in \mathbb{N}$ ▷ discretization parameters

$\mathbf{M} \in \mathbb{R}^{M_1 M_2 \times N^2}$ ▷ discrete mean value operator

$\mathbf{g} \in \mathbb{R}^{M_1 M_2}$ ▷ mean values

$\alpha > 0$ ▷ regularization parameter

$\gamma > 0$ ▷ trade-off between quadratic and TV regularization

Output

$\mathbf{f} \in \mathbb{R}^{N^2}$ ▷ reconstruction, $\mathbf{Mf} \approx \mathbf{g}$

$\mathbf{f}^{(0)} \in \mathbb{R}^{N^2}$, $\mathbf{p}^{(0)} \in \mathbb{R}^{2N^2}$ ▷ initialization

$k := 0$

repeat

  create $\mathbf{A}^{(k)}$, $\mathbf{D}^{(k)}$, $\mathbf{B}^{(k)}$ and $\boldsymbol{\chi}^{(k)}$

  $\mathbf{A} := \left(\mathbf{M}^\top \mathbf{M} - \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} \mathbf{A}^{(k)}\right)$

  $\mathbf{b} := -\mathbf{M}^\top \left(\mathbf{M}\mathbf{f}^{(k)} - \mathbf{g}\right) + \alpha \boldsymbol{\nabla}^\top \left(\mathbf{D}^{(k)}\right)^{-1} \boldsymbol{\nabla}\mathbf{f}^{(k)}$

  $\boldsymbol{\delta}_{\mathbf{f}}^{(k)} :=$ solution of $\mathbf{A}\boldsymbol{\delta}_{\mathbf{f}}^{(k)} = \mathbf{b}$ ▷ linear system of equations

  $\mathbf{f}^{(k+1)} := \mathbf{f}^{(k)} + \boldsymbol{\delta}_{\mathbf{f}}^{(k)}$ ▷ update of the primal variable

  $\mathbf{p}^{(k+1)} := -\left(\mathbf{D}^{(k)}\right)^{-1} \left(\alpha \boldsymbol{\nabla}\mathbf{f}^{(k+1)} - \boldsymbol{\chi}^{(k)} \mathbf{P}^{(k)} \mathbf{B}^{(k)} \boldsymbol{\nabla} \boldsymbol{\delta}_{\mathbf{f}}^{(k)}\right)$ ▷ update of the dual variable

  $k := k + 1$

until a convergence criterion is reached

$\mathbf{f} := \mathbf{f}^{(k)}$

5.2.1 Directional representation systems

Many applications ask for representation systems, which allow highly sparse approximations of functions. For one-dimensional signals, Wavelets [54] and their associated transforms provide those features and became a successful tool for detection and analyzing of pointwise singularities. However, Wavelets are not qualified for an optimal handling of multivariate functions, since they are deficient in dealing with anisotropic features. Starting from this fact, a variety of different representation systems were introduced with the goal to achieve an optimal sparse approximation of two-dimensional functions in the sense of the decay rate of the L^2-error of the best N-term approximation. The system of Curvelets [9] achieves such an optimal behaviour for functions that are C^2 except for discontinuities along C^2 curves, so-called cartoon-like functions, which were introduced in [14]. Since the rotation operation requires a frequency partition on polar grids, the discretization for images, sampled on Cartesian grids, is very challenging [8]. This fact motivated the development of Contourlets [12], but in contrast to Wavelets, there is no unified treatment of the continuous and the discrete theory for both Curvelets and Contourlets, and hence, the implementation differs from the continuous theory. Shearlets [52, 24] were introduced to provide a framework with an optimal sparse approximation of cartoon-like functions on the one side and with a faithful implementation on the other side. Since the further considerations focus on the application of Shearlets we will not go into any more detail, but note shortly, that there exist also many other representation systems such as anisotropic Wavelets, Bandlets, Ridgelets, Wedgelets, and so on.

5.2.2 Implementations of Shearlet transforms

From the construction of the Shearlets, the implementation can be realized by a natural discretization of the continuous theory. Currently, there are three toolboxes available, published as Local Shearlet Toolbox [15], ShearLab [50, 53, 49, 51] and fast finite Shearlet transform (FFST) [36, 35], see the given references for details. Since the algorithms related to this thesis are implemented with the FFST, we continue with a basic overview to the construction of the implemented Shearlet system following [36]. We start with a so-called mother Shearlet $\psi\colon \mathbb{R}^2 \to \mathbb{R}$, which is defined by its Fourier transform

$$\hat{\psi}\colon \mathbb{R}^2 \to \mathbb{R}, \quad \hat{\psi}(\omega_1, \omega_2) := \hat{\psi}_1(\omega_1)\hat{\psi}_2\left(\frac{\omega_2}{\omega_1}\right),$$

where $\psi_1\colon \mathbb{R} \to \mathbb{R}$ behaves like a Wavelet and $\psi_2\colon \mathbb{R} \to \mathbb{R}$ like a spline, such that $\hat{\psi}$ consists of two wedges, see Figure 5.2a. Figure 5.2b shows the mother Shearlet in the spatial domain. The whole system of Shearlets is constructed by dilations $a > 0$, shearings $s \in \mathbb{R}$, and translations $\mathbf{t} \in \mathbb{R}^2$ of the mother Shearlet ψ in the following way,

$$\psi_{a,s,\mathbf{t}}\colon \mathbb{R}^2 \to \mathbb{R}, \quad \psi_{a,s,\mathbf{t}}(\mathbf{x}) := a^{-\frac{3}{4}}\psi\left(\mathbf{A}_a^{-1}\mathbf{S}_s^{-1}(\mathbf{x} - \mathbf{t})\right),$$

with

$$\mathbf{A}_a := \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} \quad \text{and} \quad \mathbf{S}_s := \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

Figure 5.2 illustrates the dilation and shearing operations. In addition, there are some more technical issues necessary, for example the treatment of lower frequencies and the arrangement into cones. For simplicity, we skip these details in the following. The continuous Shearlet transform $\mathcal{S}$ of a square-integrable function $f \in L^2(\mathbb{R}^2)$ is defined by

$$(\mathcal{S}f)(a, s, \mathbf{t}) := \langle f, \psi_{a,s,\mathbf{t}} \rangle .$$

The discretization is done in two steps. For a discretization parameter $N \in 2\mathbb{N}$, we start with the discretization of the parameters,

$$a_j := \frac{1}{4^j}, \quad j \in \{0, \dots, j_0 - 1\}, \quad j_0 := \left\lceil \frac{1}{2} \log_2 N \right\rceil ,$$

$$s_{j,k} := \frac{k}{2^j}, \quad k \in \{-k_j, \dots, k_j\}, \quad k_j \in \mathbb{N},$$

and

$$\mathbf{t}_\mathbf{m} := \frac{\mathbf{m}}{N}, \quad \mathbf{m} \in [0, N-1]^2 \cap \mathbb{Z}^2.$$

We note, that in particular this discretization of the shearing operation $s \in \mathbb{R}$ is an oversimplification. As mentioned before, an efficient implementation requires some additional technical details. We obtain continuous Shearlets with discrete parameters,

$$\psi^{\mathrm{D}}_{j,k,\mathbf{m}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \psi^{\mathrm{D}}_{j,k,\mathbf{m}}(\mathbf{x}) := \psi_{a_j, s_{j,k}, \mathbf{t}_\mathbf{m}}(\mathbf{x}).$$

Secondly, we replace continuous functions by function samples on a grid. In particular, we define discrete Shearlets $\mathbf{\Psi}_{j,k,\mathbf{m}} \in \mathbb{C}^{N^2}$ by their discrete Fourier transforms

$$\hat{\mathbf{\Psi}}_{j,k,\mathbf{m}} := \left(\hat{\psi}^{\mathrm{D}}_{j,k,\mathbf{m}}(\boldsymbol{\omega}) \right)_{\boldsymbol{\omega} \in J_N} \in \mathbb{R}^{N^2}.$$

We denote by η the number of scales and shears,

$$\eta := \sum_{j=0}^{j_0 - 1} (2k_j + 1)$$

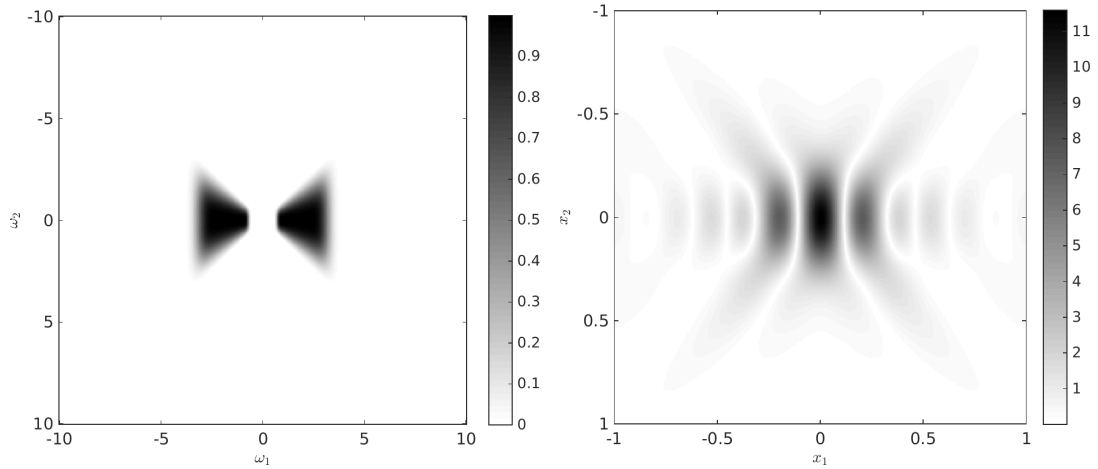
and define the discrete Shearlet transform $\mathbf{Sf} \in \mathbb{C}^{\eta N^2}$ of a vector $\mathbf{f} \in \mathbb{R}^{N^2}$ by

$$(\mathbf{Sf})_{j,k,\mathbf{m}} := \langle \mathbf{f}, \mathbf{\Psi}_{j,k,\mathbf{m}} \rangle .$$

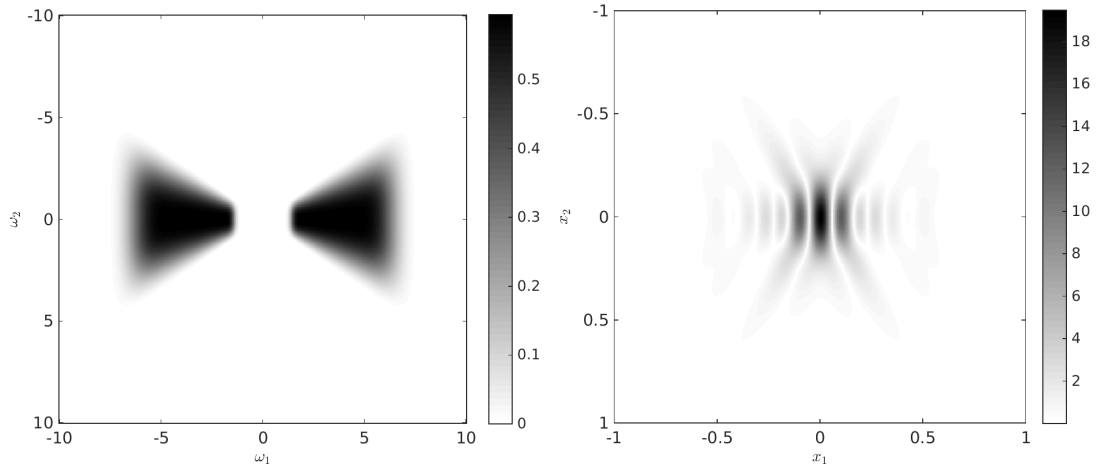
With Parseval's identity, it follows

$$(\mathbf{Sf})_{j,k,\mathbf{m}} := \left\langle \hat{\mathbf{f}}, \hat{\mathbf{\Psi}}_{j,k,\mathbf{m}} \right\rangle ,$$

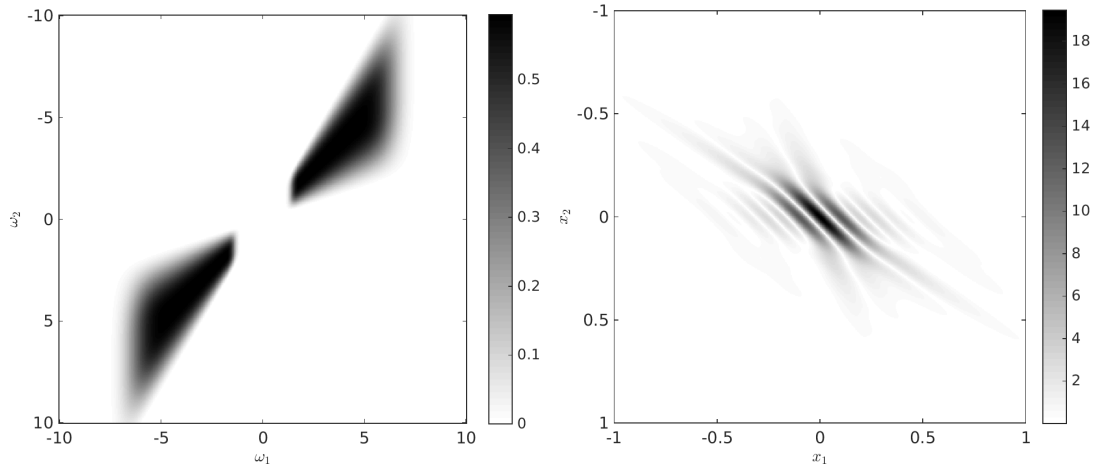
which leads to computations of fast Fourier transforms. This system of discrete Shearlets constitute a Parseval frame and consequently, the inverse transform is given by the transpose operator. Here we stop this rough introduction to the FFST and refer to [36] for a detailed and extensive description.



(a) $\hat{\psi}(\boldsymbol{\omega}) = \hat{\psi}_{1,0,\mathbf{0}}(\boldsymbol{\omega})$ (b) $\psi(\mathbf{x}) = \psi_{1,0,\mathbf{0}}(\mathbf{x})$



(c) $\hat{\psi}_{0.5,0,\mathbf{0}}(\boldsymbol{\omega})$ (d) $\psi_{0.5,0,\mathbf{0}}(\mathbf{x})$



(e) $\hat{\psi}_{0.5,1,\mathbf{0}}(\boldsymbol{\omega})$ (f) $\psi_{0.5,1,\mathbf{0}}(\mathbf{x})$

Figure 5.2: The mother Shearlet (first row) and dilated and sheared versions (second and third row). The left column shows the Fourier transforms and the right the appropriate functions in the spatial domain.

5.2.3 Numerical solution

In order to solve the minimization problem in equation (5.19), we use the strategy of a similar work about seismic data reconstruction via Shearlet-regularized directional inpainting [34]. One difference is the fact, that we deal with the discrete mean operator $\mathbf{M}$ instead of a diagonal matrix, which requires to solve additionally a linear system of equations.
We start by rewriting the minimization problem (5.19) as

$$\min_{\mathbf{f}\in\mathbb{R}^{N^2},\mathbf{v}\in\mathbb{R}^{\eta N^2}} \|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2+\lambda\|\mathbf{v}\|_1 \quad \text{subject to} \quad \mathbf{S}\mathbf{f}=\mathbf{v}, \quad \lambda>0. \tag{5.20}$$

With

$$\mathbf{A}:=\mathbf{S}, \quad \mathbf{B}:=-\mathbf{I}, \quad \mathbf{c}:=\mathbf{0},$$

and the convex functions $F_1\colon \mathbb{R}^{N^2}\to\mathbb{R}$ and $F_2\colon \mathbb{R}^{\eta N^2}\to\mathbb{R}$,

$$F_1(\mathbf{f}):=\|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2 \quad \text{and} \quad F_2(\mathbf{v}):=\lambda\|\mathbf{v}\|_1,$$

equation (5.20) is equivalent to

$$\min_{\mathbf{x}\in\mathbb{R}^{n_1},\mathbf{y}\in\mathbb{R}^{n_2}} F_1(\mathbf{x})+F_2(\mathbf{y}) \quad \text{subject to} \quad \mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{y}=\mathbf{c}.$$

As well as in [35, 34], we apply the alternating direction method of multipliers (ADMM) [7] to solve this minimization problem. The scaled form of the ADMM consists of the iterations

$$\mathbf{f}^{(k+1)}:=\operatorname*{argmin}_{\mathbf{f}\in\mathbb{R}^{N^2}}\left(\|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2+\frac{\rho}{2}\left\|\mathbf{S}\mathbf{f}-\mathbf{v}^{(k)}+\mathbf{u}^{(k)}\right\|_2^2\right), \tag{5.21}$$

$$\mathbf{v}^{(k+1)}:=\operatorname*{argmin}_{\mathbf{v}\in\mathbb{R}^{\eta N^2}}\left(\lambda\|\mathbf{v}\|_1+\frac{\rho}{2}\left\|\mathbf{S}\mathbf{f}^{(k+1)}-\mathbf{v}+\mathbf{u}^{(k)}\right\|_2^2\right), \text{ and} \tag{5.22}$$

$$\mathbf{u}^{(k+1)}:=\mathbf{u}^{(k)}+\mathbf{S}\mathbf{f}^{(k+1)}-\mathbf{v}^{(k+1)}, \tag{5.23}$$

where $\rho>0$ penalizes the constraint $\mathbf{S}\mathbf{f}=\mathbf{v}$, see Algorithm 9 for the complete procedure.
While the computation of the update of the variable $\mathbf{u}$ in equation (5.23) is straightforward, we comment on the first two steps. Since the square of the ℓ^2-norm is differentiable, we compute

$$\begin{aligned}\nabla_{\mathbf{f}}\left(\|\mathbf{M}\mathbf{f}-\mathbf{g}\|_2^2+\frac{\rho}{2}\left\|\mathbf{S}\mathbf{f}-\mathbf{v}^{(k)}+\mathbf{u}^{(k)}\right\|_2^2\right)&\\ =2\mathbf{M}^\top(\mathbf{M}\mathbf{f}-\mathbf{g})+\rho\mathbf{S}^\top\left(\mathbf{S}\mathbf{f}-\mathbf{v}^{(k)}+\mathbf{u}^{(k)}\right)&\\ =\left(2\mathbf{M}^\top\mathbf{M}+\rho\mathbf{S}^\top\mathbf{S}\right)\mathbf{f}-2\mathbf{M}^\top\mathbf{g}+\rho\mathbf{S}^\top\left(\mathbf{u}^{(k)}-\mathbf{v}^{(k)}\right)&\end{aligned}$$

and due to the convexity, the solution for the minimization problem (5.21) is given by the solution of the symmetric system of linear equations

$$\left(\mathbf{M}^\top\mathbf{M}+\frac{\rho}{2}\mathbf{S}^\top\mathbf{S}\right)\mathbf{f}^{(k+1)}=\mathbf{M}^\top\mathbf{g}+\frac{\rho}{2}\mathbf{S}^\top\left(\mathbf{v}^{(k)}-\mathbf{u}^{(k)}\right). \tag{5.24}$$

Algorithm 9 Inversion of the discrete mean value operator, using ADMM and Shearlet based regularization.

Input

$N \in 2\mathbb{N}$, $M_1 \in \mathbb{N}$, $M_2 \in \mathbb{N}$ ▷ discretization parameters
$\mathbf{M} \in \mathbb{R}^{M_1 M_2 \times N^2}$ ▷ discrete mean value operator
$\mathbf{g} \in \mathbb{R}^{M_1 M_2}$ ▷ mean values
$\eta \in \mathbb{N}$ ▷ number of scales and shears
$\mathbf{S} \in \mathbb{R}^{\eta N^2 \times N^2}$ ▷ discrete Shearlet operator
$\lambda, \rho > 0$ ▷ penalty parameters

Output

$\mathbf{f} \in \mathbb{R}^{N^2}$ ▷ reconstruction, $\mathbf{Mf} \approx \mathbf{g}$

$\mathbf{v}^{(0)}, \mathbf{u}^{(0)} \in \mathbb{R}^{\eta N^2}$ ▷ initialization
$k := 0$
repeat
  $\mathbf{f}^{(k+1)} := \operatorname{argmin}_{\mathbf{f} \in \mathbb{R}^{N^2}} \|\mathbf{Mf} - \mathbf{g}\|_2^2 + \frac{\rho}{2} \left\|\mathbf{Sf} - \mathbf{v}^{(k)} + \mathbf{u}^{(k)}\right\|_2^2$
  $\mathbf{v}^{(k+1)} := \operatorname{argmin}_{\mathbf{v} \in \mathbb{R}^{\eta N^2}} \lambda \|\mathbf{v}\|_1 + \frac{\rho}{2} \left\|\mathbf{Sf}^{(k+1)} - \mathbf{v} + \mathbf{u}^{(k)}\right\|_2^2$
  $\mathbf{u}^{(k+1)} := \mathbf{u}^{(k)} + \mathbf{Sf}^{(k+1)} - \mathbf{v}^{(k+1)}$
  $k := k + 1$
until a convergence criterion is reached
$\mathbf{f} := \mathbf{f}^{(k)}$

The second step in equation (5.23) can be minimized by

$$\mathbf{v}^{(k+1)} := \mathcal{T}_{\frac{\lambda}{\rho}} \left(\mathbf{Sf}^{(k+1)} + \mathbf{u}^{(k)}\right),$$

see [19], where the componentwise soft-shrinkage operator $\mathcal{T}_\gamma \colon \mathbb{R}^{\eta N^2} \to \mathbb{R}^{\eta N^2}$ for a threshold $\gamma > 0$ is defined by

$$\left(\mathcal{T}_\gamma(\mathbf{x})\right)_k := \begin{cases} x_k - \gamma & \text{for } x_k > \gamma, \\ 0 & \text{for } x_k \in [-\gamma, \gamma], \\ x_k + \gamma & \text{for } x_k < -\gamma, \end{cases}$$

with $k = 1, 2, \ldots, \eta N^2$.

5.2.4 Shearlets on polar frequency grids

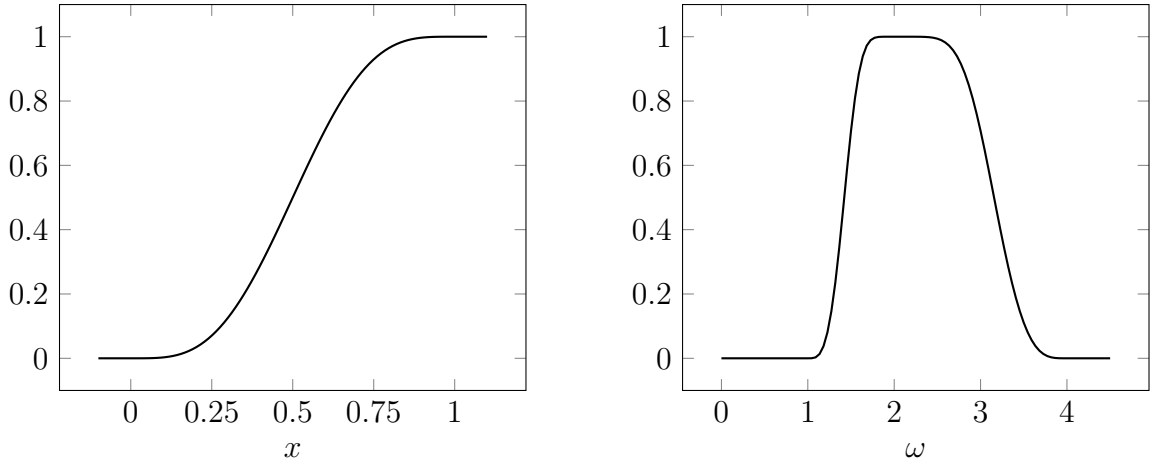
The presented Shearlets in the previous section are defined by samples on a Cartesian Fourier grid. While the appropriate shearing operation leads to deformations of the wedges, a more natural approach is given by using rotations instead of shearings. Actually, the continuous theory of Curvelets [9] is based on this fact, but the rotational invariance is not preserved in the associated discretization [8]. In the following, we suggest a modification of the Shearlet system [36, 35] for polar frequency grids, where rotations are also used in the discretization. This approach fits perfectly to the computation of mean values of functions, which are given by Fourier series with frequencies

on polar grids, see Section 4.2.2. Analog to [36], we start with auxiliary functions

$$v\colon \mathbb{R} \to \mathbb{R}, \quad v(x) := \begin{cases} 0 & \text{for } x < 0, \\ 35x^4 - 84x^5 + 70x^6 - 20x^7 & \text{for } 0 \le x \le 1, \\ 1 & \text{for } x > 1, \end{cases} \tag{5.25}$$

and

$$b\colon [0,\infty) \to \mathbb{R}, \quad b(r) := \begin{cases} \sin\left(\frac{\pi}{2} v\left(r-1\right)\right) & \text{for } 1 \le r \le 2, \\ \cos\left(\frac{\pi}{2} v\left(\frac{1}{2}r-1\right)\right) & \text{for } 2 < r \le 4, \\ 0 & \text{otherwise}, \end{cases} \tag{5.26}$$



(a) $v(x)$ (b) $b(x)$

Figure 5.3: Auxiliary functions v (left) and b (right) from equations (5.25) and (5.26).

see also Figure 5.3. The radial part is defined by dilations of a Wavelet,

$$\begin{aligned} \hat{\psi}_1\colon [0,\infty) \to \mathbb{R}, \quad \hat{\psi}_1(r) &:= \sqrt{b^2(2r) + b^2(r)}, \\ \hat{\psi}_{1,j}\colon [0,\infty) \to \mathbb{R}, \quad \hat{\psi}_{1,j}(r) &:= \hat{\psi}_1(4^{-j}r), \quad j \in \mathbb{N}, \end{aligned} \tag{5.27}$$

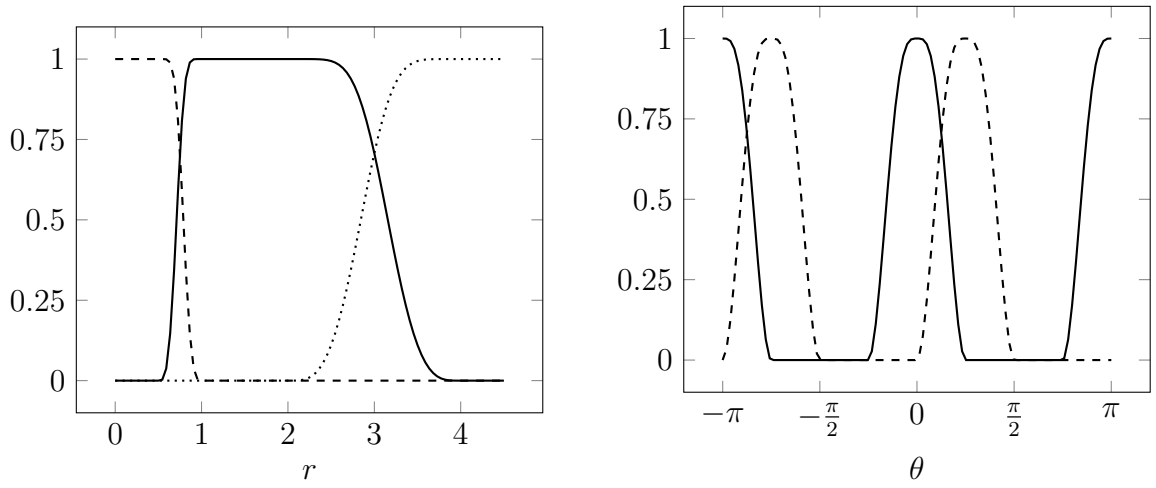
and for the lower frequencies we use

$$\hat{\varphi}\colon \mathbb{R} \to \mathbb{R}, \quad \hat{\varphi}(r) := \begin{cases} 1 & \text{for } 0 \le r \le \frac{1}{2}, \\ \cos\left(\frac{\pi}{2} v\left(2r-1\right)\right) & \text{for } \frac{1}{2} < r < 1, \\ 0 & \text{otherwise}. \end{cases} \tag{5.28}$$

Figure 5.4 shows this functions and the following Lemma 56 states some remarkable properties.

Lemma 56. *Let a discretization parameter $N \in 2\mathbb{N}$ be given. For $j \in \mathbb{N}$, the functions $\hat{\psi}_{1,j}$, see equation (5.27), fulfill*

$$\operatorname{supp} \hat{\psi}_{1,j} = \left[2^{2j-1}, 2^{2j+2}\right] \quad \textit{and} \quad \hat{\psi}_{1,j}(r) = 1 \quad \textit{for} \quad r \in [2^{2j}, 2^{2j+1}]$$



(a) $\hat{\varphi}(r)$ (dashed), $\hat{\psi}_1(r) = \hat{\psi}_{1,0}(r)$ (solid), and $\hat{\psi}_{1,1}(r)$ (dotted).

(b) $\hat{\psi}_{2,0}(\theta) = \hat{\psi}_{2,0,0}(\theta)$ (solid) and $\hat{\psi}_{2,0,1}(\theta)$ (dashed).

Figure 5.4: Wavelets for the radial part (left) and splines for the angular part (right) from equations (5.27), (5.28) and (5.29).

and for $r \in \left[0, \frac{N}{2}\right]$ *it follows*

$$|\hat{\varphi}(r)|^2 + \sum_{j=0}^{j_0-1} \left|\hat{\psi}_{1,j}(r)\right|^2 = 1, \quad j_0 := \frac{1}{2}\log_2 N.$$

Proof. The assertion follows from [36, Theorem 2.2]. □

We build the angular part with the splines

$$\hat{\psi}_{2,j} \colon \pi\mathbb{T} \to \mathbb{R}, \quad \hat{\psi}_{2,j}(\theta) := \begin{cases} \sqrt{v\left(1 + \frac{2^{j+2}\theta}{\pi}\right)} & \text{for} \quad \theta \leq 0, \\ \sqrt{v\left(1 - \frac{2^{j+2}\theta}{\pi}\right)} & \text{for} \quad \theta > 0, \end{cases}$$

and

$$\hat{\psi}_{2,j,k} \colon \pi\mathbb{T} \to \mathbb{R}, \quad \hat{\psi}_{2,j,k}(\theta) := \hat{\psi}_2\left(\theta - \frac{k\pi}{2^{j+2}}\right) + \hat{\psi}_2\left(\theta + \pi - \frac{k\pi}{2^{j+2}}\right), \tag{5.29}$$

where $j \in \mathbb{N}_0$ and $k \in \mathbb{Z}$. Similar to Lemma 56, we obtain the following statement.

Lemma 57. *For* $j \in \mathbb{N}_0$ *and* $\theta \in \pi\mathbb{T}$*, the functions* $\hat{\psi}_{2,j,k}$*, see equation* (5.29)*, fulfill*

$$\sum_{k=0}^{2^{j+2}-1} |\hat{\psi}_{2,j,k}(\theta)|^2 = 1.$$

Proof. The assertion follows from [36, Theorem 2.5]. □

Merging the radial and the angular parts leads to

$$\hat{\psi}_{j,k} \colon \pi\mathbb{T} \times [0, \infty) \to \mathbb{R}, \quad \hat{\psi}_{j,k}(\theta, r) := \hat{\psi}_{1,j}(r)\hat{\psi}_{2,j,k}(\theta) \tag{5.30}$$

and

$$\hat{\phi}\colon \pi\mathbb{T} \times [0,\infty) \to \mathbb{R}, \quad \hat{\phi}(\theta, r) := \hat{\varphi}(r),$$

which produces again a partition of unity.

Corollary 58. *For a discretization parameter $N \in 2\mathbb{N}$ and $(r,\theta) \in \left[0, \frac{N}{2}\right] \times \pi\mathbb{T}$ it follows*

$$\left|\hat{\phi}(\theta, r)\right|^2 + \sum_{j=0}^{j_0-1} \sum_{k=0}^{2^{j+2}-1} \left|\hat{\psi}_{j,k}(\theta, r)\right|^2 = 1, \quad j_0 = \frac{1}{2}\log_2 N.$$

Proof. The assertion is a direct consequence of Lemmas 56 and 57. □

Summing up, with equation (5.30), we obtain a function system, which is similar to the Shearlets system in [36], see also Figure 5.5 for a visualization. The use of polar coordinates causes the replacement of the shearing operations by rotations, which are more natural. Since the design of Curvelets [9] is based on this concept, we can also interpret the following procedure as an alternative discretization of Curvelets with polar frequencies. A straightforward discretization of (5.30) leads to the following definition.

Definition 59. Let discretization parameters $N \in 2\mathbb{N}$ and $M_1, M_3 \in \mathbb{N}$ be given. Analog to equation (4.11), for $s \in \{0, \dots, M_1 - 1\}$ and $l \in \{0, \dots, M_3 - 1\}$ we denote

$$\theta_s := \cos\frac{2\pi s}{M_1}, \quad r_l := \frac{lN}{2M_3}, \quad \text{and} \quad \boldsymbol{\xi}_{s,l} := r_l \begin{pmatrix} \cos\theta_s \\ \sin\theta_s \end{pmatrix},$$

and define the system of discrete polar Curvelets by

$$\hat{\boldsymbol{\Phi}}_{\mathbf{m}} := \left(\hat{\phi}\left(\theta_s, r_l\right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\xi}_{s,l} \cdot \left(\frac{\mathbf{m}}{N} + \frac{1}{2N}\right)}\right)_{s \in \{0,\dots,M_1-1\}, l \in \{0,\dots,M_3-1\}} \in \mathbb{C}^{M_1 \times M_3},$$

and

$$\hat{\boldsymbol{\Psi}}_{j,k,\mathbf{m}} := \left(\hat{\psi}_{j,k}\left(\theta_s, r_l\right) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{\xi}_{s,l} \cdot \left(\frac{\mathbf{m}}{N} + \frac{1}{2N}\right)}\right)_{s \in \{0,\dots,M_1-1\}, l \in \{0,\dots,M_3-1\}} \in \mathbb{C}^{M_1 \times M_3}$$

with

$$j \in \left\{0, \dots, \frac{1}{2}\log_2 N - 1\right\}, \quad k \in \left\{0, \dots, 2^{j+2} - 1\right\}, \quad \text{and} \quad \mathbf{m} \in J_N.$$

Finally, the discrete polar Curvelet coefficients of a vector $\hat{\mathbf{f}} \in \mathbb{C}^{M_1 M_3}$ are given by

$$c_{\mathbf{m}}\left(\hat{\mathbf{f}}\right) := \left\langle \hat{\mathbf{f}}, \hat{\boldsymbol{\Phi}}_{\mathbf{m}} \right\rangle \quad \text{and} \quad c_{j,k,\mathbf{m}}\left(\hat{\mathbf{f}}\right) := \left\langle \hat{\mathbf{f}}, \hat{\boldsymbol{\Psi}}_{j,k,\mathbf{m}} \right\rangle.$$

□

For the implementation of the polar Curvelet transform from Definition 59 we fix $\mathbf{m} \in J_N$ and consider

$$\begin{aligned} \left\langle \hat{\mathbf{f}}, \hat{\boldsymbol{\Phi}}_{\mathbf{m}} \right\rangle &= \sum_{s=0}^{M_1-1} \sum_{l=0}^{M_3-1} \hat{f}_{s,l} \left(\hat{\boldsymbol{\Phi}}_{\mathbf{m}}\right)_{s,l} = \sum_{s=0}^{M_1-1} \sum_{l=0}^{M_3-1} \hat{f}_{s,l} \left(\hat{\boldsymbol{\Phi}}_{\mathbf{0}}\right)_{s,l} \mathrm{e}^{-\frac{2\pi \mathrm{i}}{N} \boldsymbol{\xi}_{s,l} \cdot \left(\mathbf{m} + \frac{1}{2}\right)} \\ &= \sum_{s=0}^{M_1-1} \sum_{l=0}^{M_3-1} \left(\hat{f}_{s,l} \left(\hat{\boldsymbol{\Phi}}_{\mathbf{0}}\right)_{s,l} \mathrm{e}^{-\frac{\pi \mathrm{i}}{N} \boldsymbol{\xi}_{s,l} \cdot \mathbf{1}}\right) \mathrm{e}^{2\pi \mathrm{i} \left(-\frac{\boldsymbol{\xi}_{k,l}}{N}\right) \cdot \mathbf{m}}. \end{aligned}$$

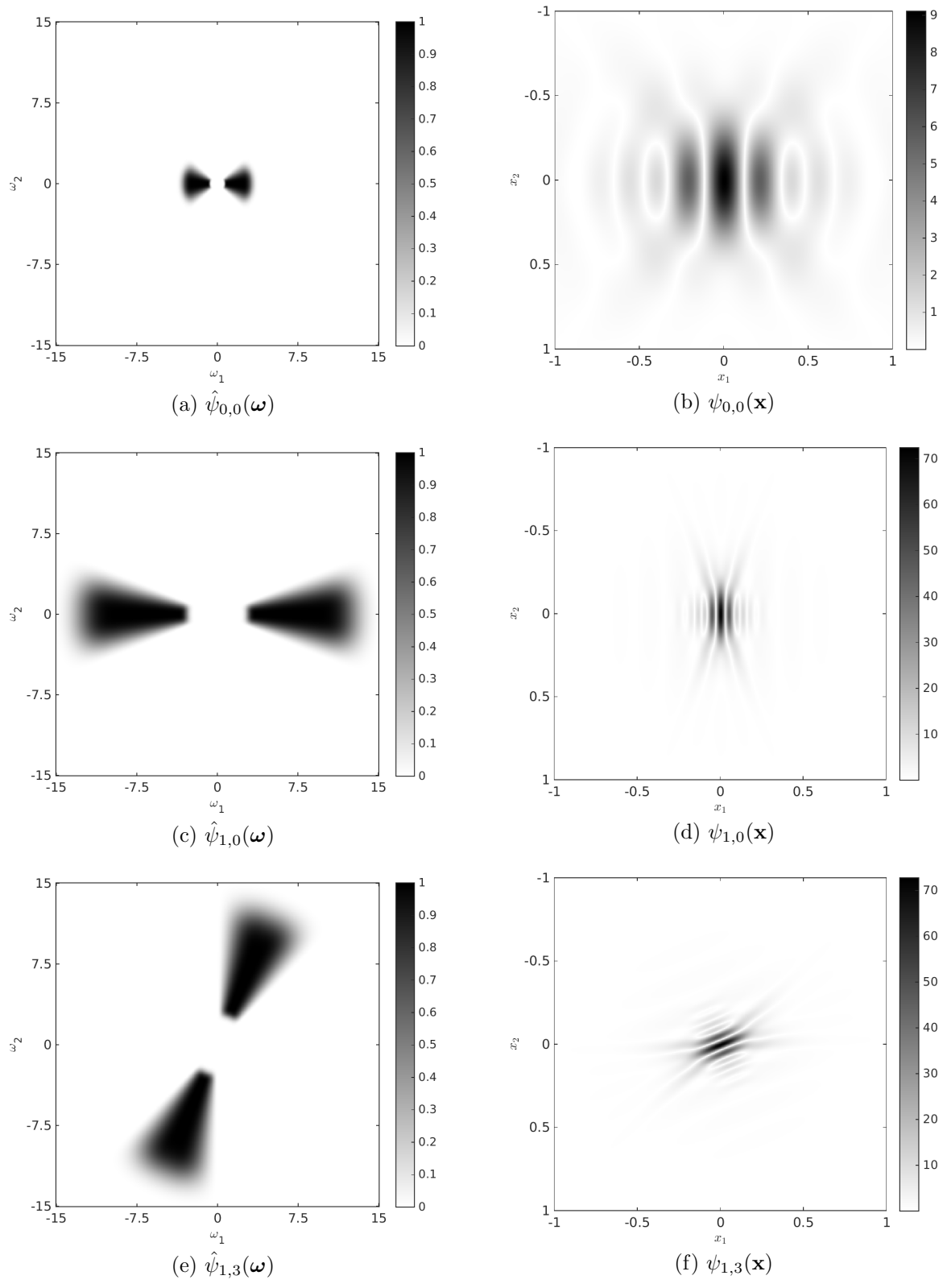
For all $\mathbf{m} \in J_N$, this can efficiently be computed by an adjoint nonequispaced fast Fourier transform [43]. For the remaining Curvelets $\mathbf{\Psi}_{j,k,\mathbf{m}}$, exactly the same procedure leads to an adjoint NFFT for each (j,k). For fixed $(s,l) \in \{0,\dots,M_1-1\}\times\{0,\dots,M_3-1\}$, the transpose operator is given by

$$\begin{aligned}\hat{a}_{s,l} &:= \sum_{\mathbf{m}\in J_N} c_{\mathbf{m}} \left(\hat{\mathbf{\Phi}}_{\mathbf{m}}\right)_{s,l} = \sum_{\mathbf{m}\in J_N} c_{\mathbf{m}} \left(\hat{\mathbf{\Phi}}_{\mathbf{0}}\right)_{s,l} \mathrm{e}^{-\frac{2\pi \mathrm{i}}{N}\boldsymbol{\xi}_{s,l}\cdot\left(\mathbf{m}+\frac{1}{2}\right)} \\ &= \left(\hat{\mathbf{\Phi}}_{\mathbf{0}}\right)_{s,l} \mathrm{e}^{-\frac{\pi \mathrm{i}}{N}\boldsymbol{\xi}_{s,l}\cdot\mathbf{1}} \overline{\sum_{\mathbf{m}} \overline{c_{\mathbf{m}}} \mathrm{e}^{-2\pi \mathrm{i}\left(-\frac{\boldsymbol{\xi}_{s,l}}{N}\right)\cdot\mathbf{m}}}\end{aligned}$$

and the application of a NFFT produces the entire vector $\mathbf{a} \in \mathbb{R}^{M_1 M_3}$. If we perform a NFFT for the remaining blocks $c_{j,k,:}$ and denote the output by $\mathbf{a}_{j,k}$, then the transpose Curvelet transform of coefficients $\mathbf{c}$ is given by

$$\mathbf{a} + \sum_{j=0}^{j_0-1} \sum_{k=0}^{2^{j+2}-1} \mathbf{a}_{j,k}.$$

Remark 60. *Since Parseval's identity does not hold for the polar frequencies, the derived Curvelet system provides no Parseval frame and consequently, in contrast to the Shearlet system in [36, section 3.3], the inverse transform is not given by the transpose operator and requires a more sophisticated approach.*



(a) $\hat{\psi}_{0,0}(\boldsymbol{\omega})$ (b) $\psi_{0,0}(\mathbf{x})$

(c) $\hat{\psi}_{1,0}(\boldsymbol{\omega})$ (d) $\psi_{1,0}(\mathbf{x})$

(e) $\hat{\psi}_{1,3}(\boldsymbol{\omega})$ (f) $\psi_{1,3}(\mathbf{x})$

Figure 5.5: The mother Curvelet (first row) and scaled and rotated versions (second and third row). The left column shows the Fourier transforms and on the right are the appropriate functions in the spatial domain.

6

Implementation

In the previous chapters we presented different algorithms for an efficient computation and inversion of mean value operators. The implementation of these methods is an essential part of this thesis and the result is a Matlab toolbox, that contains the derived algorithms as well as thematically related work. The software is designed to enable the improvement and extension potentially by multiple developers, realized by thematically splitting of the code and organizing of the files in clear structured directories. Further features are the verification of input data and the output of meaningful error messages. Since the considered operators are linear, a theoretically realization of the algorithms is given by the creation of full matrices. This concept involves a well-known handling of the operators with the built-in capabilities of matrix operations in Matlab, but is also limited by the amount of the main memory and therefore not qualified for typical applications. Indeed, the previous algorithms are formulated with other techniques, but in order to provide a simple usage of the toolbox, the implementation is designed to support relevant methods from Matlab for working with matrices, for example `mtimes`, `mldivide`, and `transpose`. As consequence, many commands are similar to the matrix operations in Matlab and independent of the specific realization.
The developed toolbox is implemented by object-oriented programming in Matlab, which is based on structuring the code into classes and complies perfectly with the desired concept. Figure 6.1 shows the hierarchical structured division of the algorithms into classes. The file names are chosen to be as much as possible self-explanatory and numbers indicates the spatial dimension. Many examples for the different settings are given by the included scripts starting with `test`, and documentation can be obtained with the `help` or `doc` command.
Due to the object-oriented design, other existing projects with compatible licenses can efficiently be integrated into the toolbox with the possibility for continuing developments. Our toolbox [27] is currently available on

`http://torstengoerner.de/software/`

under the GNU General Public License version 3 as published by the Free Software Foundation [20].
We proceed with a short overview to object-oriented programming, and describe the main concepts of the different classes in the toolbox afterwards.

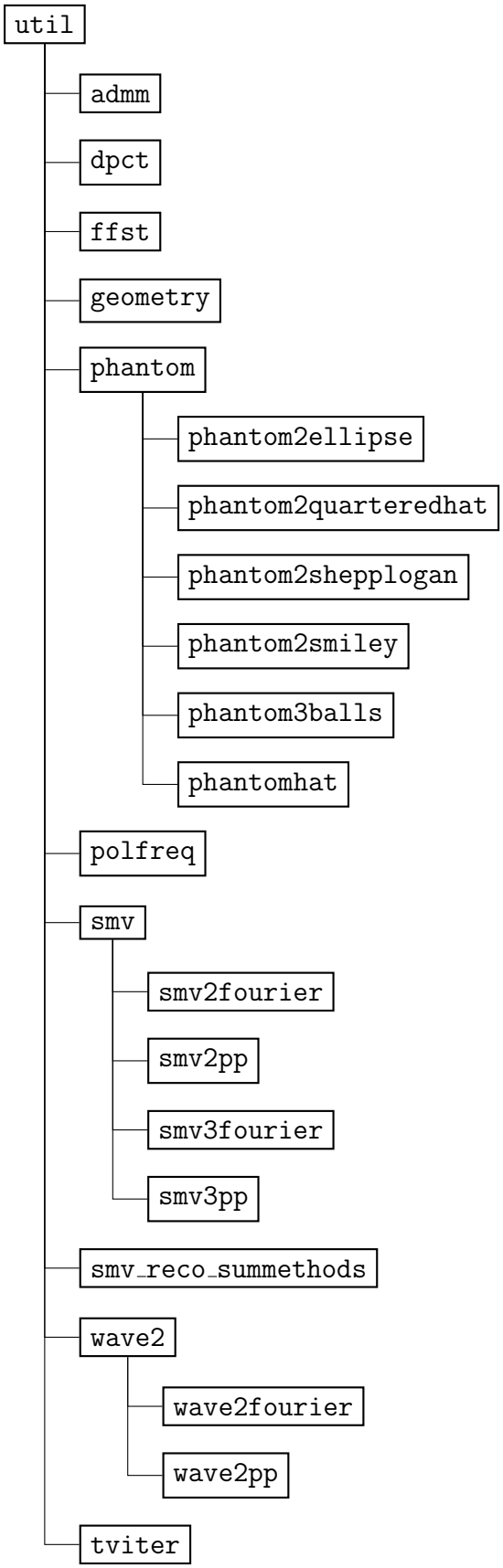


Figure 6.1: Hierarchical structure of the implemented classes in the Matlab toolbox.

6.1 Object-oriented programming in Matlab

Object-oriented programming is a model, which describes the connections between the algorithms and the data by objects, in most popular languages so-called instances of classes. This approach confirms basically with the human mind and produces a clear structured code with an improved readability. It allows for a manageable realization of complex problems and provides simplified maintenance and extension of projects. This structured framework leads also to a consistent modelling of the problems and implies a higher independence of the specific hardware and software environment. However, the overhead from the organization of the complete data structure can cause some additional computational effort and might result in a lower performance.
We continue with a short introduction to the workflow of object-oriented programming in Matlab and give some examples from our toolbox. The basis builds the definition of classes, which specifies mainly the properties, the constructor, the destructor and further methods. For utilizing a class, we have to create an instance, also called object, by invoking the constructor, which has the same name as the class. Each data structure, which results from this procedure, is characterized by the assigned values in the properties. Afterwards, we can work with these objects by applying methods from the class definition. Here, the specific processing and output depends on the set properties in the object. In order to delete objects, in some cases it is necessary to execute additional commands. For this purpose, we can define a destructor, which is automatically called during the delete process and implemented by the method `delete` in Matlab.
For instance, an object of the `geometry` class for spatial dimension $d = 2$ with discretization parameter $N = 64$ is generated by

```
obj=geometry(2,'N',64).
```

We can retrieve the properties for the spatial dimension d or the discretization parameter N by calling

```
d=obj.d   or   N=obj.N
```

and the current grid can be obtained by invoking the method

```
[X1,X2]=mk_grid(obj).
```

Finally, we point out concepts in Matlab, which improve the organization of the code and are also used in the developed toolbox.
For complex projects, it is advantageous to split the code into files and folders, which can be realized by using @-folders. This allows to put all files for a class into an own folder, which has the same name as the class starting with the symbol @, and offers the possibility to save methods in separate files. Instead of adding all needed @-folders to the Matlab path, we have only to ensure, that their parent folders are on the path.
It is also possible to generate a hierarchy of classes to avoid duplicating code. The main idea is to define a superclass as root and further subclasses for a more specific modelling of the process. A subclass inherits all properties and methods from the superclass, but can also have some extra features. This procedure can recursively be continued, such that a subclass can also act as a superclass for deeper subclasses. In the example of our

toolbox, we implemented a `phantom` superclass with general functionalities for all test functions and the details for each particular test functions, mainly specific properties, are implemented by appropriate subclasses.

6.2 Utility and geometry class

All classes of the toolbox are derived from the superclass `util`, which collects some auxiliary tools, that are needed by almost all components of the library. For example the implemented `size` acts as an analogon to the Matlab built-in and provides a similar interface. There are also other equivalents to basic Matlab operations such as `transpose`, `ctranspose` and `conj` included.

Another fundamental part of the toolbox represents the `geometry` class, which organizes all information about the specific setting, for example discretization parameters, grid types, arrangement of the center points and radii, and visualization options. Such a `geometry` object is used to create efficiently instances of the other classes and avoids the repeated specification of the used setting. One useful method is `mk_grid`, which outputs the coordinates of the points in the current grid.

6.3 Test functions

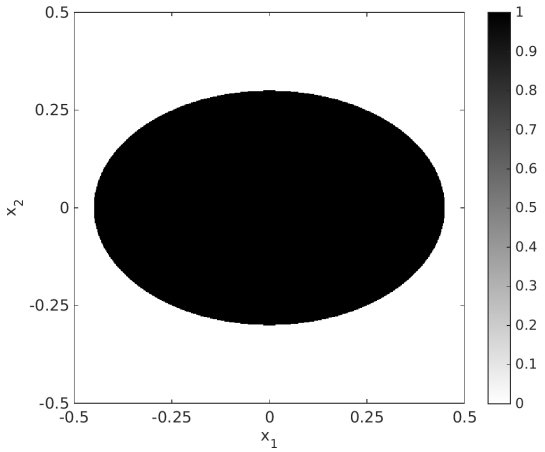
The Matlab toolbox contains multiple test functions, shown in Figure 6.2. They are suitable in particular for medical imaging and implemented as subclasses of the superclass `phantom`. The hat function, which is shown in Figure 6.2f and defined in equation (7.1), is a radial function with controllable smoothness and the quarteredhat phantom in Figure 6.2b is a modification with added singularities. Another variant is given by the balls phantom for spatial dimension $d = 3$ in Figure 6.2e, created by a composition of piecewise constant hat functions.

The ellipse test function represents a constant function with support inside of an ellipse and the summation of ten such objects with specified parameters [61, Table B.1] leads to the Shepp-Logan phantom in Figure 6.2c.

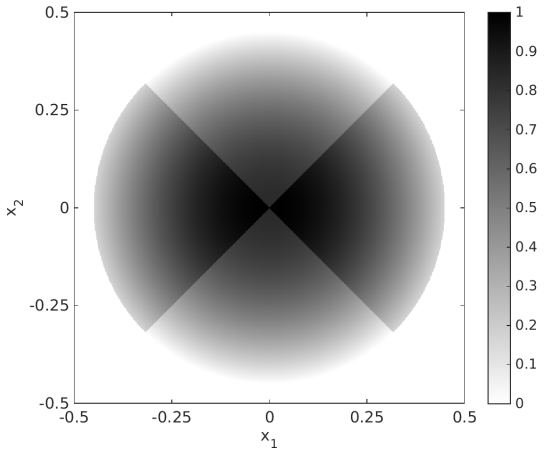
Figure 6.2d shows a further piecewise constant test function, the laughing Smiley phantom. In addition to smooth edges, it contains also some singularities at corners, whose reconstruction is a challenging task.

The `phantom` class provides mainly three methods. The function samples on the current grid in the geometry object can be computed by the method `full` and the method `visualise` produces a graphical output. In addition, the method `means` implements for some test functions the analytically computation of mean values. This is realized by the solution of quartic equations for the class `phantom2ellipse` and by the considerations in Examples 62 and 63 for the class `phantomhat`.

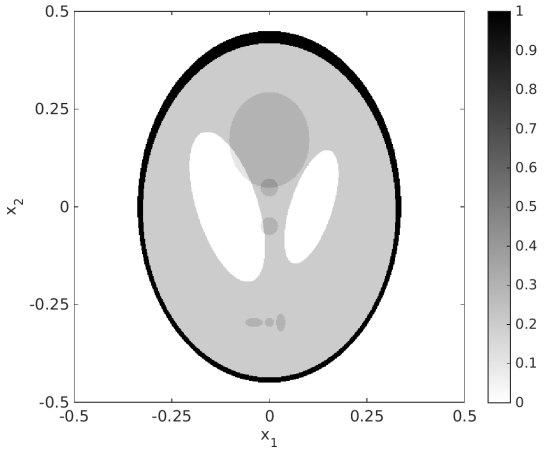
Depending on the test function, there are also other methods available or can be implemented at a later time. For instance, the class `phantomhat` provides also methods for the computation of the Fourier coefficients and of mean values from the Fourier partial sum, which is based on equation (7.2).



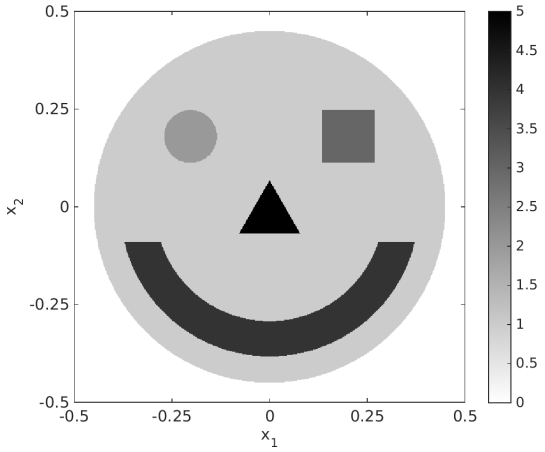
(a) `phantom2ellipse`



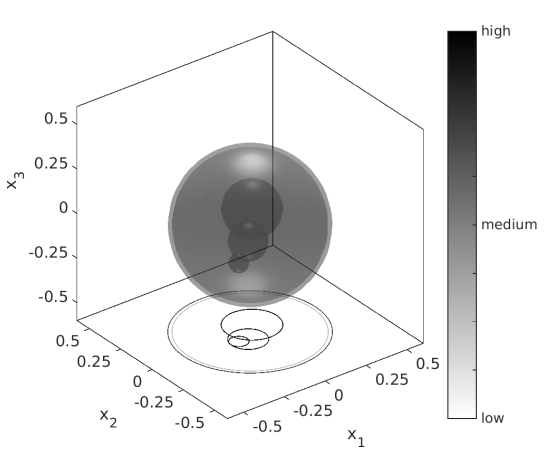
(b) `phantom2quarteredhat`



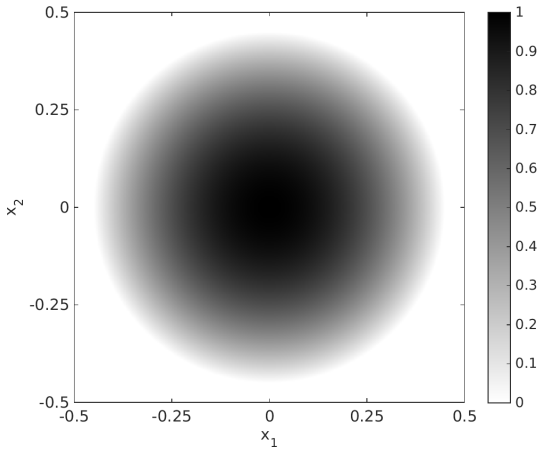
(c) `phantom2shepplogan`



(d) `phantom2smiley`



(e) `phantom3balls`



(f) `phantomhat` ($d = 2$)

Figure 6.2: Available test functions of the Matlab toolbox, implemented as subclasses of the superclass `phantom`.

6.4 Forward operators

In Chapter 4, we introduced discretizations of mean value operators. The algorithms are implemented as the superclasses `smv` for the spherical mean value operator $\mathcal{M}$ and `wave2` for the mean value operator $\mathcal{N}$, where the different variants are realized by all the subclasses. An appropriate object behaves like a matrix, in particular the Matlab functions `conj`, `ctranspose`, `transpose` and `size` can be used in the usual way, `full` outputs the matrix and the evaluation of the operator is provided by the `mtimes` method.
Since the Fourier based approaches make use of specific fast Fourier transforms, non-equispaced fast Fourier transform [43] and butterfly sparse fast Fourier transforms [45], it is neccessary for these discretization variants to install further software libraries and assign the according directories to the property `libdir`.

6.5 Reconstructions

Currently, the toolbox contains in addition to a simple least squares method three variants for the reconstruction of function samples from mean values. The code is organized in three classes, `tviter` for the generalized Newton method with total variation based regularization, see Section 5.1, `admm` for the alternating direction method of multipliers with Shearlet based regularization, see Section 5.2, and `smv_reco_summethods` for the summability methods, see [60, 17]. In fact, the user has not to deal with these classes directly. The implementation is motivated by the procedure for the solution of systems of linear equations and can be invoked by the `mldivide` method from instances of the classes for the forward operators in Section 6.4. These instances have also properties for specifying the particular setting. More precise, the property `rec_method` manages the reconstruction method and specific parameters, for example regularization parameters or error tolerances, are assigned by structure arrays in the properties `param_lsqr`, `param_tv_iter`, `param_summethods`, and `param_admm_shear` respectively. The class `dpct` implements the Curvelet system, which we proposed in Section 5.2.4.

6.6 Auxiliary classes

To complete this description of the toolbox, we give also some remarks to the remaining classes. They are needed by other classes and were originally not intended to be used directly, but they are also applicable for related problems.
The `ffst` class forms an object-oriented interface to the FFST library [36, 35], which is consistent with the concept of our toolbox. However, the code is not reimplemented and we still use the original code, which requires the additional installation of the FFST library and the declaration of the according directory in the property `ffstdir`. Currently, our toolbox supports version 2.0 of the FFST library.
Another auxiliary tool is the `polfreq` class. This part of the library represents the operator for the evaluation of a Fourier sum with polar frequencies, realized by non-equispaced fast Fourier transforms [43]. As for the `ffst` class, this requires also an additional installation of the NFFT library and the specification of the according directory in the property `nfftdir`.

7

Numerical experiments

After the derivation and theoretical analysis of the discretization and reconstruction algorithms in the previous chapters, we examine in the following their numerical properties. All algorithms were implemented in Matlab and the experiments were performed on a Dell PowerEdge R900-rack-server, equipped with Intel Xeon E7450 processors (2.40 GHz) and 94 GByte main memory.
The numerical experiments are grouped into three parts. For spatial dimensions $d = 2$ and $d = 3$, we start with the analysis of the discretizations from Chapter 4. We consider for the mean value operator $\mathcal{M}$ the error behavior and measure the increase of the running times for a typical setting afterwards. The experiments regarding the running times are also done for the mean value operator $\mathcal{N}$ in spatial dimension $d = 2$. Since we are currently not aware of a suitable test function, whose mean values for the operator $\mathcal{N}$ are known analytically, there are no error tests for this operator. Finally, we demonstrate the capabilities of the reconstruction methods from Chapter 5. The presented numerical experiments in Sections 7.1 and 7.2 are published in [25, 26].

7.1 Error analysis

We start by defining some test functions. Let a size parameter $t \in (0, 0.5]$, a smoothness parameter $s \in \mathbb{N}_0$, and the radial test function $f_{d,s,t} \colon \mathbb{R}^d \to \mathbb{R}$,

$$f_{d,s,t}(\mathbf{x}) := \begin{cases} \varphi_{s,t}\left(|\mathbf{x}|^2\right) & \text{for } |\mathbf{x}| \leq t, \\ 0 & \text{otherwise,} \end{cases} \quad \varphi_{s,t} \colon [0,t] \to \mathbb{R}, \quad \varphi_{s,t}(\tau) := \left(1 - \frac{\tau}{t^2}\right)^s, \tag{7.1}$$

be given, see also Figure 7.1. We compute spherical means of these test functions analytically by using the following statement for radial functions.

Lemma 61. *Let a spatial dimension $d \in \mathbb{N}$, $d \geq 2$, a center point $\mathbf{y} \in \mathbb{R}^d$, a radius $r \geq 0$, and a radial function $f \colon \mathbb{R}^d \to \mathbb{R}$, $f(\mathbf{x}) = \varphi\left(|\mathbf{x}|^2\right)$, for some function $\varphi \colon [0, \infty) \to \mathbb{R}$ be given. Then the spherical mean values of the function f can be computed by*

$$\mathcal{M}f(\mathbf{y}, r) = \frac{\omega_{d-2}}{\omega_{d-1}} \int_{-1}^{1} \varphi\left(|\mathbf{y}|^2 + r^2 + 2r|\mathbf{y}|\tau\right) (1 - \tau^2)^{\frac{d-3}{2}} \mathrm{d}\tau.$$

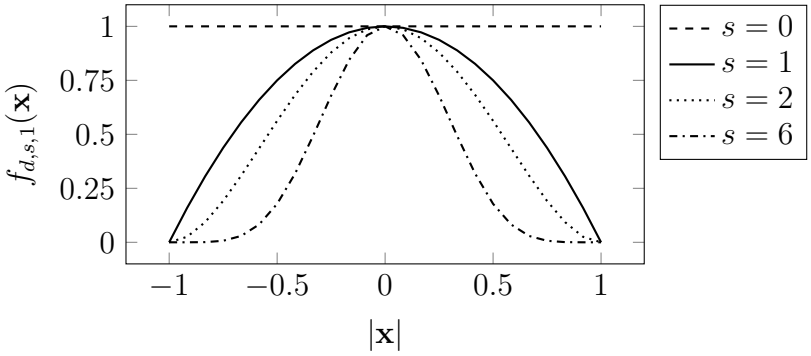


Figure 7.1: Test function $f_{d,s,t}$ from equation (7.1) for size parameter $t = 1$, and different smoothness parameters s.

Proof. First, we assume without loss of generality $\mathbf{y} = |\mathbf{y}|\mathbf{e}_d$, where

$$\mathbf{e}_d := (0, \dots, 0, 1)^\top \in \mathbb{R}^d.$$

Then

$$\begin{aligned}
\mathcal{M}f(\mathbf{y}, r) &= \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} \varphi\left(|\mathbf{y} + r\xi|^2\right) \mathrm{d}\sigma(\boldsymbol{\xi}) \\
&= \frac{1}{\omega_{d-1}} \int_{\mathbb{S}^{d-1}} \varphi\left(|\mathbf{y}|^2 + r^2 + 2r|\mathbf{y}|\boldsymbol{\xi} \cdot \mathbf{e}_d\right) \mathrm{d}\sigma(\boldsymbol{\xi}) \\
&= \frac{1}{\omega_{d-1}} \int_{-1}^{1} \int_{\mathbb{S}^{d-2}} \varphi\left(|\mathbf{y}|^2 + r^2 + 2r|\mathbf{y}|\tau\right) \mathrm{d}\sigma\left(\tilde{\boldsymbol{\xi}}\right) (1 - \tau^2)^{\frac{d-3}{2}} \mathrm{d}\tau.
\end{aligned}$$

□

Next we use the above lemma to compute the spherical mean values of some of the test functions (7.1) explicitly.

Example 62. Let us start with spatial dimension $d = 2$. Then we have

$$\mathcal{M}f_{2,s,t}(\mathbf{y}, r) = \frac{1}{\pi} \int_0^{\vartheta_0} \varphi_{s,t}\left(|\mathbf{y}|^2 + r^2 - 2r|\mathbf{y}| \cos\vartheta\right) \mathrm{d}\vartheta,$$

where

$$\vartheta_0 := \vartheta_0\left(|\mathbf{y}|, r, t\right) := \begin{cases} \pi & \text{for} \quad t^2 \geq (|\mathbf{y}| + r)^2, \\ 0 & \text{for} \quad t^2 < (|\mathbf{y}| - r)^2, \\ \arccos \frac{|\mathbf{y}|^2 + r^2 - t^2}{2r|\mathbf{y}|} & \text{else.} \end{cases}$$

Fixing the smoothness parameter $s \in \mathbb{N}_0$ gives an explicit solution of the integral. With adequate coefficients $b_{s,k} := b_{s,k}(\mathbf{y}, r, t) \in \mathbb{R}$, $k = 0, \dots, s$, we have

$$\varphi_{s,t}\left(|\mathbf{y}|^2 + r^2 - 2r|\mathbf{y}| \cos\vartheta\right) = \sum_{k=0}^{s} b_{s,k} (\cos\vartheta)^k$$

and it follows

$$\mathcal{M}f_{2,s,t}(\mathbf{y},r) = \frac{1}{\pi}\sum_{k=0}^{s} b_{s,k}\int_0^{\vartheta_0}(\cos\vartheta)^k \mathrm{d}\vartheta.$$

For example $s=0$ yields the coefficient $b_{0,0}=1$ and $s=1$ yields the coefficients $b_{1,0}=\frac{t^2-|\mathbf{y}|^2-r^2}{t^2}$ and $b_{1,1}=\frac{2r|\mathbf{y}|}{t^2}$. Hence, we obtain

$$\mathcal{M}f_{2,0,t}(\mathbf{y},r) = \frac{\vartheta_0}{\pi} \quad \text{and}$$

$$\mathcal{M}f_{2,1,t}(\mathbf{y},r) = \frac{1}{\pi}\left(b_{1,0}\cdot\vartheta_0 + b_{1,1}\cdot\sin\vartheta_0\right) = \frac{\vartheta_0\cdot(t^2-|\mathbf{y}|^2-r^2)+2r|\mathbf{y}|\sin\vartheta_0}{\pi t^2}.$$

□

Example 63. Since the weight disappears for dimension $d=3$ we define a primitive $\Phi_{s,t}\colon\mathbb{R}\to\mathbb{R}$ of $\varphi_{s,t}$ and an auxiliary quantity $\tau_0\in\mathbb{R}_+$,

$$\Phi_{s,t}(\tau) := -\frac{(t^2-\tau)^{s+1}}{t^{2s}(1+s)}, \qquad \tau_0 := \begin{cases} (|\mathbf{y}|+r)^2 & \text{for} \quad t^2 \geq (|\mathbf{y}|+r)^2, \\ (|\mathbf{y}|-r)^2 & \text{for} \quad t^2 < (|\mathbf{y}|-r)^2, \\ t^2 & \text{otherwise,} \end{cases}$$

and obtain for $\mathbf{y}\neq\mathbf{0}$ and $r\neq 0$ the closed-form expression

$$\mathcal{M}f_{3,s,t}(\mathbf{y},r) = \frac{1}{4r|\mathbf{y}|}\cdot\left(\Phi_{s,t}(\tau_0) - \Phi_{s,t}\left((|\mathbf{y}|-r)^2\right)\right).$$

For example with parameters $s=0$ and $s=1$ we get the formulas

$$\mathcal{M}f_{3,0,t} = \frac{\tau_0-(|\mathbf{y}|-r)^2}{4r|\mathbf{y}|}, \qquad \mathcal{M}f_{3,1,t} = \frac{\left(t^2-(|\mathbf{y}|-r)^2\right)^2-(t^2-\tau_0)^2}{8r|\mathbf{y}|t^2}.$$

□

Moreover, Sonine's integral [62, Chapter 12.11] yields the Fourier transform of these test functions

$$\hat{f}_{d,s,t}(\mathbf{z}) = \frac{\Gamma(s+1)t^{\frac{d}{2}-s}}{\pi^s}\cdot\frac{\mathcal{J}_{s+\frac{d}{2}}(2\pi|\mathbf{z}|t)}{|\mathbf{z}|^{s+\frac{d}{2}}}, \tag{7.2}$$

which, in conjunction with [62, Sec. 7.21], allows for the estimate

$$\left|\hat{f}_{d,s,t}(\mathbf{z})\right| \leq C_{d,s,t}\cdot(1+|\mathbf{z}|^2)^{-\frac{1}{2}\left(s+\frac{d+1}{2}\right)}$$

and thus $f_{d,s,t}\in\mathcal{H}_{s+\frac{1}{2}-\varepsilon}(\mathbb{T}^d)$ for every $\varepsilon>0$. Theorems 32 and 33 imply error bounds

$$|\mathcal{M}f_{d,s,t}(\mathbf{y},r) - \mathcal{M}\mathcal{I}_N f_{d,s,t}(\mathbf{y},r)| \leq C_{d,s,t,f,\varepsilon}\cdot N^{\frac{d}{2}-s-\frac{1}{2}+\varepsilon}$$

and

$$|\mathcal{M}f_{d,s,t}(\mathbf{y},r) - \mathcal{M}\mathcal{S}_N f_{d,s,t}(\mathbf{y},r)| \leq C_{d,s,t,f,\varepsilon}\cdot N^{-s+\varepsilon},$$

which are supported by the following numerical results.
Let the size parameter $t = 0.2$ be fixed and a set of discretization parameters $N \in 2\mathbb{N}$ and smoothness parameters $s \in \mathbb{N}_0$ be given. For spatial dimensions $d \in \{2, 3\}$, we consider settings with $M_1 = N^{d-1}$ center points, which are arranged on a circular line ($d = 2$) or a surface of a sphere ($d = 3$), and $M_2 = N$ equispaced radii $r_1, \dots, r_N \in (0.46]$. In the numerical experiments, we compute the mean values $\mathbf{Mf}$ of the test function, given by samples

$$\mathbf{f} := \left(f_{d,s,t}\left(\mathbf{x}\right)\right)_{\mathbf{x}\in X_N} \in \mathbb{R}^{N^d},$$

with different discretizations $\mathbf{M} \in \mathbb{R}^{N^d \times N^d}$ of the mean value operator $\mathcal{M}$ from Section 4 and measure the absolute error

$$\varepsilon_{\text{abs}} := \varepsilon_{\text{abs}}(N, s, d) := \max_{\substack{j\in\{1,\dots,N^{d-1}\}\\ k\in\{1,\dots,N\}}} \left|\mathcal{M} f_{d,s,t}\left(\mathbf{y}_j, r_k\right) - \left(\mathbf{Mf}\right)_{j,k}\right|.$$

We utilize the quadrature based approach with piecewise constant ansatz functions and the application of a simple rectangular rule, see Algorithms 1 ($d = 2$) and 2 ($d = 3$), and Algorithm 4, that computes the mean values of the trigonometric interpolant $\mathcal{I}_N$. Additionally, we consider an approximation with bilinear functions and the application of a trapezoidal rule, see Remark 22, and a variant that uses the exact Fourier coefficients (7.2) of our test functions and implements $\mathcal{M}\mathcal{S}_N f_{d,s,t}$.

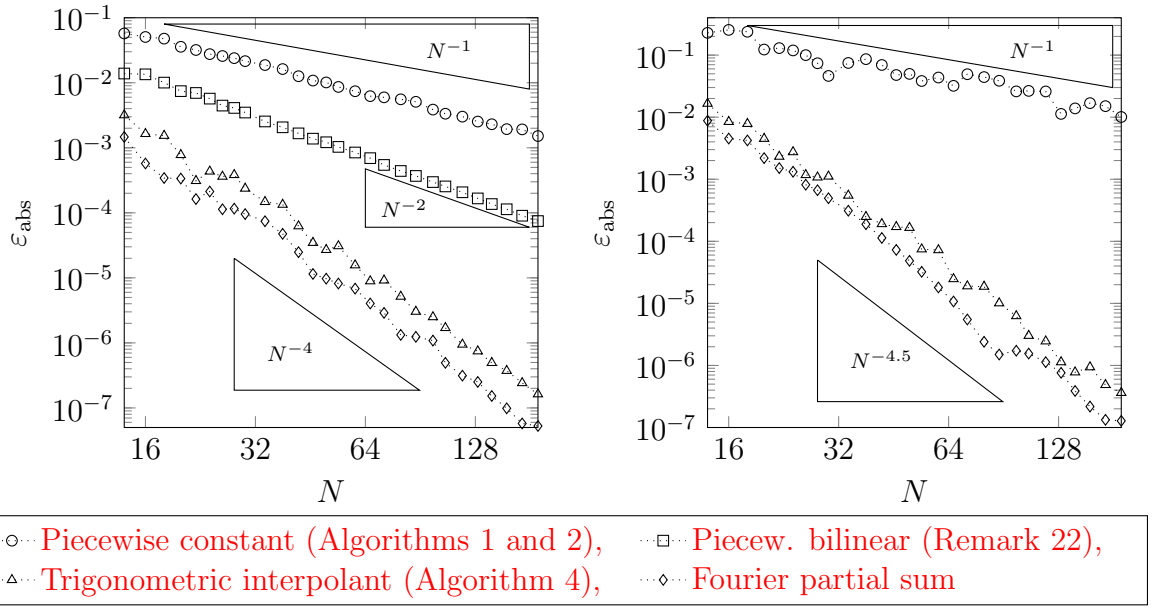


Figure 7.2: Accuracy with respect to the discretization parameter N for spatial dimension $d = 2$ (left) and $d = 3$ (right), and test functions $f_{d,s,t}$ with smoothness parameter $s = 3$.

For fixed smoothness parameter $s = 3$, Figure 7.2 illustrates the behavior of the error ε_{abs} with respect to the discretization parameter N. We observe, that the convergence rate for the considered test function is even better than the theoretical bound proven in Theorem 33. We suspect that this is related to the fact, that the Fourier coefficients of $f_{s,t,d}$ oscillate. Table 7.1 summarizes for further smoothness parameters $s \in \{0, \dots, 6\}$ the estimates of the convergence rates obtained from a least squares fit.

	$d=2$				$d=3$		
s	p. const.	p. bil.	$\mathcal{I}_N$	$\mathcal{S}_N$	p. const.	$\mathcal{I}_N$	$\mathcal{S}_N$
0	0.76	0.60	0.83	0.88	1.10	1.12	1.37
1	1.25	1.55	1.79	1.97	1.10	2.03	2.45
2	1.34	2.04	2.86	3.03	1.06	3.23	3.29
3	1.42	2.06	3.78	3.89	1.10	4.21	4.49
4	1.37	2.00	4.84	5.02	1.14	5.40	5.41
5	1.32	1.96	5.75	5.91	1.18	6.42	6.57
6	1.30	1.95	6.68	6.92	1.24	7.40	7.49
conjecture	1	2	$s+1$	$s+1$	1	$s+1.5$	$s+1.5$
theory	–	–	$s-0.5$	s	–	$s-1$	s

Table 7.1: Estimated orders $-\log \varepsilon_{\text{abs}}/\log N$ of convergence with respect to the smoothness parameter s. These are derived by a least square fit of the computed errors in Figure 7.2.

7.2 Running times

For comparing the running times, we consider analog to Section 7.1 for different discretization parameters $N \in 2\mathbb{N}$ and a spatial dimensions $d \in \{2,3\}$ a setting with $M_1 = N^{d-1}$ center points and $M_2 = N$ radii. We measure the computation time for the evaluation of mean values $\mathbf{Mf} \in \mathbb{R}^{N^2}$ from samples $\mathbf{f} \in \mathbb{R}^{N^2}$ with different discretizations $\mathbf{M} \in \mathbb{R}^{N^2 \times N^2}$ of the mean value operator $\mathcal{M}$ from Section 4. Focusing on the running times, we have the following classification of the discretization variants. For dimensions $d \in \{2,3\}$, the used algorithms are based on applying a simple rectangular rule, see Algorithms 1 and 2, and nonequispaced fast Fourier transforms, see Algorithm 4. Additionally, for spatial dimension $d=2$, we utilize an evaluation of the integral with a trapezoidal rule, see Remark 22, and for spatial dimension $d=3$, we consider a variant, which is based on sparse fast Fourier transforms, see Algorithm 6.

	$d=2$		$d=3$
	Input/Output: $\mathcal{O}(N^2)$		Input/Output: $\mathcal{O}(N^3)$
	$\mathcal{M}$	$\mathcal{N}$	$\mathcal{M}$
Rectangular rule	$\mathcal{O}(N^3)$	$\mathcal{O}(N^4)$	$\mathcal{O}(N^5)$
Trapezoidal rule	$\mathcal{O}(N^3)$	–	–
NFFT	$\mathcal{O}(N^3 \log N)$	$\mathcal{O}(N^3 \log N)$	$\mathcal{O}(N^4 \log N)$
Sparse FFT	–	$\mathcal{O}\left((\log N)^5 N^2\right)$	$\mathcal{O}\left((\log N)^6 N^3\right)$
Polar frequencies	–	$\mathcal{O}(N^2 \log N)$	–

Table 7.2: Proven rates of the running times for $M_1 = \mathcal{O}\left(N^{d-1}\right)$ center points and $M_2 = \mathcal{O}(N)$ radii.

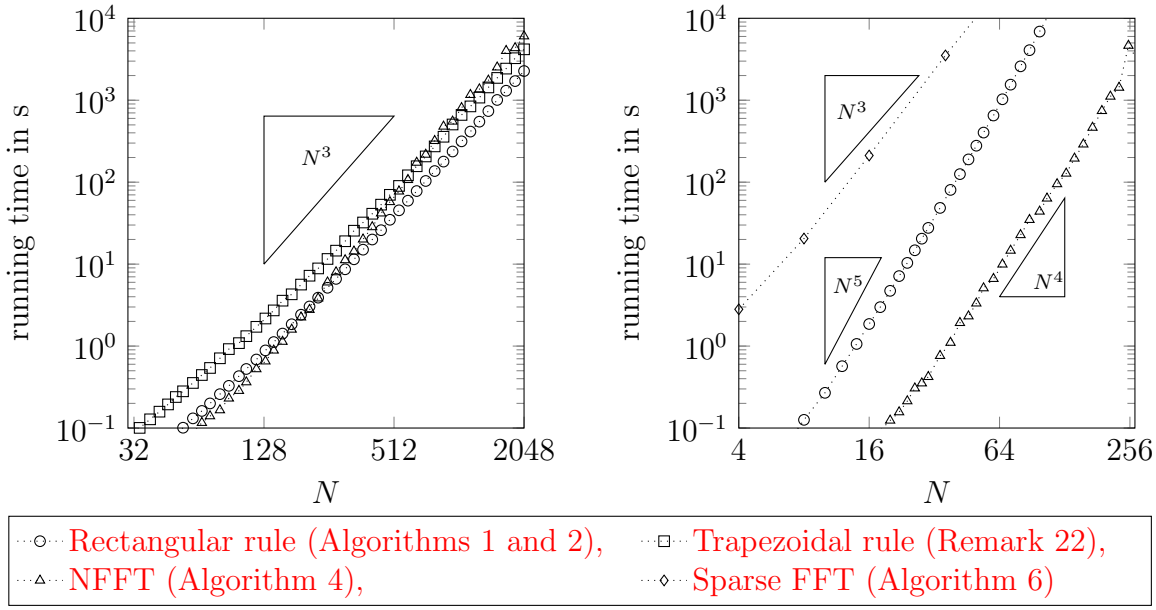


Figure 7.3: Running times for the discretizations of $\mathcal{M}$ with respect to the discretization parameter N for spatial dimension $d = 2$ (left) and $d = 3$ (right).

Table 7.2 lists the proven rates and Figure 7.3 summarizes the numerically obtained running times with respect to the discretization parameter N. For spatial dimension $d = 2$, the differences are not significant, but for spatial dimension $d = 3$, the NFFT based variant outperforms the other algorithms. However, the best increasing rate is given by Algorithm 6 and consequently, we expect benefits from using this methods for sufficient large problem sizes.

Let us repeat the same procedure for the discretization of the mean value operator $\mathcal{N}$ with the spatial dimension $d = 2$. In addition to the algorithms, which are based on a rectangular rule (Algorithm 3), nonequispaced fast Fourier transforms (Algorithm 5) and sparse fast Fourier transforms (Corollary 29), we consider the variant with polar frequencies (Algorithm 7). The results are shown in Figure 7.4. We observe a superior performance of the Fourier based algorithms. The use of polar frequencies leads to best running times, but is also restricted to a particular setting. In general situations, the NFFT based approach is in our experiments optimal regarding the running times, but similar to the operator $\mathcal{M}$, we expect advantages for sufficient large problem sizes from using sparse FFTs.

7.3 Reconstructions

In this section, we demonstrate the reconstruction of objects from their mean values. We start with a least squares solution as an initial guess and utilize afterwards iterative algorithms with regularization techniques to improve the reconstruction quality. Here we use the Fourier based forward operator from Algorithm 4.

We consider two settings, which differ in the arrangement of the center points. Firstly, the center points lie on a circular line around the object. In a second experiment, we analyze the situation, where the center points partially surround the object.

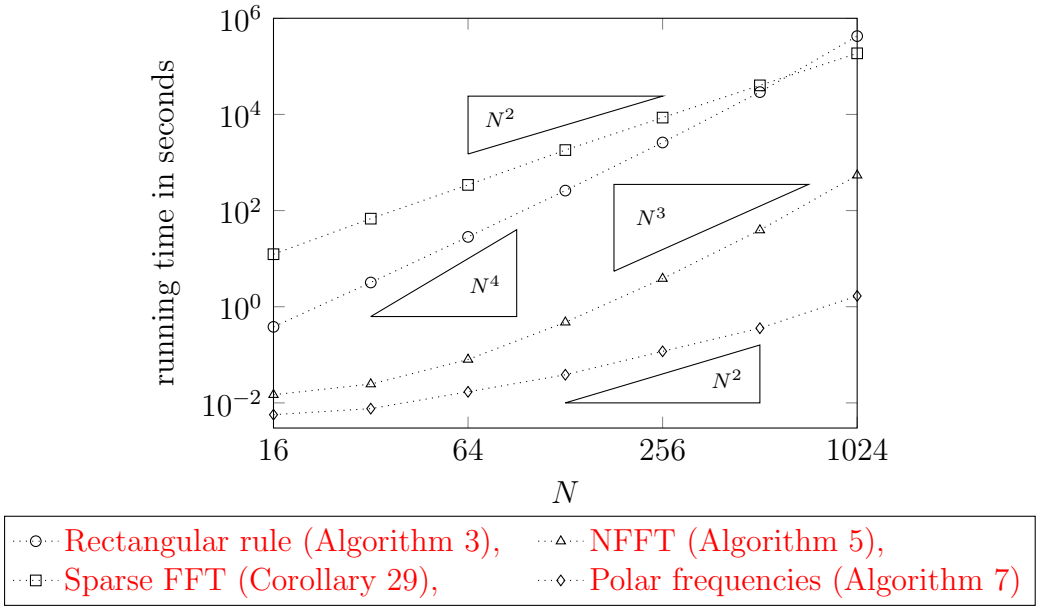


Figure 7.4: Running times for the discretizations of $\mathcal{N}$ with respect to the discretization parameter N.

7.3.1 Surrounding center points

We start with fixed discretization parameter $N = 128$ and given mean values $\mathbf{g} \in \mathbb{R}^{N^2}$ according to $M_1 = N$ center points $\mathbf{y}_1, \dots \mathbf{y}_N \in \mathbb{R}^2$ and $M_2 = N$ radii $r_1, \dots, r_N \in [0, 1]$,

$$\mathbf{y}_j := \frac{1}{2} \begin{pmatrix} \cos \frac{2\pi(j-1)}{N} \\ \sin \frac{2\pi(j-1)}{N} \end{pmatrix} \quad \text{and} \quad r_k := \frac{k-1}{N}, \quad j, k \in \{1, \dots, N\}.$$

The task is to find a solution $\mathbf{f} \in \mathbb{R}^{N^2}$ of equation (5.2). Instead of the exact data $\mathbf{g}$, we work with noisy data $\tilde{\mathbf{g}}_a \in \mathbb{R}^{N^2}$,

$$\tilde{\mathbf{g}}_a := \mathbf{g} + a\,(-0.5 + \boldsymbol{\eta}) \cdot \left(\max_{n \in \{1,\dots,N^2\}} |g_n| \right), \quad a > 0,$$

with uniformly distributed pseudorandom numbers $\boldsymbol{\eta} \in [0, 1]^{N^2}$.
Beside visual ratings, we evaluate the quality of the reconstructions by using the peak signal-to-noise ratio (PSNR) [21], which is a widely used image quality assessment measure. With the mean squared error $\varepsilon_{\text{mse}} \geq 0$ between $\mathbf{f} \in \mathbb{R}^{N^2}$ and $\tilde{\mathbf{f}} \in \mathbb{R}^{N^2}$,

$$\varepsilon_{\text{mse}}\left(\mathbf{f}, \tilde{\mathbf{f}}\right) = \frac{1}{N^2} \sum_{n=1}^{N^2} \left| f_n - \tilde{f}_n \right|^2, \quad \mathbf{f}, \tilde{\mathbf{f}} \in \mathbb{R}^{N^2},$$

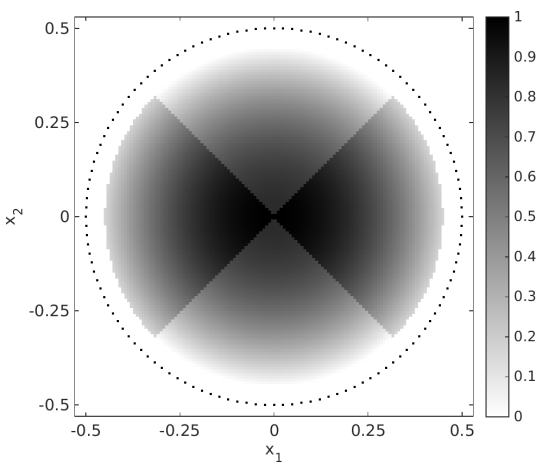
the in decibel measured PSNR value between an exact signal $\mathbf{f}$ and a noisy signal $\tilde{\mathbf{f}}$ is defined as

$$\text{psnr}\left(\mathbf{f}, \tilde{\mathbf{f}}\right) := 10 \log_{10} \left(\frac{(\max \mathbf{f})^2}{\varepsilon_{\text{mse}}\left(\mathbf{f}, \tilde{\mathbf{f}}\right)} \right).$$

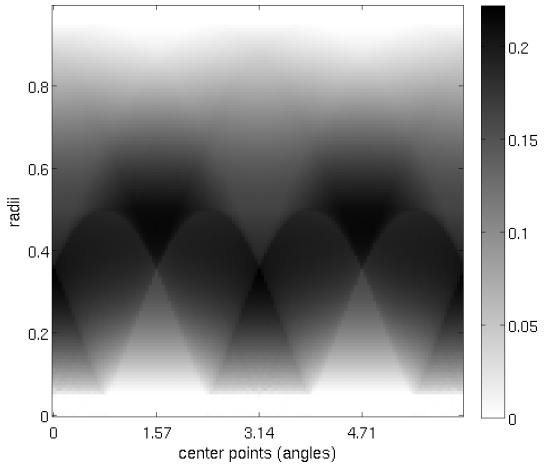
In principle, a higher PSNR value indicates a higher reconstruction quality, but it is advisable to consider also other criterions for a substantial statement.
The used object $f\colon \mathbb{R}^2 \to \mathbb{R}$ in the experiments is a modification of the test function (7.1) with smoothness parameter $s = 1$,

$$f(\mathbf{x}) := \frac{1}{1.2}\left(f_{2,1,0.45} + 0.2\chi_{\{\mathbf{x}\in\mathbb{R}^2 : |x_2|<|x_1| \text{ and } \|\mathbf{x}\|_2^2 \leq 0.45^2\}}(\mathbf{x})\right),$$

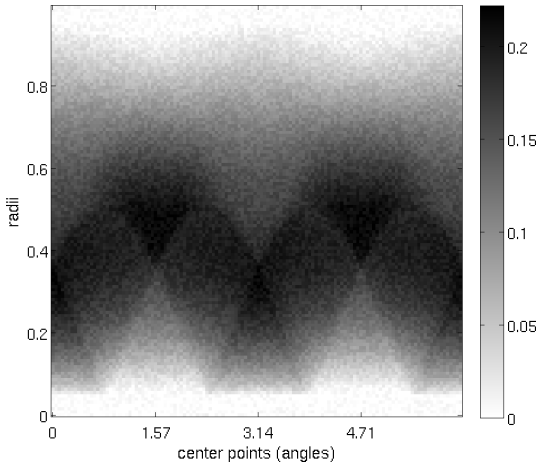
such that we obtain a piecewise smooth function, which is shown in Figure 7.5a. We compute the mean values $\mathbf{g} = \mathbf{M}\mathbf{f}$ with the Fourier based discretization $\mathbf{M}$ of the mean value operator $\mathcal{M}$, see Figure 7.5b, consider the noise levels $a = 10\%$ and $a = 20\%$, and obtain the data in Figures 7.5c and 7.5d.



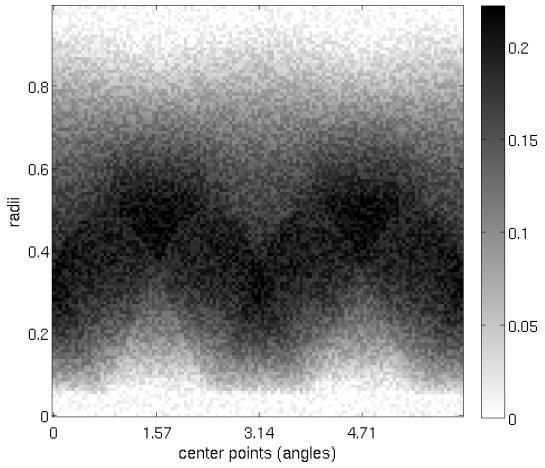
(a) Test function $\mathbf{f}$ and center points $\mathbf{y}$.



(b) Mean values $\mathbf{g} = \mathbf{M}\mathbf{f}$.



(c) Noisy mean values $\tilde{\mathbf{g}}_{0.1}$, PSNR=30.76 dB.



(d) Noisy mean values $\tilde{\mathbf{g}}_{0.2}$, PSNR=24.78 dB.

Figure 7.5: Test function with center points on a circular line, according mean values, and noisy mean values for the noise levels $a = 10\%$ and $a = 20\%$.

We start with the solutions from the least squares method, presented in the first row of Figure 7.6, as an initial guess and apply the derived inversion algorithms from Section 5 to obtain regularized reconstructions. The results in the second row are produced by the total variation based Algorithm 8, where the biconjugate gradient stabilized method [59] is utilized to solve the linear system of equations in line 13. The Shearlet based

Algorithm 9 leads to the third row. Here, we use the symmetric LQ method [6, 56], available in Matlab as function symmlq, for the implementation of equation (5.24). All parameters of the experiments are listed in Tables 7.3 and 7.4

noise level	α	γ	Netwon steps
$a = 10\%$	0.0013	2.7826e−05	6
$a = 20\%$	0.0100	2.7826e−05	5

Table 7.3: Used parameters for the total variation based reconstruction Algorithm 8.

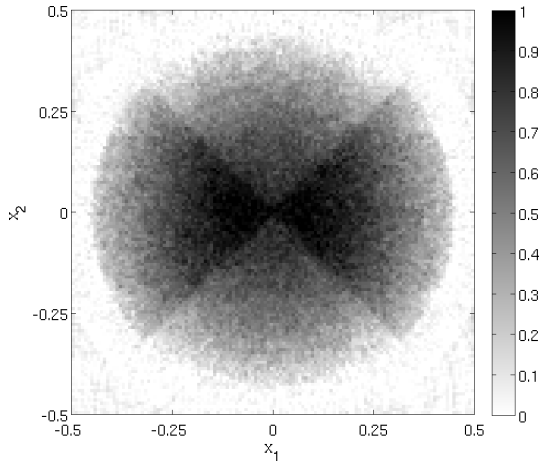
noise level	λ	ρ	ADMM steps	symmlq maxit	symmlq tol
$a = 10\%$	1.7857e−04	0.0721	13	3	1e−3
$a = 20\%$	3.5000e−04	0.0514	13	4	1e−3

Table 7.4: Used parameters for the Shearlet based reconstruction Algorithm 9 with center points on a circular line.

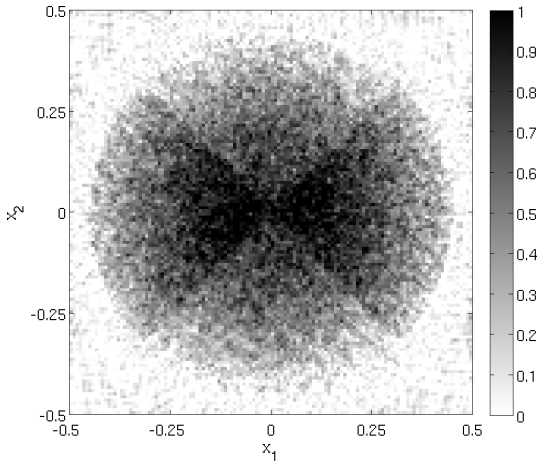
Starting with a noisy solution from the least squares method as initial guess, the regularization techniques in Algorithms 8 and 9 improve the reconstruction quality significantly. While the PSNR values are of the same magnitude, a visual inspection indicates the typical properties of each of the two regularization variants. The total variation based algorithm is well-suited for piecewise constant objects. However, in our setting with a piecewise smooth object, we obtain images of flat regions separated by artifact boundaries, which is also known as the staircasing effect. The Shearlet based approach leads to reconstructions with smoother areas and sharper edges, and requires less computation time. Altogether, for this particular setting with a piecewise smooth object, the Shearlet based regularization outperforms the total variation based approach.

7.3.2 Incomplete data

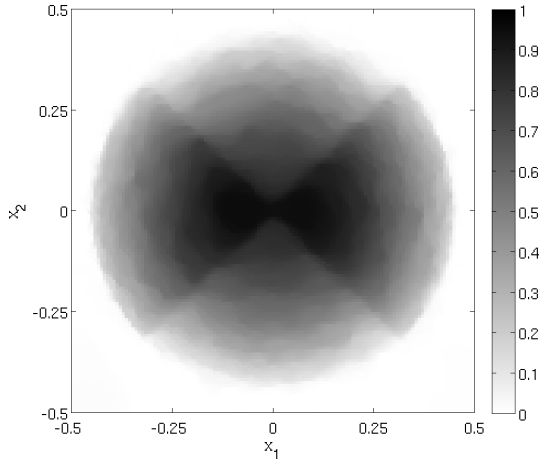
For various reasons, it is not in every application possible to place measuring points completely around the object. This fact motivates the following experiment, where we use mainly the setting from Section 7.3.1. The difference is, that we consider center points on a mirrored "J"-shape, see Figure 7.7a, and equispaced radii $r \in \left[0, \frac{1+\sqrt{2}}{2}\right]$. The related mean values are presented in Figure 7.7b, and Figures 7.7c and 7.7d show this input data with added noise. The following procedure is equal to the previous experiments. We start with the least squares solutions from Figures 7.8a and 7.8b. Afterwards, we improve this reconstructions with total variation based regularization, see Figures 7.8c and 7.8d, and Shearlet based regularization, see Figures 7.8e and 7.8f. The used parameters are listed in the Tables 7.3 and 7.5.
Due to the missing center points on the right side, the reconstructions are imprecise on the right part of the object. Apart from that, we observe the same characteristics as in



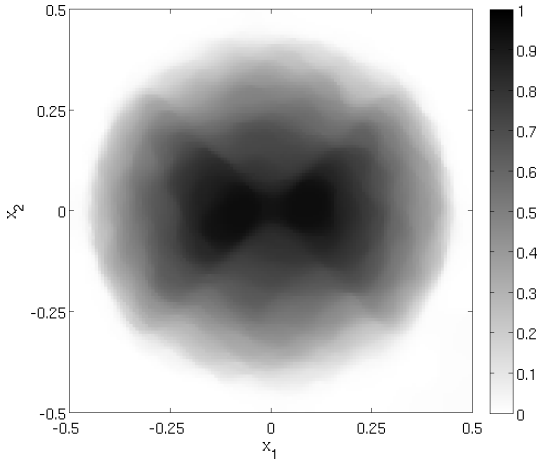
(a) PSNR=26.04 dB



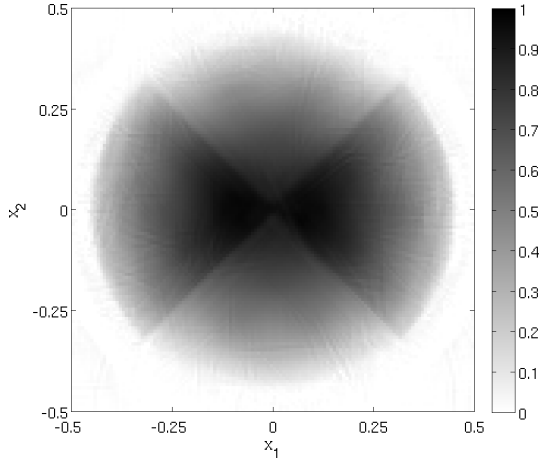
(b) PSNR=19.29 dB



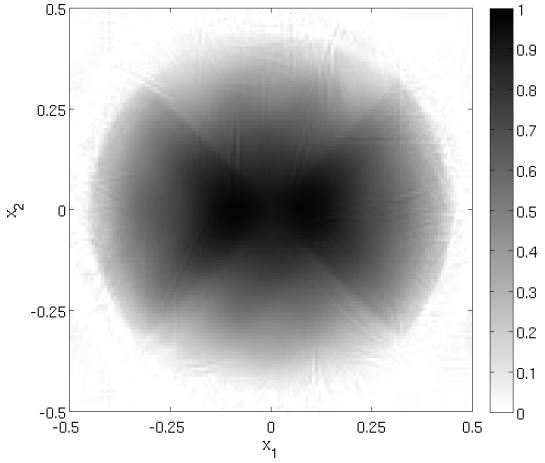
(c) PSNR=34.21 dB, τ = 38:18



(d) PSNR=32.53 dB, τ = 30:21



(e) PSNR=32.92 dB, τ = 7:28



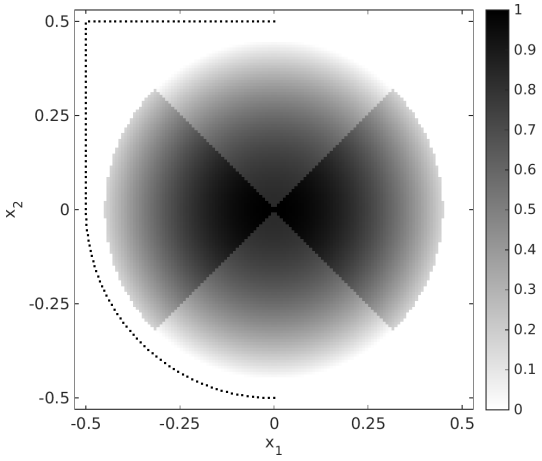
(f) PSNR=30.41 dB, τ = 7:57

Figure 7.6: Reconstructions of the test function for center points on a circular line and noise levels $a = 10\%$ (left) and $a = 20\%$ (right) with least squares method (first row), total variation based regularization (second row), and Shearlet based regularization (last row). The computation time in minutes:seconds is denoted by τ.

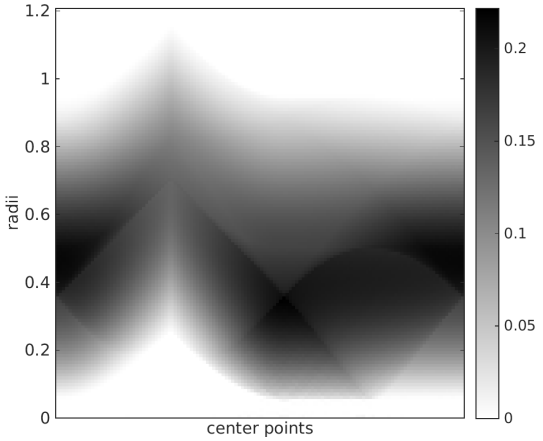
Section 7.3.1. In particular, both reconstruction methods from Chapter 5 increase the reconstruction quality from the least squares solutions significantly. While the total variation based regularization produces the typical staircasing effects, the Shearlet based regularization stands out with sharper edges and faster computations.

noise level	λ	ρ	ADMM steps	symmlq maxit	symmlq tol
$a = 10\%$	1.3053e−04	0.0695	17	3	1e−3
$a = 20\%$	3.5000e−04	0.0514	4	4	1e−3

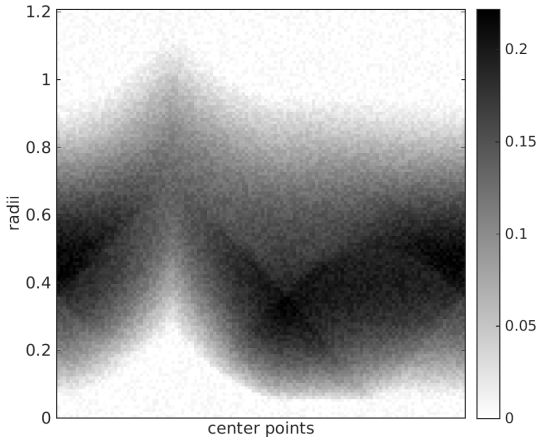
Table 7.5: Used parameters for the Shearlet based reconstruction Algorithm 9 with center points on a J-shape.



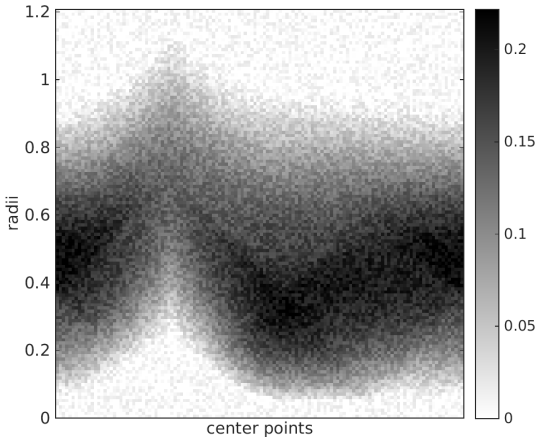
(a) Test function $\mathbf{f}$ and center points $\mathbf{y}$.



(b) Mean values $\mathbf{g} = \mathbf{M}\mathbf{f}$.

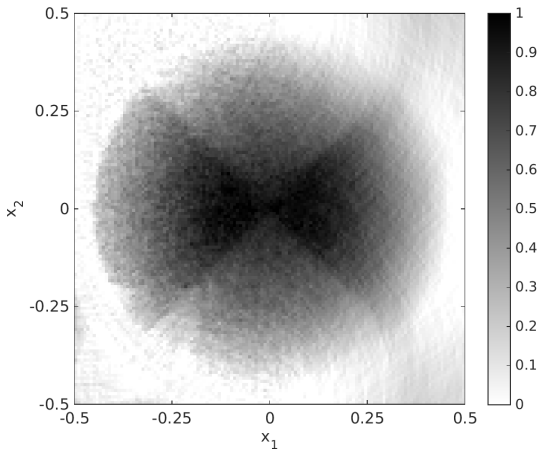


(c) Noisy mean values $\tilde{\mathbf{g}}_{0.1}$, PSNR=30.77 dB.

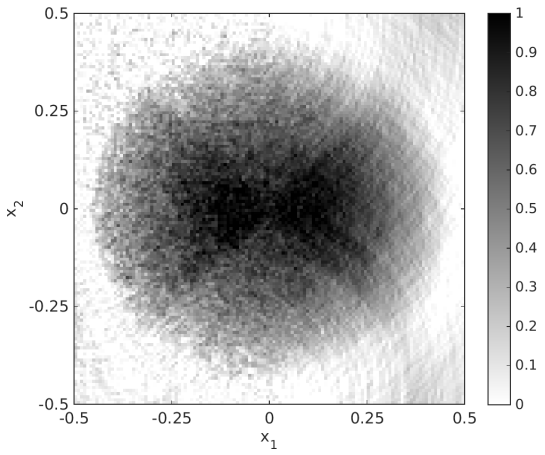


(d) Noisy mean values $\tilde{\mathbf{g}}_{0.2}$, PSNR=24.74 dB.

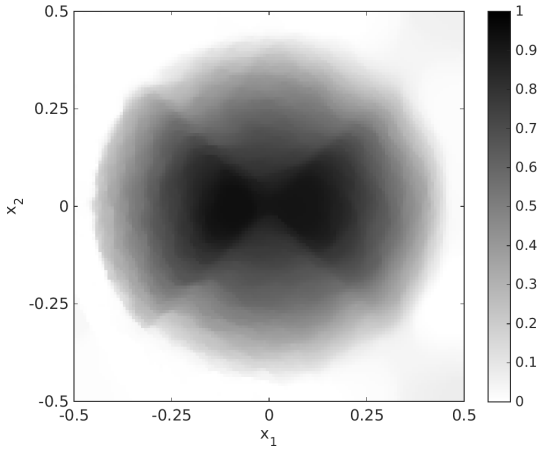
Figure 7.7: Test function with center points on a J-shape, according mean values, and noisy mean values for the noise levels $a = 10\%$ and $a = 20\%$.



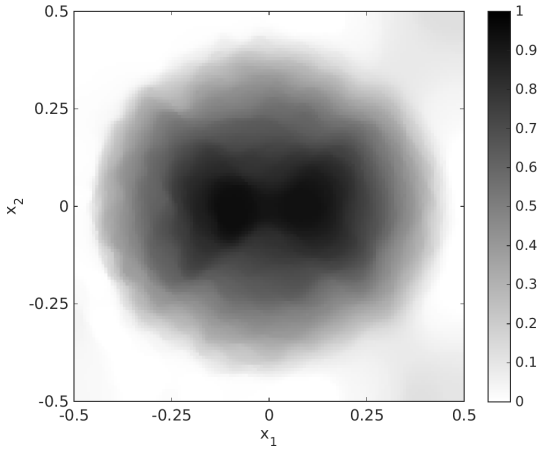
(a) PSNR=24.93 dB



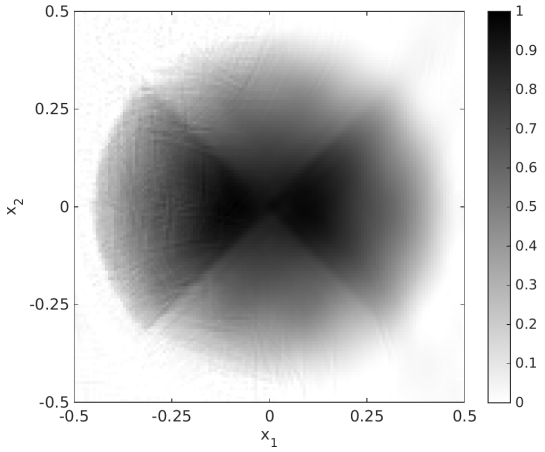
(b) PSNR=21.83 dB



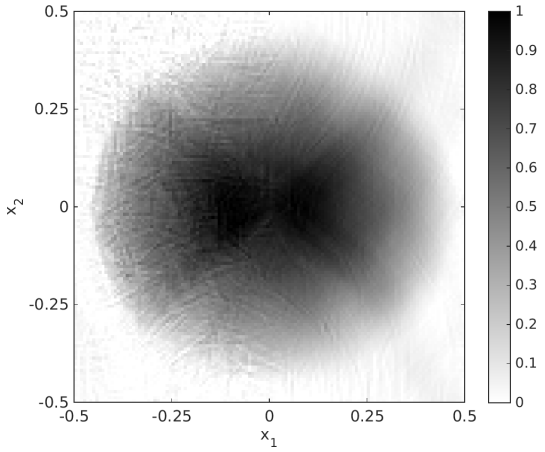
(c) PSNR=30.55 dB, $\tau = 77{:}12$



(d) PSNR=26.93 dB, $\tau = 49{:}50$



(e) PSNR=30.81 dB, $\tau = 21{:}53$



(f) PSNR=27.68 dB, $\tau = 6{:}07$

Figure 7.8: Reconstructions of the test function for center points on a J-shape and noise levels $a = 10\%$ (left) and $a = 20\%$ (right) with least squares method (first row), total variation based regularization (second row), and Shearlet based regularization (last row). The computation time in minutes:seconds is denoted by τ.

8

Conclusion

In this thesis, we developed efficient and accurate reconstruction methods for photoacoustic imaging. Under some assumptions, in particular a constant speed of sound, this process can be modelled as the inversion of an integral operator, the so-called spherical mean value operator, which assigns functions their mean values over spheres. The presented algorithms are iterative methods, which require the fast computation of mean values, the so-called forward problem. For this purpose, we use finite Fourier series as approximations of continuous functions. This allows to evaluate the specific integrals, which are related to the computation of mean values, by multiplications with Bessel functions in the Fourier domain.

In practice, the reconstruction of objects from given data poses some additional challenges. Different issues, for example imprecise measurements, lead to undesirable effects, such as noise or particular artefacts. To avoid this problems, we employ regularization techniques, which are realized by forcing the reconstructions to fulfill certain properties, which are given by the application. In photoacoustic imaging, piecewise smooth objects are a suitable assumption. This motivates our approach to prefer reconstructions with a small total variation or sparse Shearlet coefficients.

The extensive numerical experiments demonstrate the fundamental capabilities of the derived algorithms. Even in the case of incomplete data, we can produce accurate reconstructions. For a complete validation of this promising results, it remains to extend the theoretical background to more precise error bounds and a deeper analysis of the approximation by Fourier series with polar frequency grids.

In addition, if the detailed setting of the application is known, for example the amount of data, the examined objects and the necessary accuracy, then the performance of the algorithms can be further improved. For example the use of customized parameters and programming techniques can lead to less running times and an increased reconstruction quality.

This thesis is completed by the implementation and publication of all derived algorithms as a Matlab toolbox [27]. The large amount of code is organized in an object-oriented design and numerous examples are included, which allows a further development by different researchers.

Bibliography

[1] M. Abramowitz and I. A. Stegun, eds. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, Inc., New York, 1972.

[2] B. Adcock, M. Gataric, and A. C. Hansen. Weighted frames of exponentials and stable recovery of multidimensional functions from nonuniform Fourier samples. *ArXiv e-prints*, Aug. 2014.

[3] M. Agranovsky, P. Kuchment, and L. Kunyansky. On Reconstruction Formulas and Algorithms for the Thermoacoustic Tomography. In L. V. Wang, editor, *Photoacoustic Imaging and Spectroscopy*, pp. 89–101. CRC Press, Boca Raton, FL, 2009.

[4] M. Agranovsky and P. Kuchment. Uniqueness of reconstruction and an inversion procedure for thermoacoustic and photoacoustic tomography with variable sound speed. *Inverse Problems*, 23(5):2089–2102, 2007.

[5] M. Agranovsky, P. Kuchment, and E. T. Quinto. Range descriptions for the spherical mean Radon transform. *J. Funct. Anal.*, 248(2):344–386, 2007.

[6] R. Barrett, M. Berry, T. F. Chan, et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[7] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, 2011.

[8] E. Candès, L. Demanet, D. Donoho, and L. Ying. Fast Discrete Curvelet Transforms. *Multiscale Model. Simul.*, 5(3):861–899, 2006.

[9] E. J. Candès and D. L. Donoho. New Tight Frames of Curvelets and Optimal Representations of Objects with Piecewise C^2 Singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.

[10] J. W. Cooley and J. W. Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *Math. Comp.*, 19(90):297–301, 1965.

[11] R. Courant and D. Hilbert. *Methods of Mathematical Physics. Vol. II: Partial Differential Equations*. Interscience Publishers (a division of John Wiley & Sons), New York-London, 1962.

[12] M. Do and M. Vetterli. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Trans. Image Process.*, 14(12):2091–2106, 2005.

[13] Y. Dong, T. Görner, and S. Kunis. An algorithm for total variation regularized photoacoustic imaging. *Adv. Comput. Math.*, 41(2):423–438, 2015.

[14] D. L. Donoho. Wedgelets: nearly minimax estimation of edges. *Ann. Statist.*, 27(3):859–897, 1999.

[15] G. Easley, D. Labate, and W.-Q. Lim. Sparse directional image representations using the discrete shearlet transform. *Appl. Comput. Harmon. Anal.*, 25(1):25–46, 2008.

[16] F. Filbir, R. Hielscher, and W. R. Madych. Reconstruction from circular and spherical mean data. *Appl. Comput. Harmon. Anal.*, 29(1):111–120, 2010.

[17] F. Filbir, S. Kunis, and R. Seyfried. Effective Discretization of Direct Reconstruction Schemes For Photoacoustic Imaging In Spherical Geometries. *SIAM J. Numer. Anal.*, 52(6):2722–2742, 2014.

[18] D. Finch, M. Haltmeier, and Rakesh. Inversion of Spherical Means and the Wave Equation in Even Dimensions. *SIAM J. Appl. Math.*, 68(2):392–412, 2007.

[19] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser/Springer, New York, 2013.

[20] Free Software Foundation. GNU General Public License version 3 (GPLv3). http://www.gnu.org/licenses/. June 2007.

[21] J. D. Gibson and A. Bovik, eds. *Handbook of Image and Video Processing*. Academic Press, Inc., Orlando, FL, USA, 2000.

[22] R. Griesse and D. A. Lorenz. A semismooth Newton method for Tikhonov functionals with sparsity constraints. *Inverse Problems*, 24(3):035007, 2008.

[23] K. Gröchenig. Reconstruction Algorithms in Irregular Sampling. *Math. Comp.*, 59(199):181–194, 1992.

[24] K. Guo and D. Labate. Optimally Sparse Multidimensional Representation Using Shearlets. *SIAM J. Math. Anal.*, 39(1):298–318, 2007.

[25] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. *Inverse Probl. Imag.*, 6(4):645–661, 2012.

[26] T. Görner and S. Kunis. Effective discretization of the two-dimensional wave equation. *PAMM*, 14(1):947–948, 2014.

[27] T. Görner and S. Kunis. SMV, Matlab toolbox for computing spherical mean values. http://torstengoerner.de/software/. May 2015.

[28] M. Haltmeier. Inversion of circular means and the wave equation on convex planar domains. *Comput. Math. Appl.*, 65(7), 2013.

[29] M. Haltmeier, O. Scherzer, and G. Zangerl. A Reconstruction Algorithm for Photoacoustic Imaging Based on the Nonuniform FFT. *IEEE Trans. Med. Imag.*, 28(11):1727–1735, 2009.

[30] M. Haltmeier, T. Schuster, and O. Scherzer. Filtered backprojection for thermoacoustic computed tomography in spherical geometry. *Math. Methods Appl. Sci.*, 28(16):1919–1937, 2005.

[31] M. Haltmeier. A Mollification Approach for Inverting the Spherical Mean Radon Transform. *SIAM J. Appl. Math.*, 71(5):1637–1652, 2011.

[32] M. Haltmeier. Universal Inversion Formulas for Recovering a Function from Spherical Means. *SIAM J. Math. Anal.*, 46(1):214–232, 2014.

[33] S. Hassani. *Mathematical Physics: A Modern Introduction to Its Foundations*. Springer-Verlag, New York, 1999.

[34] S. Häuser and J. Ma. Seismic Data Reconstruction via Shearlet-Regularized Directional Inpainting. *Preprint*, May 2012.

[35] S. Häuser and G. Steidl. Convex multiclass segmentation with shearlet regularization. *Int. J. Comput. Math.*, 90(1):62–81, 2013.

[36] S. Häuser and G. Steidl. Fast Finite Shearlet Transform: a tutorial. *ArXiv e-prints*, July 2014.

[37] M. Hintermüller, K. Ito, and K. Kunisch. The Primal-Dual Active Set Strategy as a Semismooth Newton Method. *SIAM J. Optim.*, 13(3):865–888 (2003), 2002.

[38] M. Hintermüller and G. Stadler. An Infeasible Primal-Dual Algorithm for Total Bounded Variation-Based Inf-Convolution-Type Image Restoration. *SIAM J. Sci. Comput.*, 28(1):1–23, 2006.

[39] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms I: Fundamentals*. Springer-Verlag, Berlin, 1993.

[40] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods*. Springer-Verlag, Berlin, 1993.

[41] P. J. Huber. Robust Estimation of a Location Parameter. *Ann. Math. Statist.*, 35:73–101, 1964.

[42] F. John. *Plane Waves and Spherical Means Applied to Partial Differential Equations*. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1955 original.

[43] J. Keiner, S. Kunis, and D. Potts. Using NFFT 3 – A Software Library for Various Nonequispaced Fast Fourier transforms. *ACM Trans. Math. Software*, 36(4):Art. 19, 30, 2009.

[44] P. Kuchment and L. Kunyansky. Mathematics of thermoacoustic tomography. *European J. Appl. Math.*, 19(2):191–224, 2008.

[45] S. Kunis and I. Melzer. A Stable and Accurate Butterfly Sparse Fourier Transform. *SIAM J. Numer. Anal.*, 50(3):1777–1800, 2012.

[46] L. Kunyansky. Reconstruction of a function from its spherical (circular) means with the centers lying on the surface of certain polygons and polyhedra. *Inverse Problems*, 27(2):025012, 22, 2011.

[47] L. A. Kunyansky. A series solution and a fast algorithm for the inversion of the spherical mean Radon transform. *Inverse Problems*, 23(6):S11–S20, 2007.

[48] L. A. Kunyansky. Explicit inversion formulae for the spherical mean Radon transform. *Inverse Problems*, 23(1):373–383, 2007.

[49] G. Kutyniok, W.-Q. Lim, and X. Zhuang. Digital Shearlet Transforms. In G. Kutyniok and D. Labate, editors, *Shearlets*, pp. 239–282. Birkhäuser/Springer, New York, 2012.

[50] G. Kutyniok, M. Shahram, and D. L. Donoho. Development of a digital shearlet transform based on Pseudo-Polar FFT. In *Proc. SPIE 7446*, 2009, 74460B–74460B–13.

[51] G. Kutyniok, M. Shahram, and X. Zhuang. ShearLab: A Rational Design of a Digital Parabolic Scaling Algorithm. *SIAM J. Imaging Sci.*, 5(4):1291–1332, 2012.

[52] D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss. Sparse multidimensional representation using shearlets. In *Proc. SPIE 5914*, 2005, 59140U–59140U–9.

[53] W.-Q. Lim. The Discrete Shearlet Transform: A New Directional Transform and Compactly Supported Shearlet Frames. *IEEE Trans. Image Process.*, 19(5):1166–1180, 2010.

[54] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way*. Elsevier/Academic Press, Amsterdam, third ed., 2009.

[55] F. Natterer. Photo-acoustic inversion in convex domains. *Inverse Probl. Imaging*, 6(2):1–6, 2012.

[56] C. C. Paige and M. A. Saunders. Solutions of Sparse Indefinite Systems of Linear Equations. *SIAM J. Numer. Anal.*, 12(4):617–629, 1975.

[57] E. T. Quinto. Helgason's support theorem and spherical Radon transforms. In, *Radon Transforms, Geometry, and Wavelets*. Vol. 464, pp. 249–264. Amer. Math. Soc., Providence, RI, 2008.

[58] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[59] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second ed., 2003.

[60] R. Seyfried. *Summability Methods for the Inversion of the Spherical Mean Operator*. Shaker, Aachen, DE, 2014.

[61] P. Toft. *The Radon Transform: Theory and Implementation*. Department of Mathematical Modelling, Section for Digital Signal Processing, Technical University of Denmark, 1996.

[62] G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Cambridge, GB, 1966.

[63] L. Ying. Sparse Fourier Transform via Butterfly Algorithm. *SIAM J. Sci. Comput.*, 31(3):1678–1694, 2009.

[64] A. Zygmund. *Trigonometric Series, Vol. I and II*. Cambridge University Press, Cambridge, GB, third ed., 2002.

List of Figures

3.1 Spherical coordinate systems. 9

4.1 Regular Cartesian frequency grid. 25
4.2 Distribution of the frequencies on a double cone. 28
4.3 Setting for a fast computation of mean values of functions, which are given by Fourier sums with frequencies on a polar grid. 35

5.1 Maps for the TV seminorm and the Huber function. 48
5.2 The mother Shearlet and dilated and sheared versions. 56
5.3 Auxiliary function v and b. 59
5.4 Wavelets for the radial part and splines for the angular part. 60
5.5 The mother Curvelet and scaled and rotated versions. 63

6.1 Hierarchical structure of the implemented classes. 66
6.2 Available test functions of the Matlab toolbox. 69

7.1 Test function $f_{d,s,t}$. 72
7.2 Accuracy of the discretization of $\mathcal{M}$. 74
7.3 Running times for the discretizations of $\mathcal{M}$. 76
7.4 Running times for the discretizations of $\mathcal{N}$. 77
7.5 Test function, according mean values, and noisy mean values for center points an a circular line. 78
7.6 Reconstructions of the test function for center points on a circular line. 80
7.7 Test function, according mean values, and noisy mean values for center points on a J-shape. 82
7.8 Reconstructions of the test function for center points on a J-shape. . . 83

List of Tables

4.1 Convergence rates from Theorems 32 and 33. 34

7.1 Estimated orders of convergence. 75
7.2 Proven rates of the running times. 75
7.3 Used parameters for the total variation based reconstruction. 79
7.4 Used parameters for the Shearlet based reconstruction with center points on a circular line. 79
7.5 Used parameters for the Shearlet based reconstruction Algorithm with center points on a J-shape. 81

List of Algorithms

1 Discrete mean value operator $\mathcal{M}$, using quadrature, $d = 2$. 21
2 Discrete mean value operator $\mathcal{M}$, using quadrature, $d = 3$. 22
3 Discrete mean value operator $\mathcal{N}$, using quadrature, $d = 2$. 23
4 Discrete mean value operator $\mathcal{M}$, using NFFT. 26
5 Discrete mean value operator $\mathcal{N}$, using NFFT, $d = 2$. 27
6 Discrete mean value operator $\mathcal{M}$, using sparse FFT, $d = 3$. 30
7 Discrete mean value operator $\mathcal{N}$, using a polar frequency grid, $d = 2$. . 37

8 Inversion of the discrete mean value operator, using a generalized Newton method and total variation based regularization. 53
9 Inversion of the discrete mean value operator, using ADMM and Shearlet based regularization. 58