Universität Hamburg

Dissertation

Error Estimates and Convergence Rates for Filtered Back Projection Reconstructions

---

Dissertation zur Erlangung des Doktorgrades
an der
Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Mathematik
der Universität Hamburg

vorgelegt von
Matthias Beckmann

Hamburg, 2018

Als Dissertation angenommen vom Fachbereich Mathematik der Universität Hamburg auf Grund der Gutachten von:

Prof. Dr. Armin Iske
Prof. Dr. Peter Maaß

Datum der Disputation: 17.05.2018

Abstract

The method of filtered back projection (FBP) is a commonly used reconstruction technique in computerized tomography, which allows us to recover an unknown bivariate function from the knowledge of its Radon data. The reconstruction is based on the classical FBP formula, which yields an analytical inversion of the Radon transform provided that the complete Radon data is available. The FBP formula, however, is highly sensitive with respect to noise and, hence, numerically unstable. To overcome this problem, suitable low-pass filters of finite bandwidth and with compactly supported window functions are employed. This reduces the noise sensitivity, but only leads to an inexact approximation of the target function.

The main objective of this thesis is to analyse the inherent FBP reconstruction error which is incurred by the application of the low-pass filter. To this end, we present error estimates in Sobolev spaces of fractional order and provide quantitative criteria to a priori evaluate the performance of the utilized low-pass filter by means of its window function. The obtained error bounds depend on the bandwidth of the low-pass filter, on the flatness of the filter's window function at the origin, on the smoothness of the target function, and on the order of the considered Sobolev norm in which the reconstruction error is measured.

Further, we prove convergence for the approximate FBP reconstruction in the treated Sobolev norms along with asymptotic convergence rates as the filter's bandwidth goes to infinity, where we in particular observe saturation at fractional order depending on smoothness properties of the filter's window function.

Finally, we develop convergence rates for noisy data as the noise level goes to zero, where we prove estimates for the data error and combine these with our results for the approximation error. Furthermore, the filter's bandwidth is coupled with the noise level to achieve the convergence.

The theoretical results are supported by numerical experiments.

Acknowledgements

First and foremost, I want to express my deep gratitude to my supervisor Prof. Dr. Armin Iske for his perpetual support and encouragement. In particular, I thank him for always having a sympathetic ear for my questions, problems and doubts and for helping me to find my own way in mathematics. Thanks to his support I had the opportunity to attend various workshops and conferences and, hence, to make new contacts and meet many great people. But most of all I appreciate his warm-heartedness and the familial atmosphere in his working group.

I cordially thank Prof. Dr. Peter Maaß for writing the second report on my thesis as well as Prof. Dr. Stefan Kunis for acting as external examiner in my thesis defense.

Special thanks go to my colleagues, to my fellow PhD students and, specifically, to my fantastic working group including all former members. I truly enjoyed the time together at Geomatikum and will never forget our discussions during the relaxing and inspiring coffee breaks. Furthermore, I am very grateful to Benedikt, Niklas and Sara for proofreading this thesis.

My deepest thanks go to my family and friends. This thesis would not have been possible without their love and support. Most notably, I want to thank my mother for always listening to my doubts and getting me back on track. I highly appreciate her loving support and constant encouragement throughout the years.

I dedicate this work to my mother and to the memory of my grandmother.

Contents

Chapter 1

Introduction

The development of *computerized tomography* (CT) has revolutionized the field of diagnostic radiology and CT is by now one of the standard modalities in medical imaging. Its goal consists in imaging the interior structure of a scanned object by measuring and processing the attenuation of X-rays along a large number of lines through the cross-sections of the object to be examined. In this process, a fundamental feature of CT is the mathematical reconstruction of an image by the application of a suitable and sophisticated algorithm.

The impact of CT in diagnostic medicine has been revolutionary, since it has provided a non-invasive imaging modality and has enabled doctors to view internal organs with mould-breaking precision and safety for the patient. Since the invention of the first CT scanner in the 1970s the number of CT scans for diagnostic purpose has been growing extensively. In addition, there are numerous non-medical imaging applications which are also based on the methods of computerized tomography. One example is non-destructive testing (NDT) in materials science, where we want to evaluate the properties of a material without causing damage. Another application is electron microscopy, which is a typical example for an incomplete data problem, because only observations in a limited angular range are available.

Mathematically, an X-ray scan provides the line integral values of the object's attenuation function along lines in the plane. Hence, the CT reconstruction problem requires the recovery of a bivariate function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ from the knowledge of its line integrals

$$\mathcal{R}f(t,\theta) = \int_{\{x\cos(\theta)+y\sin(\theta)=t\}} f(x,y)\,\mathrm{d}(x,y) \quad \text{for } (t,\theta) \in \mathbb{R}\times[0,\pi).$$

The purely mathematical problem of reconstructing a function from its line integral values was first studied and analytically solved by the Austrian mathematician Johann Radon in 1917 in his pioneering paper "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten", cf. [94]. In that work, Radon derived an explicit inversion formula for the linear integral transform

$$\mathcal{R} : f \longmapsto \mathcal{R}f$$

under the assumption that the data $\mathcal{R}f(t,\theta)$ is complete, i.e., available for all possible values $(t,\theta) \in \mathbb{R}\times[0,\pi)$. In his honour, the operator $\mathcal{R}$ is now known as the *Radon transform* and the corresponding integral values are called *Radon data*.

Historically, the foundation of CT was laid in 1895 by the German physicist Wilhelm C. Röntgen, who discovered a new kind of radiation, which he called *X-radiation* to emphasize its unknown type. Immediately after the discovery, X-rays have been used to image the interior of the human body. In 1901 his achievements earned Röntgen the first Nobel Prize in Physics. The two pioneering scientists who were primarily responsible for the development of computer assisted tomography in the 1960s and 1970s were Allan M. Cormack and Godfrey N. Hounsfield. With their work, the hitherto purely mathematical problem of reconstructing a bivariate function from the knowledge of its Radon transform has finally become relevant for practical applications.

In [20], [21], Cormack developed mathematical algorithms to create an image from X-ray scans. At about the same time, but working completely independently of Cormack, Hounsfield designed the first operational CT scanner as well as the first commercially available model, see [47]. In 1979 the Nobel Prize for Medicine and Physiology was jointly awarded to Cormack and Hounsfield for their fundamental achievements.

This breakthrough has attracted very much attention in engineering science as well as in the mathematical community. In particular, the tomographic reconstruction problem has been studied extensively and many different reconstruction algorithms were developed, cf. [30], [45], [82]. One of the most applied techniques is still the method of *filtered back projection* (FBP), which is based on an analytical inversion formula for the Radon transform and where low-pass filters with finite bandwidth and compactly supported window functions are employed.

Although the FBP algorithm has been one of the standard reconstruction algorithms in computerized tomography for more than 40 years, its error analysis and convergence behaviour are not completely settled so far. In [90], Popov showed pointwise convergence restricted to a small class of functions which are piecewise smooth and have jump discontinuities only along smooth curves. The approach of Rieder and Schuster [103] leads to L^2-convergence with suboptimal rates for compactly supported target functions from Sobolev spaces with smoothness greater than $\frac{1}{2}$. Here, the authors assume that the convolution kernel of the utilized low-pass filter has compact support. However, this is not satisfied for typical choices of the filter, as, for example, for the well-established Shepp-Logan or Ram-Lak filter. A different approach is taken by Rieder and Faridani in [100] as well as by Rieder and Schneck in [101], where optimal L^2-convergence rates are proven for sufficiently smooth target functions. However, the authors verify the assumptions of their theory only for a restricted class of filters which are based on (orthogonalized) B-splines. Recently, Qu [92] considered the FBP method in the continuous setting and showed convergence without rates in the L^1-norm and the L^2-norm for essentially bounded target functions with compact support as well as in every point of continuity under additional assumptions.

In this thesis, we focus also on the continuous setting and analyse the inherent FBP reconstruction error which is incurred by the application of the low-pass filter. To this end, we prove novel error estimates in Sobolev spaces of fractional order and provide quantitative criteria to a priori evaluate the performance of the utilized low-pass filter by means of its window function. Further, we prove convergence for the approximate FBP reconstruction in the treated Sobolev norms along with asymptotic convergence rates as the filter's bandwidth goes to infinity, where the smoothness order of the target function is only required to be positive. In addition, our results allow us to predict saturation of the order of convergence at fractional rates depending on smoothness properties of the filter's window function, which can easily be evaluated.

In contrast to that, the estimates in [82, Chapter V.1] by Natterer and in [32] by Faridani and Ritman deal with different types of errors which are incurred by the discretization of the FBP method and by the sampling of the Radon transform. In addition, the results are of qualitative nature in terms of essentially band-limited functions and, since the main tool used is the Poisson summation formula, the considered target functions are required to be continuous at least.

We finally remark that pointwise error formulas and L^∞-error bounds for the inherent FBP reconstruction error are discussed by Munshi et al. in [78] under rather restrictive differentiability conditions on the target function. In particular, the authors observe an affine-linear behaviour of the error with respect to the second derivative of the filter's window function at the origin. In this thesis, the above observation is substantiated and generalized under weaker assumptions.

To give a broad overview, the present work contributes to the analysis and assessment of low-pass filters in the filtered back projection method for approximately solving the two-dimensional tomographic reconstruction problem. Let us remark that the construction and calculation of reconstruction filters is still an important and active field of research, both in the two-dimensional case, cf. [87], [88], as well as in the three-dimensional setting for different scanning geometries, see, for example, [62], [64], [70], where the method of approximate inverse [65], [68] is applied.

With designing suitable reconstruction filters one can also tackle the problem of reconstructing certain features of the unknown target function by only using incomplete or local data, as, for instance, in contour tomography [66], or of combining the reconstruction with image analysis tasks like edge-detection or the direct calculation of wavelet coefficients, see [38], [39], [63], [69]. We finally remark that the FBP algorithm and its variants undergo steady improvements, as recently in [4], [74], [75], and have been adopted to various different settings, cf. [41], [104], [105], to name just a few recent developments.

The main part of this thesis is organized in four chapters as follows.

Chapter 2 has a introductory character and is devoted to the basic concepts of computerized tomography. We begin with a brief overview of the principle of X-ray tomography and introduce the Radon transform $\mathcal{R}$ as the mathematical model for the tomographic measuring process. Furthermore, we present some fundamental properties of the Radon transform and define the back projection operator $\mathcal{B}$. This leads us to an inversion formula for the Radon transform given by the classical filtered back projection formula, where we also study the ill-posedness of the tomographic reconstruction problem. In the last section of this chapter, we finally discuss classical reconstruction techniques, where we distinguish between *analytical* and *algebraic* methods. Most of the material presented in this Chapter is well-known and can be found, for example, in the textbooks [33], [82] or in the overview article [93].

In Chapter 3 we thoroughly introduce the method of filtered back projection, where low-pass filters A_L with finite bandwidth and compactly supported window functions are employed. This reduces the noise sensitivity of the FBP formula, but only leads to an inexact reconstruction. For target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ we show that the approximate FBP reconstruction f_L satisfies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ and can be written in standard form as

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f) = f * K_L,$$

where the convolution kernel $K_L \in \mathrm{L}^2(\mathbb{R}^2)$ is not compactly supported and not necessarily integrable on $\mathbb{R}^2$. To this end, we first collect some preliminary properties of the Radon transform $\mathcal{R}$ and the back projection operator $\mathcal{B}$. In the last section of this chapter, we then focus on the convolution kernel K_L and investigate some of its properties. We close the chapter with discussing the concrete example of the classical Ram-Lak filter.

Chapter 4 contains the main results of this thesis. Our principle goal is the analysis of the inherent FBP reconstruction error being incurred by the application of the low-pass filter. In particular, we aim at developing quantitative and easy-to-check criteria to a priori evaluate the performance of the filter by means of its window function W. In Section 4.1 we start with describing related techniques and results, which can be found in the literature, and explain their differences to our approach. Based on our L^2-error estimates from [6], we then develop refined error estimates for target functions from Sobolev spaces of fractional order. More precisely, for functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$ we analyse the H^σ-norm of the FBP reconstruction error

$$e_L = f - f_L$$

for all $0 \leq \sigma \leq \alpha$ and, under suitable assumptions on the filter's window function, we prove convergence of the approximate FBP reconstruction f_L to the target function f in the H^σ-norm as the filter's bandwidth L tends to infinity. Section 4.3 is then devoted to determining the rate of convergence in terms of L subject to properties of the window W. In particular, we show that the decay rate of the error bound is determined by the difference between the smoothness α of the target function and the order σ of the Sobolev norm in which the error e_L is measures, but saturates at (possibly) fractional rates depending on smoothness properties of W. In Section 4.4 we then prove asymptotic error estimates under weaker assumptions, that predict an affine-linear behaviour of the error with respect to the k-th derivative of the filter's window at the origin.

Finally, in Section 4.5 we consider the practically important case of noisy Radon data, where we combine our results for the approximation error with estimates for the data error and couple the filter's bandwidth with the noise level, as standard in the regularization theory of ill-posed problems. This enables us to determine asymptotic convergence rates for the overall error of the FBP reconstruction from noisy data as the noise level tends to zero.

In Chapter 5 we finally provide numerical experiments to validate the error theory we derived in Chapter 4. To this end, in Section 5.1 we first describe a standard discretization of the FBP method for the reconstruction of target functions from finitely many Radon samples in parallel beam geometry, cf. [33], [82]. Following this, we introduce two mathematical phantoms of different smoothness as test cases, for which we compute the Radon transform analytically. Section 5.3 then contains our numerical results, where we reconstruct both phantoms from finite Radon data by utilizing the discrete FBP algorithm we derived before. Furthermore, we restrict ourselves to the L^2-case, i.e., $\sigma = 0$, and use the standard root mean square error (RMSE) to measure the reconstruction error and investigate its behaviour for various low-pass filters. In our first numerical experiments we validate the proven convergence order of the error, where we in particular observe saturation at the fractional rates predicted by our theory. In addition, our second set of numerical experiments shows that the error behaves affine-linearly with respect to different quality indicators which we identified in Chapter 4.

Appendix A contains supplementary material. For the reader's convenience, in Section A.1 we review the definitions and some basic properties of the Fourier transform and the convolution product. Following this, we introduce (tempered) distributions in Section A.2 and define Sobolev spaces of fractional order in Section A.3. We close the appendix with collecting some further mathematical tools we need for our analysis in Chapters 3 and 4. All presented results are well-known and can be found, for example, in the textbooks [16], [34], [35], [46], [108], [119].

Parts of the results of this thesis have been published in the conference proceedings

- M. Beckmann and A. Iske: Error Estimates for Filtered Back Projection, *IEEE International Conference on Sampling Theory and Applications* (SampTA), 2015, 553–557,
- M. Beckmann and A. Iske: On the error behaviour of the filtered back projection, *Proceedings in Applied Mathematics and Mechanics* (PAMM) 16(1), 2016, 833–834,
- M. Beckmann and A. Iske: Sobolev Error Estimates for Filtered Back Projection Reconstructions, *IEEE International Conference on Sampling Theory and Applications* (SampTA), 2017, 251–255,

or have been disseminated in the preprints

- M. Beckmann and A. Iske: Analysis of the Inherent Reconstruction Error in Filtered Back Projection, *Hamburger Beiträge zur Angewandten Mathematik* (HBAM) 2016-01, University of Hamburg, 2016,
- M. Beckmann and A. Iske: Approximation of Bivariate Functions from Fractional Sobolev Spaces by Filtered Back Projection, *Hamburger Beiträge zur Angewandten Mathematik* (HBAM) 2017-05, University of Hamburg, 2017,
- M. Beckmann and A. Iske: Error Estimates and Convergence Rates for Filtered Back Projection, *Hamburger Beiträge zur Angewandten Mathematik* (HBAM) 2016-06, University of Hamburg, 2016, accepted for publication in *Mathematics of Computation*.

Chapter 2

Computerized tomography

In this chapter we give an overview of the basic concept of computerized tomography (CT). We start with explaining the principle of X-ray tomography and stating the underlying reconstruction problem. Following this, we introduce the Radon transform as the mathematical model for the measurement process in CT and study some of its fundamental properties. The main part of this chapter then deals with the inversion of the Radon transform, given by the classical filtered back projection formula, and the ill-posedness of the CT reconstruction problem. In the last section we finally describe some classical reconstruction techniques. We remark that most of the presented results are well-known and can be found in the monographs [30], [33], [44], [45], [51], [82], [83] or in the overview articles [93], [112], [116], [117].

2.1 The principle of X-ray tomography

The term *computerized tomography* (CT) refers to the reconstruction of a bivariate function from its line integral values. The expression 'tomography' is derived from the Ancient Greek words $\tau\acute{o}\mu o\varsigma$, which means 'slice', and $\gamma\rho\acute{\alpha}\varphi\omega$ meaning 'to write'.

One of the most prominent examples of X-ray tomography is still transmission CT in medical imaging and non-destructive testing. Here, the aim consists in recovering the interior of an unknown two-dimensional object from measurements of one-dimensional X-ray projections. In order to explore the two-dimensional structure of the object, the X-ray projections are taken from different views. To this end, a source-detector pair is rotated around the object, see Figure 2.1, where the source emits X-ray beams of a given initial intensity and the detector measures the intensity of the beams after passing the object.

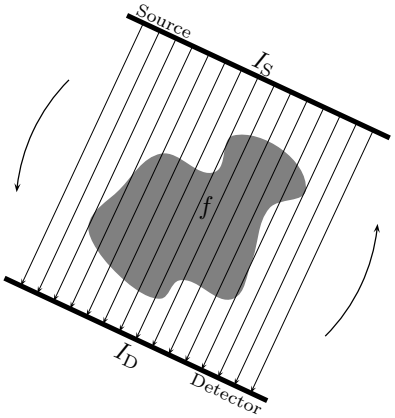


Figure 2.1: Imaging principle of X-ray tomography. To explore the inner structure of an object f, a source-detector pair is rotated around the object collecting X-ray projections at different views.

The imaging principle of X-ray tomography is now based on the fact that the X-ray beams are attenuated when passing the object. The attenuation of the X-rays depends on the inner structure of the scanned medium and, thus, carries information about the interior of the unknown object. When a single X-ray beam of known intensity travels along a line $\ell \subset \mathbb{R}^2$ from the X-ray source to the X-ray detector, a fraction of the energy present in the beam is absorbed by the material on the line ℓ and the remaining portion passes through. The intensity of the beam, as it emerges from the medium, is measured by the detector and the difference of the initial and the final intensities describes the ability of the material to absorb energy.

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ denote the so called *attenuation function* of the scanned medium, which describes the proportion of energy being absorbed by the material and whose support is assumed to lie in a convex set $\Omega \subseteq \mathbb{R}^2$. Thus, f is a characteristic quantity of the scanned object and, mathematically, the goal of X-ray tomography is to recover f from the given X-ray scans.

If the X-rays are monochromatic and $I(\mathbf{x})$ denotes the intensity of the X-ray beam at position $\mathbf{x} \in \Omega$, the intensity loss $\Delta I(\mathbf{x})$ in a small segment of ℓ of length $\Delta\mathbf{x}$ is approximately given by

$$\Delta I(\mathbf{x}) \approx -f(\mathbf{x}) \cdot I(\mathbf{x}) \cdot \Delta\mathbf{x}.$$

Taking the limit $\Delta\mathbf{x} \longrightarrow 0$ leads us to the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}} I(\mathbf{x}) = -f(\mathbf{x}) \cdot I(\mathbf{x}), \tag{2.1}$$

which is known as *Beer's law*, cf. [33, Section 1.2]. If I_{S} denotes the initial intensity of the X-ray at the source and I_{D} its final intensity at the detector, integrating (2.1) from source to detector yields

$$\int_{I_{\mathrm{S}}}^{I_{\mathrm{D}}} \frac{1}{I} \,\mathrm{d}I = -\int_{\ell} f(\mathbf{x}) \,\mathrm{d}\mathbf{x},$$

which implies that

$$\log\left(\frac{I_{\mathrm{S}}}{I_{\mathrm{D}}}\right) = \int_{\ell} f(\mathbf{x}) \,\mathrm{d}\mathbf{x}. \tag{2.2}$$

Since the values I_{S} and I_{D} are measured during the scanning process, the X-ray data provides us with the line integral values of the attenuation function f along the straight line $\ell \subset \mathbb{R}^2$. Recovering the absorption coefficient f or, equivalently, reconstructing an object from X-ray projections therefore reduces to solving the integral equation (2.2).

Consequently, the basic CT reconstruction problem can be formulated as follows.

Problem 2.1.1 (Basic reconstruction problem)**.** *Reconstruct a bivariate function $f \equiv f(x, y)$ on its domain $\Omega \subseteq \mathbb{R}^2$ from given line integral values*

$$\int_{\ell} f(x, y) \,\mathrm{d}(x, y)$$

along all straight lines $\ell \subset \mathbb{R}^2$ passing through Ω.

We have seen that the CT scanner provides line integral values of the function to be reconstructed along straight lines in the plane. In order to derive a reconstruction theory, we now introduce a suitable parametrization of these lines, as illustrated in Figure 2.2.

Definition 2.1.2 (Straight line in the plane, see [33, Definition 1.4.])**.** *For any pair $(t, \theta) \in \mathbb{R}^2$ of parameters, we define $\ell_{t,\theta} \subset \mathbb{R}^2$ to be the unique* straight line *that passes through the point $(t\cos(\theta), t\sin(\theta)) \in \mathbb{R}^2$ and is perpendicular to the unit vector $\mathbf{n}_\theta = (\cos(\theta), \sin(\theta)) \in \mathbb{R}^2$.*

We remark that every straight line in the plane can be characterized as $\ell_{t,\theta}$ for some real numbers $t, \theta \in \mathbb{R}$. Further, for any pair of parameters $(t, \theta) \in \mathbb{R}^2$ holds that

$$\ell_{t,\theta+2\pi} = \ell_{t,\theta} \quad \text{and} \quad \ell_{t,\theta+\pi} = \ell_{-t,\theta}. \tag{2.3}$$

Thus, each line in the plane has infinitely many different representations of the form $\ell_{t,\theta}$. To enforce uniqueness, we restrict the parameter pair (t,θ) to $t \in \mathbb{R}$ and $\theta \in [0,\pi)$ so that the set

$$\{\ell_{t,\theta} \mid t \in \mathbb{R},\, \theta \in [0,\pi)\}$$

contains exactly all straight lines in the plane. For fixed $t \in \mathbb{R}$ and $\theta \in [0,\pi)$ we now parametrize the line $\ell_{t,\theta}$ as follows.

Remark 2.1.3 (Parametrization of the straight line $\ell_{t,\theta}$)**.** *For any fixed angle $\theta \in [0,\pi)$, we have*

$$\mathbf{n}_\theta = (\cos(\theta), \sin(\theta)) \perp (-\sin(\theta), \cos(\theta)) = \mathbf{n}_\theta^\perp.$$

Therefore, every point (x,y) on the line $\ell_{t,\theta}$ is of the form

$$(x,y) = t \cdot \boldsymbol{n}_\theta + s \cdot \boldsymbol{n}_\theta^\perp = (t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \tag{2.4}$$

for some $s \in \mathbb{R}$ and, consequently,

$$\ell_{t,\theta} = \{(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \mid s \in \mathbb{R}\}.$$

We close this paragraph on lines in the plane by noting that for fixed angle $\theta \in [0,\pi)$ there is a unique straight line $\ell_{t,\theta}$ passing through a given point $(x,y) \in \mathbb{R}^2$.

Remark 2.1.4. *For any point $(x,y) \in \mathbb{R}^2$ in the plane and a given angle $\theta \in [0,\pi)$ there exists a unique value for $t \in \mathbb{R}$ such that the straight line $\ell_{t,\theta}$ passes through (x,y). The unique $t, s \in \mathbb{R}$ satisfying*

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta)$$

are given by

$$t = x\cos(\theta) + y\sin(\theta) \quad \text{and} \quad s = -x\sin(\theta) + y\cos(\theta)$$

and meet the relation

$$x^2 + y^2 = t^2 + s^2. \tag{2.5}$$

With the described parametrization of straight lines in the plane, we finally reformulate the basic CT reconstruction Problem 2.1.1 as follows.

Problem 2.1.5 (Basic CT reconstruction problem)**.** *On given domain $\Omega \subseteq \mathbb{R}^2$, reconstruct a bivariate function $f \equiv f(x,y)$ from its line integral values*

$$\int_{\ell_{t,\theta}} f(x,y)\, \mathrm{d}(x,y), \tag{2.6}$$

which are assumed to be given for all parameters $(t,\theta) \in \mathbb{R} \times [0,\pi)$.

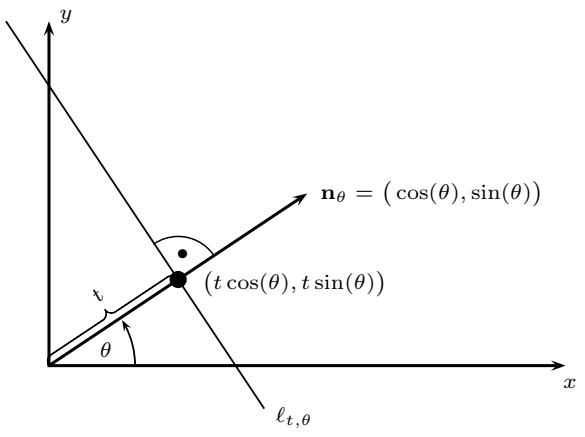


Figure 2.2: Representation of the straight line $\ell_{t,\theta} \subset \mathbb{R}^2$ with parameters $(t,\theta) \in \mathbb{R} \times [0,\pi)$.

2.2 Mathematics of computerized tomography

In this section we give an overview of some basics facts about the mathematics of computerized tomography. To this end, we follow the proceedings in [33], [82], but we also include additional theorems, relax some assumptions and extend the proofs.

2.2.1 The Radon transform

In Section 2.1 we have seen that the scanning process in CT provides us with the line integrals of a bivariate function f along straight lines in the plane, from which we have to reconstruct f. For $f \in \mathrm{L}^1(\mathbb{R}^2)$ and any parameter pair $(t, \theta) \in \mathbb{R}^2$ the line integral in (2.6) can be rewritten as

$$\int_{\ell_{t,\theta}} f(x,y)\,\mathrm{d}(x,y) = \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}s,$$

where we used the parametrization (2.4) with the arclength element

$$\|(\dot{x}(s), \dot{y}(s))\|_{\mathbb{R}^2}\,\mathrm{d}s = \sqrt{(-\sin(\theta))^2 + (\cos(\theta))^2}\,\mathrm{d}s = \mathrm{d}s.$$

The integral transform which maps a function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ into the set of its line integrals was firstly investigated by Johann Radon. In his honour, this integral transform is called *Radon transform* and, thus, the CT reconstruction Problem 2.1.5 simply seeks for the inversion of the Radon transform on $\mathbb{R}^2$.

Definition and basic properties

In the following, we investigate some basic properties of the Radon transform. For a comprehensive treatment of the Radon transform we refer the reader to the monograph [44].

Definition 2.2.1 (Radon transform, see [33, Definition 2.1.]). *Let $f \in \mathrm{L}^1(\mathbb{R}^2)$ be a bivariate function in Cartesian coordinates. Then, the* Radon transform $\mathcal{R}f$ *of f at the point $(t, \theta) \in \mathbb{R}^2$ is defined as*

$$\mathcal{R}f(t,\theta) = \int_{\ell_{t,\theta}} f(x,y)\,\mathrm{d}(x,y) = \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}s. \tag{2.7}$$

The graph of the Radon transform $\mathcal{R}f$ in the (t,θ)-plane is called sinogram *of f.*

Note that the Radon transform $\mathcal{R}f$ of a bivariate function f is 2π-periodic in the angular variable θ and, due to relation (2.3), it suffices to consider $\mathcal{R}f$ only on the domain $\mathbb{R} \times [0, \pi)$. Further, an application of Fubini's theorem shows that the Radon transform $\mathcal{R}f$ of $f \in \mathrm{L}^1(\mathbb{R}^2)$ is well-defined almost everywhere on $\mathbb{R} \times [0, \pi)$ in the sense that for any angle $\theta \in [0, \pi)$ the integral in (2.7) is well-defined for almost all $t \in \mathbb{R}$.

Proposition 2.2.2 (Well-definedness of the Radon transform). *For $f \in \mathrm{L}^1(\mathbb{R}^2)$, the Radon transform $\mathcal{R}f$ is well-defined almost everywhere on $\mathbb{R} \times [0, \pi)$.*

Proof. Let $f \in \mathrm{L}^1(\mathbb{R}^2)$. We define the auxiliary function $H : \mathbb{R}^2 \times [0, \pi) \longrightarrow \mathbb{R}^2$ as

$$H(s,t,\theta) = (t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \quad \text{for } (s,t,\theta) \in \mathbb{R}^2 \times [0,\pi).$$

Then, H is continuous on $\mathbb{R}^2 \times [0, \pi)$ and we have

$$\mathcal{R}f(t,\theta) = \int_{\mathbb{R}} f(H(s,t,\theta))\,\mathrm{d}s \quad \forall\,(t,\theta) \in \mathbb{R} \times [0,\pi).$$

Note that, for fixed angle $\theta \in [0, \pi)$, the mapping

$$(s, t) \longmapsto (t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))$$

is a rotation in $\mathbb{R}^2$ and, thus, measure preserving. In particular, the mapping

$$(s, t) \longmapsto f(H(s, t, \theta))$$

is in $\mathrm{L}^1(\mathbb{R}^2)$ for all $\theta \in [0, \pi)$, since $f \in \mathrm{L}^1(\mathbb{R}^2)$ by assumption. Therefore, Fubini's theorem shows that the partial mapping

$$s \longmapsto f(H(s, t, \theta))$$

is integrable on $\mathbb{R}$ for almost all $t \in \mathbb{R}$ and all $\theta \in [0, \pi)$. This implies that

$$(t, \theta) \longmapsto \int_{\mathbb{R}} f(H(s, t, \theta))\, \mathrm{d}s = \mathcal{R}f(t, \theta)$$

is well-defined for almost all $t \in \mathbb{R}$ and all $\theta \in [0, \pi)$. □

Remark 2.2.3 (see [33, Chapter 2.6])**.** *The Radon transform $\mathcal{R}f$ of a function f is defined everywhere on $\mathbb{R} \times [0, \pi)$ if the integral of f along the line $\ell_{t,\theta}$ exists for all pairs $(t, \theta) \in \mathbb{R} \times [0, \pi)$. To ensure this, it suffices to require that f is continuous on $\mathbb{R}^2$ and has compact support.*

With Definition 2.2.1 the classical CT reconstruction problem can be reformulated as follows.

Problem 2.2.4 (Reconstruction problem)**.** *For given domain $\Omega \subseteq \mathbb{R}^2$, reconstruct a bivariate function $f \in \mathrm{L}^1(\Omega)$ from its Radon data*

$$\{\mathcal{R}f(t, \theta) \mid t \in \mathbb{R},\, \theta \in [0, \pi)\}\,.$$

Therefore, the CT reconstruction problem seeks for the inversion of the Radon transform $\mathcal{R}$. However, before explaining the inversion of $\mathcal{R}$, we collect some of its fundamental properties.

Observation 2.2.5. *The Radon transform $\mathcal{R}$ maps a bivariate function $f \equiv f(x, y)$ in Cartesian coordinates onto a bivariate function $\mathcal{R}f \equiv \mathcal{R}f(t, \theta)$ in polar coordinates.*

The first two important properties of the Radon transform are its linearity and evenness.

Proposition 2.2.6 (Linearity of the Radon transform, see [33, Proposition 2.6.])**.** *The Radon transform $\mathcal{R}$ is a positive linear integral operator, i.e., for all $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathrm{L}^1(\mathbb{R}^2)$ we have*

$$\mathcal{R}(\alpha f + \beta g) = \alpha\, \mathcal{R}f + \beta\, \mathcal{R}g.$$

and

$$f \geq 0 \quad \Longrightarrow \quad \mathcal{R}f \geq 0.$$

Further, the Radon transform $\mathcal{R}f$ of $f \in \mathrm{L}^1(\mathbb{R}^2)$ satisfies the evenness condition

$$\mathcal{R}f(-t, \theta + \pi) = \mathcal{R}f(t, \theta) \quad \forall\, (t, \theta) \in \mathbb{R} \times [0, \pi).$$

Proof. The statement follows from the positivity and linearity of the integral and (2.3). □

The next theorem is concerned with the continuity of the Radon transform as a mapping

$$\mathcal{R} : \mathrm{L}^1(\mathbb{R}^2) \longrightarrow \mathrm{L}^1(\mathbb{R} \times [0, \pi)).$$

Proposition 2.2.7 (Continuity of the Radon transform, see [93, Theorem 2.1.])**.** *The Radon transform $\mathcal{R}$ is a continuous operator from $\mathrm{L}^1(\mathbb{R}^2)$ to $\mathrm{L}^1(\mathbb{R} \times [0, \pi))$. In particular, for $f \in \mathrm{L}^1(\mathbb{R}^2)$ we have*

$$\|\mathcal{R}f\|_{\mathrm{L}^1(\mathbb{R} \times [0, \pi))} \leq \pi\, \|f\|_{\mathrm{L}^1(\mathbb{R}^2)}.$$

Proof. Let $f \in \mathrm{L}^1(\mathbb{R}^2)$ and $\theta \in [0,\pi)$ be fixed. By the definition of the Radon transform $\mathcal{R}$ follows that

$$\begin{aligned}\int_{\mathbb{R}} |\mathcal{R}f(t,\theta)| \,\mathrm{d}t &= \int_{\mathbb{R}} \left| \int_{\ell_{t,\theta}} f(x,y) \,\mathrm{d}(x,y) \right| \mathrm{d}t \\ &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))| \,\mathrm{d}s \,\mathrm{d}t.\end{aligned}$$

Applying the transformation

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta),$$

we get $\mathrm{d}x\,\mathrm{d}y = \mathrm{d}t\,\mathrm{d}s$ and, consequently,

$$\int_{\mathbb{R}} |\mathcal{R}f(t,\theta)| \,\mathrm{d}t \leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x,y)| \,\mathrm{d}x \,\mathrm{d}y = \|f\|_{\mathrm{L}^1(\mathbb{R}^2)}.$$

This gives

$$\|\mathcal{R}f\|_{\mathrm{L}^1(\mathbb{R}\times[0,\pi))} = \int_0^\pi \int_{\mathbb{R}} |\mathcal{R}f(t,\theta)| \,\mathrm{d}t \,\mathrm{d}\theta \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^2)} \int_0^\pi 1 \,\mathrm{d}\theta = \pi \,\|f\|_{\mathrm{L}^1(\mathbb{R}^2)}.$$

Hence, $\mathcal{R}$ is a continuous operator from $\mathrm{L}^1(\mathbb{R}^2)$ to $\mathrm{L}^1(\mathbb{R}\times[0,\pi))$. □

A special situation occurs if the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is radially symmetric, i.e., f is invariant under rotations. This means that there exists a function $f_0 : \mathbb{R} \longrightarrow \mathbb{R}$ with

$$f(x,y) = f_0(\|(x,y)\|_{\mathbb{R}^2}) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Proposition 2.2.8. *If the function $f \in \mathrm{L}^1(\mathbb{R}^2)$ is radially symmetric, its Radon transform $\mathcal{R}f$ depends only on the modulus $|t|$ of the radial variable $t \in \mathbb{R}$, but not on the angle $\theta \in [0,\pi)$.*

Proof. Let $f \in \mathrm{L}^1(\mathbb{R}^2)$ be radially symmetric. Thus, there exists a function $f_0 : \mathbb{R} \longrightarrow \mathbb{R}$ with

$$f(x,y) = f_0(x^2+y^2) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

With this, for all $(t,\theta) \in \mathbb{R}\times[0,\pi)$ follows that

$$\begin{aligned}\mathcal{R}f(t,\theta) &= \int_{\ell_{t,\theta}} f(x,y) \,\mathrm{d}(x,y) = \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \,\mathrm{d}s \\ &= \int_{\mathbb{R}} f_0((t\cos(\theta) - s\sin(\theta))^2 + (t\sin(\theta) + s\cos(\theta))^2) \,\mathrm{d}s \\ &= \int_{\mathbb{R}} f_0(t^2+s^2) \,\mathrm{d}s.\end{aligned}$$

This shows that $\mathcal{R}f(t,\theta)$ is independent of the angular variable $\theta \in [0,\pi)$ and only depends on the absolute value $|t|$ of the radial variable $t \in \mathbb{R}$. □

Another important property of $\mathcal{R}$ is that the compact support of a function f carries over to its Radon transform $\mathcal{R}f$.

Proposition 2.2.9. *Let $f \in \mathrm{L}^1(\mathbb{R}^2)$ have compact support, i.e., there exists an $R > 0$ such that*

$$f(x,y) = 0 \quad \forall\,\|(x,y)\|_{\mathbb{R}^2} > R.$$

Then, $\mathcal{R}f$ has compact support as well with

$$\mathcal{R}f(t,\theta) = 0 \quad \forall\,|t| > R,\ \theta \in [0,\pi).$$

Proof. For each fixed $\theta \in [0, \pi)$ and $t \in \mathbb{R}$ with $|t| > R$, we have

$$f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) = 0 \quad \forall\, s \in \mathbb{R}.$$

This implies that

$$\mathcal{R}f(t, \theta) = 0 \quad \forall\, (t, \theta) \in \mathbb{R} \times [0, \pi) :\ |t| > R$$

and, hence, $\mathcal{R}f$ has compact support as well. □

Let $\mathscr{C}_c(\mathbb{R}^2)$ denote the space of continuous functions with compact support in $\mathbb{R}^2$, i.e.,

$$\mathscr{C}_c(\mathbb{R}^2) = \left\{ f \in \mathscr{C}(\mathbb{R}^2) \mid \exists\, R > 0 :\ \operatorname{supp}(f) \subseteq B_R(0) \right\},$$

where

$$B_R(0) = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq R^2 \right\}.$$

We now show that the Radon transform $\mathcal{R}f$ of a function $f \in \mathscr{C}_c(\mathbb{R}^2)$ is continuous on $\mathbb{R} \times [0, \pi)$.

Proposition 2.2.10. *Let $f \in \mathscr{C}_c(\mathbb{R}^2)$ have support in a compact set $K \subset \mathbb{R}^2$ with diameter*

$$\operatorname{diam}(K) = \sup \{ \|(x - X, y - Y)\|_{\mathbb{R}^2} \mid (x, y), (X, Y) \in K \} < \infty.$$

Then, we have $\mathcal{R}f \in \mathscr{C}_c(\mathbb{R} \times [0, \pi))$ and

$$|\mathcal{R}f(t, \theta)| \leq \operatorname{diam}(K)\, \|f\|_\infty \quad \forall\, (t, \theta) \in \mathbb{R} \times [0, \pi).$$

Proof. Let $f \in \mathscr{C}_c(\mathbb{R}^2)$ have compact support in $K \subset \mathbb{R}^2$ and choose $R > 0$ such that

$$f(x, y) = 0 \quad \forall\, \|(x, y)\|_{\mathbb{R}^2} > R.$$

Then, we also have $f \in \mathrm{L}^1(\mathbb{R}^2)$ and its Radon transform $\mathcal{R}f \in \mathrm{L}^1(\mathbb{R} \times [0, \pi))$ is compactly supported due to Proposition 2.2.9. To show the continuity of $\mathcal{R}f$, we fix $(t, \theta) \in \mathbb{R} \times [0, \pi)$ and consider a sequence $(t_n, \theta_n)_{n \in \mathbb{N}}$ in $\mathbb{R} \times [0, \pi)$ with

$$\lim_{n \to \infty} t_n = t \quad \text{and} \quad \lim_{n \to \infty} \theta_n = \theta.$$

For the sake of brevity, we introduce the notations

$$f_{(t,\theta)}(s) = f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \quad \text{for } s \in \mathbb{R}$$

and, analogously,

$$f_{(t_n,\theta_n)}(s) = f(t_n\cos(\theta_n) - s\sin(\theta_n), t_n\sin(\theta_n) + s\cos(\theta_n)) \quad \text{for } s \in \mathbb{R}.$$

Then, we obtain

$$\begin{aligned} |\mathcal{R}f(t, \theta) - \mathcal{R}f(t_n, \theta_n)| &= \left| \int_{\mathbb{R}} f_{(t,\theta)}(s) - f_{(t_n,\theta_n)}(s)\, \mathrm{d}s \right| = \left| \int_{-R}^{R} f_{(t,\theta)}(s) - f_{(t_n,\theta_n)}(s)\, \mathrm{d}s \right| \\ &\leq \int_{-R}^{R} \left| f_{(t,\theta)}(s) - f_{(t_n,\theta_n)}(s) \right| \mathrm{d}s, \end{aligned}$$

since for all $s \in \mathbb{R}$ with $|s| > R$ holds that

$$f_{(t,\theta)}(s) = 0 \quad \text{and} \quad f_{(t_n,\theta_n)}(s) = 0.$$

Because $f \in \mathscr{C}_c(\mathbb{R}^2)$ is continuous and compactly supported, it is also uniformly continuous and we obtain

$$\lim_{n \to \infty} \left| f_{(t,\theta)}(s) - f_{(t_n,\theta_n)}(s) \right| = 0$$

uniformly in $s \in [-R, R]$, i.e.,

$$\max_{|s|\leq R} \left| f_{(t,\theta)}(s) - f_{(t_n,\theta_n)}(s) \right| \xrightarrow{n\to\infty} 0.$$

This implies that

$$|\mathcal{R}f(t,\theta) - \mathcal{R}f(t_n,\theta_n)| \leq 2R \max_{|s|\leq R} \left| f_{(t,\theta)}(s) - f_{(t_n,\theta_n)}(s) \right| \xrightarrow{n\to\infty} 0,$$

which shows the continuity of $\mathcal{R}f$. In particular, we have proven that $\mathcal{R}f \in \mathscr{C}_c(\mathbb{R} \times [0,\pi))$. Further, since f is supported on K, for $(t,\theta) \in \mathbb{R} \times [0,\pi)$ holds that

$$\begin{aligned} |\mathcal{R}f(t,\theta)| &= \left| \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s \right| \\ &= \left| \int_{\mathbb{R}} (\chi_K \cdot f)(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s \right|, \end{aligned}$$

where χ_K denotes the characteristic function of K. Consequently,

$$\begin{aligned} |\mathcal{R}f(t,\theta)| &\leq \|f\|_\infty \int_{\mathbb{R}} \chi_K(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s \\ &= \|f\|_\infty \int_{\ell_{t,\theta}} \chi_K(x,y)\, \mathrm{d}(x,y) = \mathcal{R}\chi_K(t,\theta)\, \|f\|_\infty \\ &\leq \operatorname{diam}(K)\, \|f\|_\infty, \end{aligned}$$

as stated. □

In particular, Proposition 2.2.10 implies that the Radon transform $\mathcal{R}$ is continuous as a mapping

$$\mathcal{R} : \mathscr{C}(B_R(0)) \longrightarrow \mathscr{C}([-R,R] \times [0,\pi)).$$

We remark that the Radon transform also preserves smoothness and decay properties of the input function. This has been proven by Helgason in [43]. To state the result, recall that the Schwartz space $\mathcal{S}(\mathbb{R}^2)$ on $\mathbb{R}^2$ is given by all smooth functions $f \equiv f(x,y) \in \mathscr{C}^\infty(\mathbb{R}^2)$ which decay with all their derivatives faster than any power of $1/\|(x,y)\|_{\mathbb{R}^2}$ at infinity, i.e.,

$$\sup_{(x,y)\in\mathbb{R}^2} |(x,y)^\alpha\, \mathrm{D}^\beta f(x,y)| < \infty \quad \forall\, \alpha, \beta \in \mathbb{N}_0^2.$$

Analogously, the Schwartz space $\mathcal{S}(\mathbb{R} \times [0,\pi))$ can be defined on $\mathbb{R} \times [0,\pi)$ as the space of all functions $g \equiv g(t,\theta)$ that can be extended to be smooth and 2π-periodic in the angular variable θ satisfying

$$g(t,\theta) = g(-t, \theta - \pi) \quad \forall\, (t,\theta) \in \mathbb{R} \times [\pi, 2\pi)$$

and where decay is understood with respect to the radial variable t in the sense that g (along with all its derivatives in t) decays faster than any power of $1/|t|$ at infinity uniformly in θ. Then, in [43, Theorem 4.1] it is proven that the Radon transform $\mathcal{R}$ is continuous as a mapping

$$\mathcal{R} : \mathcal{S}(\mathbb{R}^2) \longrightarrow \mathcal{S}(\mathbb{R} \times [0,\pi)).$$

The Fourier slice theorem

The most important property of the Radon transform is given by the classical *Fourier slice theorem* (FST), also known as *central slice theorem*, which relates the Fourier transform of the Radon transform to the Fourier transform of the function to be reconstructed.

Let us first recall that the Fourier transform $\mathcal{F}f$ of a bivariate function $f \equiv f(x,y) \in \mathrm{L}^1(\mathbb{R}^2)$ in Cartesian coordinates is given by

$$\mathcal{F}f(X,Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)\, \mathrm{e}^{-\mathrm{i}(xX+yY)}\, \mathrm{d}x\, \mathrm{d}y \quad \text{for } (X,Y) \in \mathbb{R}^2.$$

For a bivariate function $h \equiv h(t,\theta)$ in polar coordinates satisfying $h(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R})$ for all $\theta \in [0,\pi)$ we define the Fourier transform $\mathcal{F}h$ as the univariate Fourier transform acting only on the radial variable t, i.e.,

$$\mathcal{F}h(S,\theta) = \int_{\mathbb{R}} h(t,\theta)\, \mathrm{e}^{-\mathrm{i}St}\, \mathrm{d}t \quad \text{for } (S,\theta) \in \mathbb{R} \times [0,\pi).$$

Now, the Fourier slice theorem reads as follows.

Theorem 2.2.11 (Fourier slice theorem, see [33, Theorem 6.1.]). *For $f \in \mathrm{L}^1(\mathbb{R}^2)$ we have*

$$\mathcal{F}(\mathcal{R}f)(S,\theta) = \mathcal{F}f(S\cos(\theta), S\sin(\theta)) \quad \forall\, (S,\theta) \in \mathbb{R} \times [0,\pi). \qquad \square$$

More generally, we have the following projection slice theorem.

Theorem 2.2.12 (General projection slice theorem, see [93, Theorem 2.2]). *Let $f \in \mathrm{L}^1(\mathbb{R}^2)$ and $h \in \mathrm{L}^\infty(\mathbb{R})$. Then, for all $\theta \in [0,\pi)$ we have*

$$\int_{\mathbb{R}} \mathcal{R}f(t,\theta)\, h(t)\, \mathrm{d}t = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)\, h(x\cos(\theta) + y\sin(\theta))\, \mathrm{d}x\, \mathrm{d}y.$$

The Fourier slice Theorem 2.2.11 follows by using the function $h = \mathrm{e}^{-\mathrm{i}S\cdot}$ with fixed $S \in \mathbb{R}$.

Proof. We first note that for $f \in \mathrm{L}^1(\mathbb{R}^2)$, $h \in \mathrm{L}^\infty(\mathbb{R})$ and fixed angle $\theta \in [0,\pi)$, the function $F_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}$ with

$$F_\theta(x,y) = f(x,y)\, h(x\cos(\theta) + y\sin(\theta)) \quad \text{for } (x,y) \in \mathbb{R}^2$$

satisfies $F_\theta \in \mathrm{L}^1(\mathbb{R}^2)$ with

$$\|F_\theta\|_{\mathrm{L}^1(\mathbb{R}^2)} \leq \|h\|_{\mathrm{L}^\infty(\mathbb{R})}\, \|f\|_{\mathrm{L}^1(\mathbb{R}^2)}.$$

Applying the transformation

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta),$$

i.e.,

$$t = x\cos(\theta) + y\sin(\theta) \quad \text{and} \quad s = -x\sin(\theta) + y\cos(\theta),$$

again gives $\mathrm{d}x\, \mathrm{d}y = \mathrm{d}t\, \mathrm{d}s$ and, thus, it follows that

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} F_\theta(x,y)\, \mathrm{d}x\, \mathrm{d}y &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, h(t)\, \mathrm{d}t\, \mathrm{d}s \\ &= \int_{\mathbb{R}} \mathcal{R}f(t,\theta)\, h(t)\, \mathrm{d}t \end{aligned}$$

by Fubini's theorem and the definition of the Radon transform $\mathcal{R}$. $\square$

The importance of the Fourier slice theorem lies in the fact that it links together the Radon transform of a function and its Fourier transform. This connection can be used to derive properties of the Radon transform from those properties which are known for the Fourier transform. In particular, the Fourier slice theorem shows the injectivity of the Radon transform on the domain $\mathrm{L}^1(\mathbb{R}^2)$. Indeed, if $\mathcal{R}f$ vanished on $\mathbb{R} \times [0,\pi)$, then $\mathcal{F}f$ vanished on $\mathbb{R}^2$, which implies that f is zero due to the injectivity of the Fourier transform, see Corollary A.1.6.

Corollary 2.2.13 (Injectivity of the Radon transform)**.** *For $f \in \mathrm{L}^1(\mathbb{R}^2)$ we have*

$$\mathcal{R}f = 0 \quad \Longrightarrow \quad f = 0,$$

i.e., the Radon transform $\mathcal{R}$ is injective on $\mathrm{L}^1(\mathbb{R}^2)$. □

Moreover, Theorem 2.2.11 immediately provides a scheme for reconstructing a function from the knowledge of its Radon transform. Assuming $\mathcal{R}f(t,\theta)$ to be known for all $(t,\theta) \in \mathbb{R} \times [0,\pi)$, we can gain knowledge about the two-dimensional Fourier transform of f by computing the one-dimensional Fourier transform of $\mathcal{R}f$. Subsequent application of the inverse two-dimensional Fourier transform would yield the function f we want to reconstruct. We remark that such reconstruction procedures are known as *Fourier reconstruction methods*, cf. [82], [83].

We close this section on the Radon transform by considering an analytical example.

Example 2.2.14. *Consider the characteristic function $\chi_{B_R(0)}$ of the ball $B_R(0) \subset \mathbb{R}^2$ of radius $R > 0$ around 0, i.e.,*

$$\chi_{B_R(0)}(x,y) = \begin{cases} 1 & \text{for } x^2 + y^2 \leq R^2 \\ 0 & \text{for } x^2 + y^2 > R^2. \end{cases}$$

For all $(t,\theta) \in \mathbb{R} \times [0,\pi)$ we then have

$$\chi_{B_R(0)}(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) = \begin{cases} 1 & \text{for } t^2 + s^2 \leq R^2 \\ 0 & \text{for } t^2 + s^2 > R^2 \end{cases}$$

and, thus, for the Radon transform of $\chi_{B_R(0)}$ follows that

$$\begin{aligned} \mathcal{R}\chi_{B_R(0)}(t,\theta) &= \int_{\mathbb{R}} \chi_{B_R(0)}(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}s \\ &= \begin{cases} 2\sqrt{R^2 - t^2} & \text{for } |t| \leq R \\ 0 & \text{for } |t| > R. \end{cases} \end{aligned}$$

2.2.2 The back projection

We want to recover the function $f \equiv f(x,y)$ from the values $\mathcal{R}f(t,\theta)$ with $t \in \mathbb{R}$ and $\theta \in [0,\pi)$. First, we observe that each fixed point $(x_0,y_0) \in \mathbb{R}^2$ lies on infinitely many different lines $\ell_{t,\theta}$. But for a fixed angle $\theta \in [0,\pi)$ there exists exactly one $t \in \mathbb{R}$ for which $\ell_{t,\theta}$ passes through the point (x_0,y_0). Indeed, for suitable $s \in \mathbb{R}$ we have the representation

$$x_0 = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y_0 = t\sin(\theta) + s\cos(\theta)$$

if and only if $t = x_0\cos(\theta) + y_0\sin(\theta)$, see Remark 2.1.4. Consequently, the lines passing through (x_0,y_0) are of the form

$$\ell_{x_0\cos(\theta)+y_0\sin(\theta),\theta} \quad \text{for } \theta \in [0,\pi).$$

The first naive idea is now the following: To recover $f(x_0,y_0)$, we compute the average value of the line integrals

$$\mathcal{R}f(x_0\cos(\theta) + y_0\sin(\theta), \theta) \quad \text{for } \theta \in [0,\pi)$$

over all lines passing through (x_0,y_0). This operation is called *back projection.*

Definition 2.2.15 (Back projection, see [33, Definition 3.2.])**.** *Let $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$ be a bivariate function in polar coordinates. Then, the* back projection *$\mathcal{B}g$ of g at the point $(x,y) \in \mathbb{R}^2$ is defined as*

$$\mathcal{B}g(x,y) = \frac{1}{\pi}\int_0^\pi g(x\cos(\theta) + y\sin(\theta), \theta)\,\mathrm{d}\theta.$$

The following proposition shows that the back projection $\mathcal{B}g$ of a function $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$ is defined almost everywhere and locally integrable on $\mathbb{R}^2$.

Proposition 2.2.16 (Mapping property of the back projection)**.** *For $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$, the back projection $\mathcal{B}g$ is defined almost everywhere on $\mathbb{R}^2$ and satisfies $\mathcal{B}g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$.*

Proof. Let $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$. We define the auxiliary function $H : \mathbb{R}^2 \times [0,\pi) \longrightarrow \mathbb{R} \times [0,\pi)$ as

$$H(x,y,\theta) = (x\cos(\theta) + y\sin(\theta), \theta) \quad \text{for } (x,y,\theta) \in \mathbb{R}^2 \times [0,\pi).$$

Then, H is continuous on $\mathbb{R}^2 \times [0,\pi)$ and we have

$$\mathcal{B}g(x,y) = \frac{1}{\pi} \int_0^\pi g(H(x,y,\theta))\,\mathrm{d}\theta \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Now, let $K \subset \mathbb{R}^2$ be an arbitrary compact subset of $\mathbb{R}^2$. We then obtain

$$\begin{aligned}
\|\mathcal{B}g\|_{\mathrm{L}^1(K)} &= \int_K |\mathcal{B}g(x,y)|\,\mathrm{d}(x,y) = \int_K \left|\frac{1}{\pi}\int_0^\pi g(H(x,y,\theta))\,\mathrm{d}\theta\right| \mathrm{d}(x,y) \\
&= \frac{1}{\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \left|\int_0^\pi g(H(x,y,\theta))\,\mathrm{d}\theta\right| \chi_K(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&\leq \frac{1}{\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \left(\int_0^\pi |g(H(x,y,\theta))|\,\mathrm{d}\theta\right) \chi_K(x,y)\,\mathrm{d}x\,\mathrm{d}y,
\end{aligned}$$

where χ_K denotes the characteristic function of K. By applying Fubini's theorem for non-negative functions and integration by substitution for real-valued functions with

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta),$$

i.e., $\mathrm{d}x\,\mathrm{d}y = \mathrm{d}t\,\mathrm{d}s$ and

$$t = x\cos(\theta) + y\sin(\theta) \quad \text{and} \quad s = -x\sin(\theta) + y\cos(\theta),$$

we obtain

$$\begin{aligned}
\|\mathcal{B}g\|_{\mathrm{L}^1(K)} &\leq \frac{1}{\pi}\int_0^\pi \int_{\mathbb{R}}\int_{\mathbb{R}} |g(x\cos(\theta) + y\sin(\theta),\theta)|\,\chi_K(x,y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^\pi \int_{\mathbb{R}}\int_{\mathbb{R}} |g(t,\theta)|\,\chi_K(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}t\,\mathrm{d}s\,\mathrm{d}\theta.
\end{aligned}$$

Using again Fubini's theorem, the definition of the Radon transform $\mathcal{R}$ yields

$$\begin{aligned}
\|\mathcal{B}g\|_{\mathrm{L}^1(K)} &\leq \frac{1}{\pi}\int_0^\pi \int_{\mathbb{R}} |g(t,\theta)| \left(\int_{\mathbb{R}} \chi_K(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,ds\right) \mathrm{d}t\,\mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^\pi \int_{\mathbb{R}} |g(t,\theta)|\,\mathcal{R}\chi_K(t,\theta)\,\mathrm{d}t\,\mathrm{d}\theta \leq \frac{1}{\pi}\,\mathrm{diam}(K)\,\|g\|_{\mathrm{L}^1(\mathbb{R}\times[0,\pi))} < \infty.
\end{aligned}$$

Consequently, we have $\mathcal{B}g \in \mathrm{L}^1(K)$ for all compact subsets $K \subset \mathbb{R}^2$. In particular, this shows that the back projection $\mathcal{B}g$ of $g \in \mathrm{L}^1(\mathbb{R}\times[0,\pi))$ is defined almost everywhere on $\mathbb{R}^2$ and satisfies

$$\mathcal{B}g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2). \qquad \square$$

Note that the back projection $\mathcal{B}g$ of an essentially bounded function $g \in \mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$ is also defined almost everywhere and essentially bounded on $\mathbb{R}^2$. Moreover, the back projection is continuous as a mapping

$$\mathcal{B} : \mathrm{L}^\infty(\mathbb{R} \times [0,\pi)) \longrightarrow \mathrm{L}^\infty(\mathbb{R}^2),$$

where for $g \in \mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$

$$\|\mathcal{B}g\|_{\mathrm{L}^\infty(\mathbb{R}^2)} \leq \|g\|_{\mathrm{L}^\infty(\mathbb{R}\times[0,\pi))}.$$

We continue with some basic properties of the back projection operator $\mathcal{B}$.

Observation 2.2.17. *The back projection operator $\mathcal{B}$ maps a bivariate function $g \equiv g(t,\theta)$ in polar coordinates onto a bivariate function $\mathcal{B}g \equiv \mathcal{B}g(x,y)$ in Cartesian coordinates.*

As the Radon transform $\mathcal{R}$, the back projection operator $\mathcal{B}$ is a positive linear operator.

Proposition 2.2.18 (Linearity of the back projection, see [33, Proposition 3.3.])**.** *The back projection $\mathcal{B}$ is a positive linear integral operator, i.e., for all $\alpha, \beta \in \mathbb{R}$ and $g, h \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$ we have*

$$\mathcal{B}(\alpha g + \beta h) = \alpha\, \mathcal{B}g + \beta\, \mathcal{B}h$$

and

$$g \geq 0 \quad \Longrightarrow \quad \mathcal{B}g \geq 0.$$

Proof. The statement follows from the positivity and linearity of the integral. □

However, we will observe that the back projection is *not* the inverse of the Radon transform. To this end, we first prove the following auxiliary result.

Proposition 2.2.19. *Let $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$ be independent of the angular variable $\theta \in [0,\pi)$ and only depend on the absolute value $|t|$ of the radial variable $t \in \mathbb{R}$. Then, its back projection $\mathcal{B}g$ is radially symmetric.*

Proof. Since $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$ is assumed to depend only on the absolute value $|t|$ of the radial variable $t \in \mathbb{R}$, there exists a function $g_0 \in \mathrm{L}^1(\mathbb{R})$ such that

$$g(t,\theta) = g_0(|t|) \quad \forall\, (t,\theta) \in \mathbb{R} \times [0,\pi).$$

Consider now an arbitrary rotation matrix $Q \in \mathrm{SO}(2)$ given by

$$Q = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \quad \text{for } \varphi \in [0, 2\pi).$$

To prove the radial symmetry of $\mathcal{B}g$, we have to show that

$$\mathcal{B}g(Q(x,y)) = \mathcal{B}g(x,y) \quad \forall\, (x,y) \in \mathbb{R}^2.$$

For fixed $(x,y) \in \mathbb{R}^2$, the definition of the back projection $\mathcal{B}$ gives

$$\begin{aligned}
\mathcal{B}g(Q(x,y)) &= \mathcal{B}g(x\cos(\varphi) - y\sin(\varphi), x\sin(\varphi) + y\cos(\varphi)) \\
&= \frac{1}{\pi} \int_0^\pi g((x\cos(\varphi) - y\sin(\varphi))\cos(\theta) + (x\sin(\varphi) + y\cos(\varphi))\sin(\theta), \theta)\, \mathrm{d}\theta \\
&= \frac{1}{\pi} \int_0^\pi g(x\cos(\theta - \varphi) + y\sin(\theta - \varphi), \theta)\, \mathrm{d}\theta \\
&= \frac{1}{\pi} \int_0^\pi g_0(|x\cos(\theta - \varphi) + y\sin(\theta - \varphi)|)\, \mathrm{d}\theta.
\end{aligned}$$

To continue our calculations, we note that the mapping

$$\theta \longmapsto g_0(|x\cos(\theta) + y\sin(\theta)|)$$

is π-periodic. Consequently, the transformation $\vartheta = \theta - \varphi$ gives

$$\begin{aligned}
\mathcal{B}g(Q(x,y)) &= \frac{1}{\pi} \int_{-\varphi}^{\pi-\varphi} g_0(|x\cos(\vartheta) + y\sin(\vartheta)|)\, \mathrm{d}\vartheta \\
&= \frac{1}{\pi} \int_0^\pi g_0(|x\cos(\vartheta) + y\sin(\vartheta)|)\, \mathrm{d}\vartheta \\
&= \frac{1}{\pi} \int_0^\pi g(x\cos(\vartheta) + y\sin(\vartheta), \vartheta)\, \mathrm{d}\vartheta = \mathcal{B}g(x,y),
\end{aligned}$$

which shows that $\mathcal{B}g$ is radially symmetric. □

With Proposition 2.2.19 we are now prepared to prove that back projecting the measured Radon data $\mathcal{R}f$ does not recover the unknown function f.

Observation 2.2.20. *The back projection $\mathcal{B}$ is not the inverse of the Radon transform $\mathcal{R}$.*

Proof. To prove said statement, we give a counterexample. To this end, we consider the characteristic function $f = \chi_{B_R(0)} \in \mathrm{L}^1(\mathbb{R}^2)$ of the ball $B_R(0) \subset \mathbb{R}^2$ of radius $R > 0$ around 0, i.e.,

$$f(x,y) = \begin{cases} 1 & \text{for } x^2 + y^2 \leq R^2 \\ 0 & \text{for } x^2 + y^2 > R^2, \end{cases}$$

whose nonnegative Radon transform is given by

$$\mathcal{R}f(t,\theta) = \begin{cases} 2\sqrt{R^2 - t^2} & \text{for } |t| \leq R \\ 0 & \text{for } |t| > R \end{cases} \quad \forall\, (t,\theta) \in \mathbb{R} \times [0,\pi),$$

see Example 2.2.14.

We now evaluate the back projection $\mathcal{B}(\mathcal{R}f)$ of $\mathcal{R}f$ at the points $(R+\varepsilon, 0)$ with $\varepsilon > 0$. For $\varepsilon \in (0, (\sqrt{2}-1)R)$ we have $R + \varepsilon \in (R, \sqrt{2}R)$ and, thus,

$$|(R+\varepsilon)\cos(\theta)| < R \quad \forall\, \theta \in \Big(\frac{\pi}{4}, \frac{3\pi}{4}\Big).$$

Consequently, we obtain

$$\begin{aligned} \mathcal{B}(\mathcal{R}f)(R+\varepsilon, 0) &= \frac{1}{\pi} \int_0^{\pi} \mathcal{R}f((R+\varepsilon)\cos(\theta), \theta)\, \mathrm{d}\theta \geq \frac{1}{\pi} \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} \mathcal{R}f((R+\varepsilon)\cos(\theta), \theta)\, \mathrm{d}\theta \\ &= \frac{2}{\pi} \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} \sqrt{R^2 - (R+\varepsilon)^2 \cos^2(\theta)}\, \mathrm{d}\theta > 0. \end{aligned}$$

Now, the Radon transform $\mathcal{R}f$ of f is integrable on $\mathbb{R} \times [0,\pi)$ and, further, only depends on the absolute value $|t|$ of the radial variable $t \in \mathbb{R}$. Consequently, applying Proposition 2.2.19 shows that $\mathcal{B}(\mathcal{R}f)$ is radially symmetric and, for any angle $\varphi \in [0, 2\pi)$, we obtain

$$\mathcal{B}(\mathcal{R}f)((R+\varepsilon)\cos(\varphi), (R+\varepsilon)\sin(\varphi)) > 0 \quad \forall\, \varepsilon \in (R, \sqrt{2}R),$$

i.e., $\mathcal{B}(\mathcal{R}f)$ is positive on the open annulus

$$B_R^{\sqrt{2}R}(0) = \Big\{(x,y) \in \mathbb{R}^2) \mid R^2 \leq x^2 + y^2 \leq 2R^2\Big\} \subset \mathbb{R}^2.$$

However, the function f vanishes on $B_R^{\sqrt{2}R}(0)$, which shows that f is not reconstructed by the back projection $\mathcal{B}(\mathcal{R}f)$ of its Radon transform $\mathcal{R}f$, i.e.,

$$\mathcal{B}(\mathcal{R}f) \neq f. \qquad \square$$

We close this paragraph on the back projection operator by stating the following useful relation between the convolution product, the back projection and the Radon transform. Recall that the convolution product of two bivariate functions $f \equiv f(x,y)$, $g \equiv g(x,y) \in \mathrm{L}^1(\mathbb{R}^2)$ in Cartesian coordinates is given by

$$(f * g)(x,y) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(X,Y)\, g(x-X, y-Y)\, \mathrm{d}X\, \mathrm{d}Y \quad \text{for } (x,y) \in \mathbb{R}^2.$$

Further, we define the convolution product of two bivariate functions $h \equiv h(t,\theta)$, $k \equiv k(t,\theta)$ in polar coordinates satisfying $h(\cdot,\theta)$, $k(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R})$ for all $\theta \in [0,\pi)$ as

$$(h * k)(t,\theta) = \int_{\mathbb{R}} h(S,\theta)\, k(t-S,\theta)\, \mathrm{d}S \quad \text{for } (t,\theta) \in \mathbb{R} \times [0,\pi).$$

The result now reads as follows.

Theorem 2.2.21 (See [82, Theorem II.1.3]). *Let $f \equiv f(x,y) \in \mathrm{L}^1(\mathbb{R}^2)$ be a bivariate function in Cartesian coordinates and let $g \equiv g(t,\theta) \in \mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$ be a function in polar coordinates. Then, we have*

$$\mathcal{B}g * f = \mathcal{B}(g * \mathcal{R}f). \tag{2.8}$$

Proof. First of all, we note that for $f \in \mathrm{L}^1(\mathbb{R}^2)$ and $g \in \mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$ both expressions in (2.8) are well-defined as functions in $\mathrm{L}^\infty(\mathbb{R}^2)$. Indeed, for $g \in \mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$ its back projection $\mathcal{B}g$ is essentially bounded on $\mathbb{R}^2$, i.e., $\mathcal{B}g \in \mathrm{L}^\infty(\mathbb{R}^2)$, and, therefore, Young's inequality, Theorem A.1.16, shows that

$$\mathcal{B}g * f \in \mathrm{L}^\infty(\mathbb{R}^2).$$

On the other hand, for $f \in \mathrm{L}^1(\mathbb{R}^2)$ we have $\mathcal{R}f(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R})$ for all $\theta \in [0,\pi)$ according to Proposition 2.2.7 with

$$\|\mathcal{R}f(\cdot,\theta)\|_{\mathrm{L}^1(\mathbb{R})} \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^2)} \quad \forall\, \theta \in [0,\pi).$$

Consequently, the convolution product of $\mathcal{R}f$ and g is in $\mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$ and we also obtain

$$\mathcal{B}(g * \mathcal{R}f) \in \mathrm{L}^\infty(\mathbb{R}^2).$$

Now, for $(X,Y) \in \mathbb{R}^2$ the definitions of the back projection and the convolution product give

$$(\mathcal{B}g * f)(X,Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{B}g(X-x, Y-y)\, f(x,y)\, \mathrm{d}x\, \mathrm{d}y,$$

where

$$\mathcal{B}g(X-x, Y-y) = \frac{1}{\pi} \int_0^\pi g((X-x)\cos(\theta) + (Y-y)\sin(\theta), \theta)\, \mathrm{d}\theta.$$

By substituting

$$t = x\cos(\theta) + y\sin(\theta) \quad \text{and} \quad s = -x\sin(\theta) + y\cos(\theta),$$

we get $\mathrm{d}x\, \mathrm{d}y = \mathrm{d}s\, \mathrm{d}t$ and, therefore, with

$$\mathcal{R}f(t,\theta) = \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s$$

we can conclude that

$$\begin{aligned}
(\mathcal{B}g * f)(X,Y) &= \frac{1}{\pi} \int_0^\pi \int_{\mathbb{R}} g(X\cos(\theta) + Y\sin(\theta) - t, \theta)\, \mathcal{R}f(t,\theta)\, \mathrm{d}t\, \mathrm{d}\theta \\
&= \frac{1}{\pi} \int_0^\pi (g * \mathcal{R}f)(X\cos(\theta) + Y\sin(\theta), \theta)\, \mathrm{d}\theta \\
&= \mathcal{B}(g * \mathcal{R}f)(X,Y),
\end{aligned}$$

as stated. □

2.2.3 The filtered back projection formula

Based on the Fourier slice Theorem 2.2.11, we are now prepared to prove an inversion formula for the Radon transform, which is given by the classical *filtered back projection (FBP) formula.*

Theorem 2.2.22 (Filtered back projection formula, see [33, Theorem 6.2.]). *For each bivariate function $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathscr{C}(\mathbb{R}^2)$ with $\mathcal{F}f \in \mathrm{L}^1(\mathbb{R}^2)$ the* filtered back projection formula

$$f(x,y) = \frac{1}{2}\, \mathcal{B}(\mathcal{F}^{-1}[|S|\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \tag{2.9}$$

holds for all $(x,y) \in \mathbb{R}^2$.

We remark that the FBP formula is also valid under weaker assumptions on the function f. For the purpose of this work, however, the presented version is sufficient.

Proof. Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathscr{C}(\mathbb{R}^2)$ with $\mathcal{F}f \in \mathrm{L}^1(\mathbb{R}^2)$ and let $(x, y) \in \mathbb{R}^2$ be fixed. Applying the two-dimensional Fourier inversion formula, Theorem A.1.5, to f yields the identity

$$f(x,y) = \mathcal{F}^{-1}(\mathcal{F}f)(x,y) = \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}f(X,Y)\, \mathrm{e}^{\mathrm{i}(xX+yY)}\, \mathrm{d}X\, \mathrm{d}Y.$$

By changing the variables $(X, Y) \in \mathbb{R}^2$ from Cartesian coordinates to $(S, \theta) \in \mathbb{R} \times [0, \pi)$ in polar coordinates, i.e.,

$$X = S\cos(\theta) \quad \text{and} \quad Y = S\sin(\theta),$$

we get $\mathrm{d}X\, \mathrm{d}Y = |S|\, \mathrm{d}S\, \mathrm{d}\theta$. Thus, with the Fourier slice Theorem 2.2.11 follows that

$$\begin{aligned} f(x,y) &= \frac{1}{4\pi^2} \int_0^\pi \int_{\mathbb{R}} \mathcal{F}f(S\cos(\theta), S\sin(\theta))\, \mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\, |S|\, \mathrm{d}S\, \mathrm{d}\theta \\ &\overset{\text{FST}}{=} \frac{1}{4\pi^2} \int_0^\pi \int_{\mathbb{R}} \mathcal{F}(\mathcal{R}f)(S,\theta)\, \mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\, |S|\, \mathrm{d}S\, \mathrm{d}\theta \\ &= \frac{1}{2\pi} \int_0^\pi \mathcal{F}^{-1}[|S|\, \mathcal{F}(\mathcal{R}f)(S,\theta)](x\cos(\theta) + y\sin(\theta), \theta)\, \mathrm{d}\theta \\ &= \frac{1}{2}\, \mathcal{B}(\mathcal{F}^{-1}[|S|\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \end{aligned}$$

due to the definition of the back projection. □

We remark that without the factor $|S|$ in (2.9), the Fourier transform and its inverse would cancel out and the FBP formula world reduce to simply applying the back projection operator $\mathcal{B}$ to the Radon data $\mathcal{R}f$. However, as we have seen in Observation 2.2.20, this is not sufficient for the exact recovery of the function f.

Remark 2.2.23. *The FBP formula* (2.9) *reveals that multiplying the Fourier transform of $\mathcal{R}f$ with $|S|$ and applying the inverse Fourier transform is essential before back projecting the Radon data. In the language of signal processing, we say that the Radon data $\mathcal{R}f$ is* filtered *by the multiplication with the filter $|S|$ in Fourier domain, which also explains the expression* filtered back projection.

With Theorem 2.2.22 the stated CT reconstruction Problem 2.2.4 is solved *analytically*. In practice, however, the application of the FBP formula (2.9) causes severe numerical problems.

Observation 2.2.24 (FBP is unstable)**.** *By the application of the filter $|S|$ to the Fourier transform $\mathcal{F}(\mathcal{R}f)$ in* (2.9)*, especially the high frequency components in $\mathcal{R}f$ are amplified by the magnitude of $|S|$. Since noise mainly consists of high frequencies, this shows that the filtered back projection formula is in particular highly sensitive with respect to noise and, thus, numerically unstable. In practice, a direct application of the FBP formula would lead to undesired corruptions in the reconstruction.*

This observation can be explained by analysing the ill-posedness of the CT reconstruction Problem 2.2.4. Before that, we first want to pass the following remark.

Remark 2.2.25. *The filtered back projection formula assumes the Radon data $\mathcal{R}f(t, \theta)$ to be available for all straight lines $\ell_{t,\theta}$ in the plane. But in practice, only a finite number of X-ray samples are taken and, consequently, we have to recover the function f from a finite set of Radon data*

$$\{\mathcal{R}f(t_j, \theta_j) \mid j = 1, \ldots, M\} \quad \text{for some } M \in \mathbb{N}.$$

Ill-posedness of the CT reconstruction problem

In the previous paragraph we have presented an inversion formula for the Radon transform, given by the filtered back projection formula. For the application of this reconstruction formula to real data, we have to assume that perfect Radon data are given. However, this is not a realistic scenario, because in practice the data is always corrupted by noise.

Therefore, it is important to know how small perturbations in the measurement data are propagated through the reconstruction process. In particular, we have to analyse whether the inversion of the Radon transform is continuous, i.e., whether small measurement errors only lead to small errors in the reconstruction. This leads us to Hadamard's classical definition of *well-posed* problems and the notion of *ill-posedness*.

For a comprehensive treatment of ill-posed problems and the theory of inverse problems and their regularization, we refer the reader to the monographs [29], [59], [99].

Definition 2.2.26 (Well-posedness, see [99, Definition 1.5.2]). *Let $A : \mathcal{X} \longrightarrow \mathcal{Y}$ be a mapping between topological spaces $\mathcal{X}$ and $\mathcal{Y}$. The problem $Ax = y$ is called* well-posed *in the sense of Hadamard, if the following three conditions are satisfies:*

(i) Existence: *For every $y \in \mathcal{Y}$ there exists an $x \in \mathcal{X}$ such that $Ax = y$.*

(ii) Uniqueness: *For every $y \in \mathcal{Y}$ the solution $x \in \mathcal{X}$ of $Ax = y$ is uniquely determined.*

(iii) Stability: *The inverse mapping $A^{-1} : \mathcal{Y} \longrightarrow \mathcal{X}$ is continuous, i.e., the solution $x \in \mathcal{X}$ depends continuously on the data $y \in \mathcal{Y}$.*

If at least one of the above conditions is violated, the problem is called ill-posed.

A standard method to classify the degree of ill-posedness of an operator A is based on the decay behaviour of its singular values, see [29, Section 2.2]. If A is defined on L^2-spaces, another classification of ill-posedness can be defined in terms of Sobolev spaces H^α, see [59, Section 3.2].

Definition 2.2.27 (Degree of ill-posedness). *Let $A : \mathcal{X} \longrightarrow \mathcal{Y}$ be an operator between L^2-spaces $\mathcal{X}$ and $\mathcal{Y}$. The problem $Ax = y$ is called* ill-posed of degree $\alpha > 0$ *if, for some $C_1, C_2 > 0$,*

$$C_1 \|x\|_{\mathrm{L}^2} \leq \|Ax\|_{\mathrm{H}^\alpha} \leq C_2 \|x\|_{\mathrm{L}^2}.$$

We now state a continuity result for the Radon transform on Sobolev spaces of fractional order, which implies that the CT reconstruction Problem 2.2.4 is ill-posed of degree $\frac{1}{2}$.

Let us first recall the definitions of the involved Sobolev spaces, see also Appendix A.3. For functions $f \equiv f(x, y)$ in Cartesian coordinates $(x, y) \in \mathbb{R}^2$ the Sobolev space $\mathrm{H}^\alpha(\mathbb{R}^2)$ of order $\alpha \in \mathbb{R}$, defined as

$$\mathrm{H}^\alpha(\mathbb{R}^2) = \left\{ f \in \mathcal{S}'(\mathbb{R}^2) \mid \|f\|_{\mathrm{H}^\alpha(\mathbb{R}^2)} < \infty \right\},$$

is equipped with the norm

$$\|f\|_{\mathrm{H}^\alpha(\mathbb{R}^2)} = \left(\int_{\mathbb{R}} \int_{\mathbb{R}} (1 + X^2 + Y^2)^\alpha \, |\mathcal{F}f(X, Y)|^2 \, \mathrm{d}X \, \mathrm{d}Y \right)^{1/2}.$$

Further, for an open subset $\Omega \subset \mathbb{R}^2$ the space $\mathrm{H}_0^\alpha(\Omega)$ consists of those Sobolev functions with support in $\overline{\Omega}$, i.e.,

$$\mathrm{H}_0^\alpha(\Omega) = \left\{ f \in \mathrm{H}^\alpha(\mathbb{R}^2) \mid \mathrm{supp}(f) \subset \overline{\Omega} \right\}.$$

For functions $g \equiv g(t, \theta)$ in polar coordinates $(t, \theta) \in \mathbb{R} \times [0, \pi)$, we define the Sobolev space $\mathrm{H}^\alpha(\mathbb{R} \times [0, \pi))$, for $\alpha \in \mathbb{R}$, as the space of all functions g with $g(\cdot, \theta) \in \mathrm{H}^\alpha(\mathbb{R})$ for all $\theta \in [0, \pi)$ and

$$\|g\|_{\mathrm{H}^\alpha(\mathbb{R} \times [0,\pi))} = \left(\int_0^\pi \int_{\mathbb{R}} (1 + S^2)^\alpha \, |\mathcal{F}g(S, \theta)|^2 \, \mathrm{d}S \, \mathrm{d}\theta \right)^{1/2} < \infty.$$

Then, we obtain the following continuity result for the Radon transform. Here, it is essential to assume that we deal with functions with compact support. The result is based on [82, Theorem II.5.1], but its assumptions are relaxed and the statement is adopted to our situation.

Theorem 2.2.28 (See [82, Theorem II.5.1]). *Let $\Omega \subset \mathbb{R}^2$ be an open and bounded set and let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}_0^\alpha(\Omega)$ with $\alpha \in \mathbb{R}$. Then, we have $\mathcal{R}f \in \mathrm{H}^{\alpha+1/2}(\mathbb{R} \times [0,\pi))$ and there exists a constant $C_\alpha > 1$ such that*

$$\|f\|_{\mathrm{H}^\alpha(\mathbb{R}^2)} \leq \|\mathcal{R}f\|_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} \leq C_\alpha \|f\|_{\mathrm{H}^\alpha(\mathbb{R}^2)}.$$

Proof. Since $f \in \mathrm{L}^1(\mathbb{R}^2)$, the Fourier slice Theorem 2.2.11 gives

$$\mathcal{F}(\mathcal{R}f)(S,\theta) = \mathcal{F}f(S\cos(\theta), S\sin(\theta)) \quad \forall\, (S,\theta) \in \mathbb{R} \times [0,\pi)$$

and we obtain

$$\|\mathcal{R}f\|^2_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} = \int_0^\pi \int_{\mathbb{R}} (1+S^2)^{\alpha+\frac{1}{2}} \, |\mathcal{F}f(S\cos(\theta), S\sin(\theta))|^2 \, \mathrm{d}S \, \mathrm{d}\theta.$$

Using the estimate

$$(1+S^2)^{\frac{1}{2}} \geq |S| \quad \forall\, S \in \mathbb{R},$$

this implies that

$$\|\mathcal{R}f\|^2_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} \geq \int_0^\pi \int_{\mathbb{R}} (1+S^2)^{\alpha} \, |\mathcal{F}f(S\cos(\theta), S\sin(\theta))|^2 \, |S| \, \mathrm{d}S \, \mathrm{d}\theta.$$

By applying the transformation

$$X = S\cos(\theta) \quad \text{and} \quad Y = S\sin(\theta),$$

we get $\mathrm{d}X\,\mathrm{d}Y = |S|\,\mathrm{d}S\,\mathrm{d}\theta$ and with $S^2 = X^2 + Y^2$ follows that

$$\|\mathcal{R}f\|^2_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} \geq \int_{\mathbb{R}} \int_{\mathbb{R}} (1+X^2+Y^2)^{\alpha} \, |\mathcal{F}f(X,Y)|^2 \, \mathrm{d}X \, \mathrm{d}Y = \|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^2)},$$

which proves the first inequality. On the other hand, we have

$$\begin{aligned} \|\mathcal{R}f\|^2_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} &= \int_0^\pi \int_{\mathbb{R}} (1+S^2)^{\alpha+\frac{1}{2}} \, |\mathcal{F}f(S\cos(\theta), S\sin(\theta))|^2 \, \mathrm{d}S \, \mathrm{d}\theta \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} (1+X^2+Y^2)^{\alpha+\frac{1}{2}} \, (X^2+Y^2)^{-\frac{1}{2}} \, |\mathcal{F}f(X,Y)|^2 \, \mathrm{d}X \, \mathrm{d}Y, \end{aligned}$$

where we used the transformation from above and

$$|S| = \sqrt{X^2+Y^2}.$$

We now split the above representation of the $\mathrm{H}^{\alpha+1/2}$-norm of $\mathcal{R}f$ into the sum of two integrals,

$$\|\mathcal{R}f\|^2_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} = I_1 + I_2,$$

where we let

$$I_1 = \int_{X^2+Y^2\leq 1} (1+X^2+Y^2)^{\alpha+\frac{1}{2}} \, (X^2+Y^2)^{-\frac{1}{2}} \, |\mathcal{F}f(X,Y)|^2 \, \mathrm{d}(X,Y)$$

and

$$I_2 = \int_{X^2+Y^2>1} (1+X^2+Y^2)^{\alpha+\frac{1}{2}} \, (X^2+Y^2)^{-\frac{1}{2}} \, |\mathcal{F}f(X,Y)|^2 \, \mathrm{d}(X,Y).$$

In what follows, we estimate the integrals I_1 and I_2 separately with respect to the H^α-norm of f. To bound the second integral I_2, we use the estimate

$$X^2+Y^2 \geq \frac{1}{2}\,(1+X^2+Y^2) \quad \forall\, X^2+Y^2 \geq 1.$$

With this, I_2 can be estimated from above by

$$\begin{aligned} I_2 &\leq \sqrt{2} \int_{X^2+Y^2>1} (1+X^2+Y^2)^{\alpha+\frac{1}{2}}\,(1+X^2+Y^2)^{-\frac{1}{2}}\,|\mathcal{F}f(X,Y)|^2\,\mathrm{d}(X,Y) \\ &= \sqrt{2} \int_{X^2+Y^2>1} (1+X^2+Y^2)^{\alpha}\,|\mathcal{F}f(X,Y)|^2\,\mathrm{d}(X,Y) \leq \sqrt{2}\|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^2)}. \end{aligned}$$

For the first integral I_1, we use the fact that $\mathcal{F}f \in \mathscr{C}_0(\mathbb{R}^2)$ for $f \in \mathrm{L}^1(\mathbb{R}^2)$ and get

$$\begin{aligned} I_1 &\leq \int_{X^2+Y^2\leq 1} (1+X^2+Y^2)^{\alpha+\frac{1}{2}}\,(X^2+Y^2)^{-\frac{1}{2}}\,\mathrm{d}(X,Y) \sup_{X^2+Y^2\leq 1} |\mathcal{F}f(X,Y)|^2 \\ &\leq C_1(\alpha) \sup_{X^2+Y^2\leq 1} |\mathcal{F}f(X,Y)|^2 \end{aligned}$$

for some constant $0 < C_1(\alpha) < \infty$. In order to estimate the supremum, we choose an even function $\chi \in \mathscr{C}_c^\infty(\mathbb{R}^2)$ which is 1 on Ω and, for fixed $(X,Y) \in \mathbb{R}^2$, we define the function

$$\chi_{(X,Y)}(x,y) = \mathrm{e}^{-\mathrm{i}(xX+yY)}\,\chi(x,y) \quad \text{ for } (x,y) \in \mathbb{R}^2.$$

Then, its inverse Fourier transform $\mathcal{F}^{-1}\chi_{(X,Y)}$ exists and is given by

$$\mathcal{F}^{-1}\chi_{(X,Y)}(x,y) = \mathcal{F}^{-1}\chi(x-X,y-Y) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Applying Parseval's identity, Theorem A.1.11, yields

$$\begin{aligned} |\mathcal{F}f(X,Y)|^2 &= \left| \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)\,\mathrm{e}^{-\mathrm{i}(xX+yY)}\,\mathrm{d}x\,\mathrm{d}y \right|^2 = \left| \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_{(X,Y)}(x,y)\,f(x,y)\,\mathrm{d}x\,\mathrm{d}y \right|^2 \\ &= \left| \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}^{-1}\chi_{(X,Y)}(x,y)\,\mathcal{F}f(x,y)\,\mathrm{d}x\,\mathrm{d}y \right|^2 \\ &= \left| \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}^{-1}\chi_{(X,Y)}(x,y)\,(1+x^2+y^2)^{-\frac{\alpha}{2}}\,(1+x^2+y^2)^{\frac{\alpha}{2}}\,\mathcal{F}f(x,y)\,\mathrm{d}x\,\mathrm{d}y \right|^2 \end{aligned}$$

and with Hölder's inequality follows that

$$|\mathcal{F}f(X,Y)|^2 \leq \left(\int_{\mathbb{R}} \int_{\mathbb{R}} (1+x^2+y^2)^{-\alpha}\,|\mathcal{F}^{-1}\chi_{(X,Y)}(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y \right) \|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^2)},$$

since $\chi \in \mathscr{C}_c^\infty(\mathbb{R}^2)$ and $f \in \mathrm{H}_0^\alpha(\Omega)$, i.e., $(1+x^2+y^2)^{\frac{\alpha}{2}}\,|\mathcal{F}f| \in \mathrm{L}^2(\mathbb{R}^2)$. We further have

$$\mathcal{F}^{-1}\chi_{(X,Y)}(x,y) = \frac{1}{4\pi^2}\,\mathcal{F}\chi_{(X,Y)}(-x,-y) \quad \forall\,(x,y) \in \mathbb{R}^2,$$

which implies that

$$|\mathcal{F}f(X,Y)|^2 \leq \frac{1}{16\pi^4}\,\|\chi_{(X,Y)}\|^2_{\mathrm{H}^{-\alpha}(\mathbb{R}^2)}\,\|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^2)} \quad \forall\,(X,Y) \in \mathbb{R}^2.$$

The $\mathrm{H}^{-\alpha}$-norm of $\chi_{(X,Y)}$ is a continuous function of (X,Y) and, consequently, there exists a constant $C_2(\alpha) > 0$ such that

$$\sup_{X^2+Y^2\leq 1} \frac{1}{16\pi^4}\,\|\chi_{(X,Y)}\|^2_{\mathrm{H}^{-\alpha}(\mathbb{R}^2)} \leq C_2(\alpha).$$

Combining the estimates yields

$$\|\mathcal{R}f\|^2_{\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))} \leq \left(\sqrt{2} + C_1(\alpha)\,C_2(\alpha)\right) \|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^2)} = C_\alpha^2\,\|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^2)}$$

and the second inequality is also satisfied. In particular, we have $\mathcal{R}f \in \mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))$. □

Remark 2.2.29 (See [67, Chapter III.B]). *Theorem 2.2.28 shows that the Radon transform $\mathcal{R}$ admits an inverse and that this inverse is continuous as an operator from $\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))$ into $\mathrm{H}_0^\alpha(\Omega)$. Here, the order of the latter Sobolev space is optimal in the sense that the result is wrong for any larger index. In particular, if α is chosen such that $\mathrm{H}^{\alpha+1/2}(\mathbb{R}\times[0,\pi))$ is simply the L^2-space, i.e., $\alpha=-\frac{1}{2}$, then $\mathrm{H}_0^\alpha(\Omega)$ is a Sobolev space of negative order, whose norm is weaker than the L^2-norm on $\mathrm{L}^2(\Omega)$. This shows that $\mathcal{R}^{-1}$ is not continuous in an L^2-setting, i.e., the CT reconstruction Problem 2.2.4 is* ill-posed *in the sense of Hadamard.*

We remark that choosing $\alpha=0$ in Theorem 2.2.28 shows that the problem of reconstructing a bivariate function from its Radon data is ill-posed of degree $\frac{1}{2}$ in the sense of Definition 2.2.27. In particular, the Radon transform $\mathcal{R}$ smoothes by an order of $\frac{1}{2}$ in the Sobolev scale. Since the inversion has to reverse the smoothing, the reconstruction process is unstable and regularization strategies have to be applied to stabilize the inversion.

For this purpose, we follow a standard approach in Chapter 3 and replace the filter $|S|$ in the filtered back projection formula (2.9) by a so called *low-pass filter* $A_L(S)$ of finite bandwidth L and with a compactly supported window function. For other approaches we refer to the standard literature on inverse problems and regularization, see, for example, [29], [53], [59], [99].

Reformulation of the filtered back projection formula

We close this section on the mathematics of computerized tomography by reformulating the filtered back projection formula (2.9) in standard forms, that can be found in the literature, and by stating Radon's original formulation of the inversion formula. In the following, we proceed as in [82] and, for convenience, restrict ourselves to the special case of smooth and rapidly decreasing functions, i.e., we assume that the target function f lies in the Schwartz space $\mathcal{S}(\mathbb{R}^2)$.

We start with the definition of Riesz potentials. To this end, let $\alpha\in\mathbb{R}$ satisfy $\alpha<2$. Then, the *Riesz potential* $I^\alpha f$ of a function $f\in\mathcal{S}(\mathbb{R}^2)$ is defined by means of the Fourier transform $\mathcal{F}$ as

$$I^\alpha f=\mathcal{F}^{-1}(\|\cdot\|_{\mathbb{R}^2}^{-\alpha}\,\mathcal{F}f).$$

Note that, for $\alpha<2$, we have $\mathcal{F}(I^\alpha f)\in\mathrm{L}^1(\mathbb{R}^2)$ for all $f\in\mathcal{S}(\mathbb{R}^2)$ and, thus, the Riesz potential I^α is well-defined on $\mathcal{S}(\mathbb{R}^2)$. In addition, for $|\alpha|<2$, we obtain

$$I^{-\alpha}I^\alpha f=f\quad\forall\, f\in\mathcal{S}(\mathbb{R}^2).$$

Note further that the definition of the Riesz potential I^α can be extended to a larger class of functions and, if I^α is applied to a function $g\equiv g(t,\theta)$ in polar coordinates $(t,\theta)\in\mathbb{R}\times[0,\pi)$, it is defined to act only on the radial variable t, i.e.,

$$I^\alpha g(\cdot,\theta)=\mathcal{F}^{-1}(|\cdot|^{-\alpha}\,\mathcal{F}g(\cdot,\theta))\quad\text{ for }\theta\in[0,\pi).$$

With the definition of the Riesz potential I^α we can now restate the FBP formula (2.9) in standard form as

$$f=\frac{1}{2}\,\mathcal{B}(I^{-1}\mathcal{R}f)\quad\forall\, f\in\mathcal{S}(\mathbb{R}^2),$$

see also [82, Theorem II.2.1]. For more details on Riesz potentials, we refer to [118, Chapter V].

For univariate functions $g\in\mathcal{S}(\mathbb{R})$, the definition of the Riesz potential and the differentiation property of the Fourier transform yield

$$\mathcal{F}(I^{-1}g)(\omega)=|\omega|\,\mathcal{F}g(\omega)=\operatorname{sgn}(\omega)\,\omega\,\mathcal{F}g(\omega)=-\mathrm{i}\,\operatorname{sgn}(\omega)\,\mathcal{F}g'(\omega)\quad\forall\,\omega\in\mathbb{R}.$$

Thus, we obtain the representation

$$I^{-1}g=Hg'\quad\forall\, g\in\mathcal{S}(\mathbb{R}),$$

where Hg denotes the *Hilbert transform* of $g \in \mathcal{S}(\mathbb{R})$, which can be defined by means of the Fourier transform as

$$Hg = \mathcal{F}^{-1}(-\mathrm{i}\ \mathrm{sgn}\ \mathcal{F}g).$$

With this, we found another reformulation of the filtered back projection formula (2.9) as

$$f = \frac{1}{2}\,\mathcal{B}(H[(\mathcal{R}f)']) \quad \forall\, f \in \mathcal{S}(\mathbb{R}^2),$$

where the derivative and the Hilbert transform are taken with respect to the radial variable. For a comprehensive treatment of the Hilbert transform we refer the reader to [120, Chapter 5].

Finally, for $f \in \mathcal{S}(\mathbb{R}^2)$ and fixed $(x, y) \in \mathbb{R}^2$, we define the function

$$F_{(x,y)}(q) = \frac{1}{2\pi}\int_0^{2\pi} \mathcal{R}f(x\cos(\theta) + y\sin(\theta) + q, \theta)\,\mathrm{d}\theta \quad \text{for } q \in \mathbb{R}.$$

Then, the FBP formula (2.9) can be written as

$$f(x, y) = -\frac{1}{\pi}\int_0^{\infty} \frac{1}{q}\,F'_{(x,y)}(q)\,\mathrm{d}q \quad \forall\,(x, y) \in \mathbb{R}^2, \tag{2.10}$$

which is the classical inversion formula that was published by Johann Radon in 1917, see [94]. An English translation of the original article can be found in [95] and [26, Appendix A]. Moreover, in [73] it is shown that, for functions $f \in \mathrm{L}^p(\mathbb{R}^2)$ with $\frac{4}{3} < p < 2$, the Radon transform $\mathcal{R}f$ exists almost everywhere on $\mathbb{R}^2$ and Radon's inversion formula (2.10) holds for almost all $(x, y) \in \mathbb{R}^2$.

2.3 Reconstruction techniques

We close this chapter on computerized tomography by giving a brief overview of the wide range of reconstruction methods that exist in the literature, cf. [45], [82], [83].

There are basically two major classes of reconstruction techniques, namely *analytic* (or direct) and *algebraic* methods. As an example for the analytic approach, we present the well-known *method of filtered back projection* (FBP), which is based on the FBP formula for the inversion of the Radon transform. On the other hand, one of the most popular algebraic methods is given by the so called *algebraic reconstruction technique* (ART), which is based on a fully discrete formulation of the CT reconstruction problem.

The analysis of the FBP reconstruction technique is the principle goal of this thesis. Thus, this method is thoroughly introduced and discussed in great detail in Chapter 3 and, as the main part of this work, the inherent FBP reconstruction error is investigated in Chapter 4.

Before we start with explaining the different reconstruction techniques, we remark that in practical situations the Radon data $g = \mathcal{R}f$ is usually not known precisely, but only up to an error bound $\delta > 0$. Thus, in the following we assume that we are given some data g^δ with

$$\|g - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \le \delta.$$

The goal is to compute an approximation f^δ to the target function f from the noisy data g^δ. Hence, the practical CT reconstruction problem consists in solving the linear operator equation

$$\mathcal{R}f^\delta = g^\delta. \tag{2.11}$$

2.3.1 Analytic reconstruction methods

The main characteristic of the class of analytic reconstruction methods is the fact that these techniques are based on analytic inversion formulas for the Radon transform and, hence, deal with continuous data.

The method of *filtered back projection* (FBP) is the most important analytic reconstruction technique in computerized tomography and is still used in commercial CT scanners, see [85]. It is based on the classical filtered back projection formula (2.9), which yields an analytical inversion of the Radon transform. However, as we have seen in Observation 2.2.24, the FBP formula is numerically unstable due to its sensitivity to noise and the ill-posedness of the CT reconstruction Problem 2.2.4, see Remark 2.2.29. To stabilize the inversion, the filter $|S|$ in (2.9) is truncated and smoothened by the application of a low-pass filter A_L of finite bandwidth $L > 0$ satisfying

$$A_L(S) \xrightarrow{L\to\infty} |S| \quad \forall\, S \in \mathbb{R}.$$

We show in Chapter 3 that the resulting approximate reconstruction formula can be written as

$$f \approx \frac{1}{2}\, \mathcal{B}(\mathcal{F}^{-1} A_L * \mathcal{R}f).$$

Applying this reconstruction formula to the noisy measurements g^δ yields the regularized solution

$$f^\delta = \frac{1}{2}\, \mathcal{B}(\mathcal{F}^{-1} A_L * g^\delta)$$

and the bandwidth L is usually chosen as a function of the noise level δ such that

$$f^\delta \longrightarrow f \quad \text{for} \quad \delta \longrightarrow 0.$$

There are many filters that can be used in this context. Classical examples are given by the *Ram-Lak filter*, the *Shepp-Logan filter* and the *Cosine filter*, see [83]. We will explain the method of filtered back projection in great detail in Chapter 3. In the main part of this thesis, Chapter 4, we then analyse the inherent FBP reconstruction error which is incurred by the application of the low-pass filter A_L and derive convergence rates as the bandwidth L goes to infinity.

We now list some more analytic reconstruction techniques without further explanation. The *summability method* is an approximation theoretic approach for reconstructing a function from Radon data and was introduced in [71]. This technique is related to the *method of approximate inverse*, which was developed in [65] as a general framework for linear operator equations and can also be applied to the tomographic reconstruction problem, see, for instance, [68], [102], [111]. Finally, we want to mention the *Fourier reconstruction technique*, which is based on a direct application of the Fourier slice theorem, see Theorem 2.2.11 and the subsequent paragraph.

2.3.2 Algebraic reconstruction methods

The second large class of tomographic reconstruction techniques is given by the class of algebraic reconstruction methods. In contrast to the analytic approach, these techniques are not based on analytical inversion formulas for the Radon transform and, further, deal with fully discrete data. Note that the class of algebraic methods splits into many different subclasses and, in the following, we only present the basic idea of these methods.

Let us start with the fully discrete formulation of the CT reconstruction problem. In this setting, the sought function f^δ is discretized beforehand to solving the noisy reconstruction problem (2.11). To this end, a set of basis function $\{\phi_j\}_{j=1}^N$ is fixed and the function f^δ is assumed to be given by a linear combination of these basis functions, i.e.,

$$f^\delta = \sum_{j=1}^{N} c_j\, \phi_j$$

for some coefficient vector $c = (c_1, \dots, c_N)^T \in \mathbb{R}^N$. Further, we assume that we deal with a finite number of Radon data

$$y = (\mathcal{R}f(t_1, \theta_1), \dots, \mathcal{R}f(t_M, \theta_M))^T \in \mathbb{R}^M.$$

Then, the fully discrete version of the noisy reconstruction problem (2.11) is given by the linear system of equations

$$R \cdot c = y^\delta, \tag{2.12}$$

where $y^\delta \in \mathbb{R}^M$ is the noisy measurement vector and $R \in \mathbb{R}^{M \times N}$ is the system matrix with entries

$$r_{k,j} = \mathcal{R}\phi_j(t_k, \theta_k) \quad \text{for } k = 1, \ldots, M,\ j = 1, \ldots, N.$$

There are many algorithms for solving the linear system (2.12), which is usually highly underdetermined. One of the most applied methods is the so called *algebraic reconstruction technique* (ART), which is an implementation of the classical Kaczmarz method, or one of its variants SART (simultaneous algebraic reconstruction technique) and SIRT (simultaneous iterative reconstruction technique), see [45], [51], [83] for details. However, for solving (2.12) one still has to consider the ill-posedness of the reconstruction problem and apply regularization strategies, which are not yet incorporated in the standard formulation of ART.

A typical regularization strategy is to incorporate prior knowledge into the reconstruction procedure. This can be done by applying the general framework of *variational regularization*. In this setting, a regularized solution of the linear system (2.12) is computed by minimizing a Tikhonov type functional

$$c_\gamma = \operatorname*{arg\,min}_{c \in \mathbb{R}^N} \left\| R \cdot c - y^\delta \right\|_{\mathbb{R}^M}^2 + \gamma \cdot \Lambda(c),$$

where $\gamma > 0$ is a regularization parameter and $\Lambda : \mathbb{R}^N \longrightarrow \mathbb{R}_{\geq 0}$ is the *prior function*. The first term denotes the *data fidelity term* and controls the data error, whereas the second term acts as a *penalty term* and encodes the prior knowledge about the solution.

We remark that there are many possible choices for the prior term, depending on the particular features of the unknown object that shall be preserved or emphasized. One prominent example is the total variation (TV) seminorm, which can be used for edge-preserving reconstructions, cf. [40]. For more details on variational regularization, we refer to [109].

Another regularization strategy is pursued in the *statistical reconstruction* approach. Here, the unknown coefficient vector c and the measurement vector y^δ are considered to be realizations of certain random variables. For the statistical formulation of the reconstruction problem one has to model the likelihood density $p(y|c)$, which encodes the measurement process, and the prior density $p(c)$, which encodes the prior information about the unknown object. The solution of the statistical reconstruction problem then consists in determining the posteriori density $p(c|y)$, where a statistical estimator is used to compute an explicit solution to (2.12). For more details, we refer to [13], [17], [50], [54], [114].

Inspired by multivariate approximation theory, a kernel-based method for the reconstruction from scattered Radon data is proposed in [24], [25]. Here, a generalized Hermite-Birkhoff interpolation scheme with (weighted) positive definite kernel functions is applied to reconstruct the target function f from finitely many Radon data along scattered lines. To this end, f is expanded in a data-dependent basis, which is induced by the chosen kernel function, leading to a linear system of equations for the coefficients of f with a symmetric positive definite matrix.

Finally, we remark that there are also approaches that combine the ideas from analytical and algebraic methods, see, for example, [5], [87], [89] and also [31], [84] for the special case of digital X-ray tomosynthesis.

To close this section on reconstruction techniques we want to mention that there exist various other methods in the literature and describing them all is far beyond the scopes of this thesis. To name just a few, in [23] a FBP like approach is considered with angle-dependent filter functions and, similarly, in [52], [58] the concept of ridge functions is applied. The approach in [126] is based on an expansion of the target function in orthogonal polynomials leading to a direct approximation formula and, finally, in [48], [86] (convolutional) neural networks are used for the image reconstruction from Radon data.

Chapter 3

Method of filtered back projection

In Chapter 2 we have seen that the basic reconstruction problem in computerized tomography can be formulated as the problem of reconstructing a bivariate function $f \equiv f(x,y) \in \mathrm{L}^1(\mathbb{R}^2)$ from given Radon data

$$\{\mathcal{R}f(t,\theta) \mid t \in \mathbb{R},\, \theta \in [0,\pi)\}\,.$$

Based on the filtered back projection (FBP) formula (2.9) and under suitable assumptions on f we have found the reconstruction formula

$$f(x,y) = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[|S|\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \quad \forall\,(x,y) \in \mathbb{R}^2,$$

which is highly sensitive with respect to noise and, thus, cannot be used in practice.

In this chapter we explain how the above FBP formula can be stabilized by incorporating a low-pass filter $A_L : \mathbb{R} \longrightarrow \mathbb{R}$ of the form

$$A_L(S) = |S|\, W(S/L) \quad \text{for } S \in \mathbb{R}$$

with bandwidth $L > 0$ and an even window $W \in \mathrm{L}^\infty(\mathbb{R})$ of compact support $\mathrm{supp}(W) \subseteq [-1,1]$. This standard approach reduces the noise sensitivity of the reconstruction scheme, but leads to an inexact FBP reconstruction, which we denote by f_L.

In the following, we rigorously define the approximate FBP reconstruction f_L and analyse its properties. In particular, we show that f_L can be expressed in standard form as

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f) = f * K_L \tag{3.1}$$

in the L^2-sense with the convolution kernel

$$K_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L).$$

Further, we analyse the properties of the kernel K_L, which are entirely determined by the choice of the window function W and the bandwidth L.

We remark that the approach taken in this work is standard in the mathematics of computerized tomography and some of the upcoming results are well-known from the literature, especially the representation (3.1) of the approximate FBP reconstruction f_L, see, e.g., [33], [82]. However, in this chapter we provide rigorous derivations of the statements and complete proofs. In particular, we adjust the assumptions of the results to our setting, where we deal with target functions f that are only required to satisfy $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and with convolution kernels K_L that are not necessarily integrable on $\mathbb{R}^2$. In the literature, however, it is usually assumed that f lies in the Schwartz space $\mathcal{S}(\mathbb{R}^2)$, as, for instance, in [82], [83], or that the kernel K_L is integrable, such as in [71], [118]. Furthermore, we also include additional statements and we complete existing proofs.

Finally, we wish to remark that some parts of the presented material are contained in the preprint [11].

3.1 Preliminaries

In this section we first show some preliminary propositions we need for the derivation of the approximate FBP reconstruction formula (3.1). We start with the following variant of the Fourier convolution Theorem A.1.19 where one function is assumed to be integrable on $\mathbb{R}^n$, but the other one is assumed to be square integrable on $\mathbb{R}^n$. The result can be found, for instance, in [119], but without a proof.

Proposition 3.1.1 (Fourier convolution theorem, see [119, Theorem I.2.6])**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $g \in \mathrm{L}^2(\mathbb{R}^n)$. Then,*

$$\mathcal{F}(f * g) = \mathcal{F}f \cdot \mathcal{F}g \tag{3.2}$$

holds in L^2-sense and, in particular, almost everywhere on $\mathbb{R}^n$.

For the sake of completeness, we give a proof of the above Fourier convolution theorem.

Proof. For $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $g \in \mathrm{L}^2(\mathbb{R}^n)$ we have

- $f * g \in \mathrm{L}^2(\mathbb{R}^n)$ by Young's inequality, Theorem A.1.16, with

$$\|f * g\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^n)} \, \|g\|_{\mathrm{L}^2(\mathbb{R}^n)},$$

- $\mathcal{F}(f * g) \in \mathrm{L}^2(\mathbb{R}^n)$ by the Rayleigh-Plancherel Theorem A.1.12 with

$$\|\mathcal{F}(f * g)\|_{\mathrm{L}^2(\mathbb{R}^n)} = (2\pi)^{n/2} \, \|f * g\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq (2\pi)^{n/2} \, \|f\|_{\mathrm{L}^1(\mathbb{R}^n)} \, \|g\|_{\mathrm{L}^2(\mathbb{R}^n)},$$

- $\mathcal{F}f \in \mathscr{C}_0(\mathbb{R}^n)$ by the Riemann-Lebesgue Lemma A.1.2 with

$$\|\mathcal{F}f\|_\infty \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^n)},$$

- $\mathcal{F}f \cdot \mathcal{F}g \in \mathrm{L}^2(\mathbb{R}^n)$ by the generalized Hölder inequality, Theorem A.4.1, with

$$\|\mathcal{F}f \cdot \mathcal{F}g\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq \|\mathcal{F}f\|_\infty \, \|\mathcal{F}g\|_{\mathrm{L}^2(\mathbb{R}^n)} \leq (2\pi)^{n/2} \, \|f\|_{\mathrm{L}^1(\mathbb{R}^n)} \, \|g\|_{\mathrm{L}^2(\mathbb{R}^n)}.$$

Thus, both expressions in (3.2) are well-defined in $\mathrm{L}^2(\mathbb{R}^n)$. To show the asserted identity, we approximate the function $g \in \mathrm{L}^2(\mathbb{R}^n)$ by a sequence of integrable functions $g_k \in \mathrm{L}^1(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$ with

$$\|g_k - g\|_{\mathrm{L}^2(\mathbb{R}^n)} \longrightarrow 0 \quad \text{for} \quad k \longrightarrow \infty,$$

since $\mathrm{L}^1(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$ is dense in $\mathrm{L}^2(\mathbb{R}^n)$. Then, the Rayleigh-Plancherel theorem also yields

$$\|\mathcal{F}g_k - \mathcal{F}g\|_{\mathrm{L}^2(\mathbb{R}^n)} = (2\pi)^{n/2} \, \|g_k - g\|_{\mathrm{L}^2(\mathbb{R}^n)} \longrightarrow 0 \quad \text{for} \quad k \longrightarrow \infty.$$

We now show that

$$\mathcal{F}(f * g_k) = \mathcal{F}f \cdot \mathcal{F}g_k.$$

By Young's inequality we have $f * g_k \in \mathrm{L}^1(\mathbb{R}^n)$ and, consequently, we get $\mathcal{F}(f * g_k) \in \mathscr{C}_0(\mathbb{R}^n)$ due to the Riemann-Lebesgue lemma. By applying Fubini's theorem, for all $\xi \in \mathbb{R}^n$ follows that

$$\begin{aligned}
\mathcal{F}(f * g_k)(\xi) &= \int_{\mathbb{R}^n} (f * g_k)(x) \, \mathrm{e}^{-\mathrm{i}\, x^T \xi} \, \mathrm{d}x = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(y) \, g_k(x - y) \, \mathrm{d}y \right) \mathrm{e}^{-\mathrm{i}\, x^T \xi} \, \mathrm{d}x \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) \, \mathrm{e}^{-\mathrm{i}\, y^T \xi} \, g_k(x - y) \, \mathrm{e}^{-\mathrm{i}\,(x-y)^T \xi} \, \mathrm{d}y \, \mathrm{d}x \\
&= \int_{\mathbb{R}^n} f(y) \, \mathrm{e}^{-\mathrm{i} y^T \xi} \left(\int_{\mathbb{R}^n} g_k(x - y) \, \mathrm{e}^{-\mathrm{i}(x-y)^T \xi} \, \mathrm{d}x \right) \mathrm{d}y.
\end{aligned}$$

Hence, using the transformation $z = x - y$ in the inner integral yields

$$\mathcal{F}(f * g_k)(\xi) = \int_{\mathbb{R}^n} f(y)\, \mathrm{e}^{-\mathrm{i}y^T\xi} \left(\int_{\mathbb{R}^n} g_k(z)\, \mathrm{e}^{-\mathrm{i}z^T\xi}\, \mathrm{d}z \right) \mathrm{d}y = (\mathcal{F}f \cdot \mathcal{F}g_k)(\xi) \quad \forall\, \xi \in \mathbb{R}^n,$$

i.e.,

$$\mathcal{F}(f * g_k) = \mathcal{F}f \cdot \mathcal{F}g_k \quad \forall\, k \in \mathbb{N}.$$

With this we are now prepared to show the asserted identity for $g \in \mathrm{L}^2(\mathbb{R}^n)$. We have

$$\begin{aligned}
\|\mathcal{F}(f * g) - \mathcal{F}f \cdot \mathcal{F}g\|_{\mathrm{L}^2(\mathbb{R}^n)} &\leq \|\mathcal{F}(f * g) - \mathcal{F}(f * g_k)\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|\mathcal{F}(f * g_k) - \mathcal{F}f \cdot \mathcal{F}g\|_{\mathrm{L}^2(\mathbb{R}^n)} \\
&= \|\mathcal{F}(f * (g - g_k))\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|\mathcal{F}f \cdot (\mathcal{F}g_k - \mathcal{F}g)\|_{\mathrm{L}^2(\mathbb{R}^n)} \\
&\leq (2\pi)^{n/2}\, \|f\|_{\mathrm{L}^1(\mathbb{R}^n)}\, \|g - g_k\|_{\mathrm{L}^2(\mathbb{R}^n)} + \|f\|_{\mathrm{L}^1(\mathbb{R}^n)}\, \|\mathcal{F}g - \mathcal{F}g_k\|_{\mathrm{L}^2(\mathbb{R}^n)} \\
&\xrightarrow{k\to\infty} 0.
\end{aligned}$$

Consequently,

$$\mathcal{F}(f * g) = \mathcal{F}f \cdot \mathcal{F}g$$

holds in L^2-sense and, in particular, almost everywhere on $\mathbb{R}^n$. □

We continue with some additional properties of the Radon transform $\mathcal{R}$. In Proposition 2.2.7 we have seen that $\mathcal{R}$ is a continuous operator from $\mathrm{L}^1(\mathbb{R}^2)$ to $\mathrm{L}^1(\mathbb{R} \times [0, \pi))$ with

$$\|\mathcal{R}f\|_{\mathrm{L}^1(\mathbb{R}\times[0,\pi))} \leq \pi\, \|f\|_{\mathrm{L}^1(\mathbb{R}^2)} \quad \forall\, f \in \mathrm{L}^1(\mathbb{R}^2).$$

In the next proposition we analyse the L^2-norm of the Radon transform $\mathcal{R}f$. To this end, we have to assume that the function f is in the space $\mathrm{L}^2_c(\mathbb{R}^2)$ of square integrable functions on $\mathbb{R}^2$ with compact support, i.e.,

$$\mathrm{L}^2_c(\mathbb{R}^2) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^2) \mid \exists\, R > 0 :\ \mathrm{supp}(f) \subseteq B_R(0) \right\}.$$

Proposition 3.1.2 (L^2-norm of the Radon transform)**.** *Let $f \in \mathrm{L}^2_c(\mathbb{R}^2)$ have support in a compact set $K \subset \mathbb{R}^2$ with diameter*

$$c_f = \mathrm{diam}(K) = \sup\{\|(x - X, y - Y)\|_{\mathbb{R}^2} \mid (x, y), (X, Y) \in K\} < \infty.$$

Then, we have $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$ with

$$\|\mathcal{R}f\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \leq \pi\, c_f\, \|f\|^2_{\mathrm{L}^2(\mathbb{R}^2)}.$$

Proof. Let $f \in \mathrm{L}^2_c(\mathbb{R}^2)$ be compactly supported with $\mathrm{supp}(f) \subseteq K$ and let $(t, \theta) \in \mathbb{R} \times [0, \pi)$ be fixed. Then, the definition of the Radon transform gives

$$\begin{aligned}
|\mathcal{R}f(t,\theta)|^2 &= \left| \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s \right|^2 \\
&= \left| \int_{\mathbb{R}} (\chi_K \cdot f)(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s \right|^2
\end{aligned}$$

with the characteristic function χ_K of K. For the sake of brevity, we set

$$f_{(t,\theta)}(s) = f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \quad \text{for } s \in \mathbb{R}.$$

Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned}
|\mathcal{R}f(t,\theta)|^2 &\leq \int_{\mathbb{R}} \chi_K^2(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}s \cdot \int_{\mathbb{R}} |f_{(t,\theta)}(s)|^2\, \mathrm{d}s \\
&= \int_{\ell_{t,\theta}} \chi_K(x, y)\, \mathrm{d}(x, y) \cdot \int_{\mathbb{R}} |f_{(t,\theta)}(s)|^2\, \mathrm{d}s.
\end{aligned}$$

Since K has bounded diameter, it follows that

$$|\mathcal{R}f(t,\theta)|^2 \leq \operatorname{diam}(K) \int_{\mathbb{R}} |f_{(t,\theta)}(s)|^2 \,\mathrm{d}s = c_f \int_{\mathbb{R}} |f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))|^2 \,\mathrm{d}s.$$

Thus,

$$\begin{aligned}\|\mathcal{R}f\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} &= \int_0^\pi \int_{\mathbb{R}} |\mathcal{R}f(t,\theta)|^2 \,\mathrm{d}t\,\mathrm{d}\theta \\ &\leq c_f \int_0^\pi \int_{\mathbb{R}} \int_{\mathbb{R}} |f(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))|^2 \,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}\theta.\end{aligned}$$

Applying the transformation

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta)$$

yields $\mathrm{d}x\,\mathrm{d}y = \mathrm{d}t\,\mathrm{d}s$ and, consequently,

$$\|\mathcal{R}f\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \leq c_f \int_0^\pi \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\theta = \pi\, c_f \,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^2)},$$

as stated. □

Proposition 3.1.2 indicates that the Radon transform $\mathcal{R}$ can be seen as a densely defined unbounded linear operator from $\mathrm{L}^2(\mathbb{R}^2)$ to $\mathrm{L}^2(\mathbb{R}\times[0,\pi))$ with domain $\mathrm{L}^2_c(\mathbb{R}^2) \subset \mathrm{L}^2(\mathbb{R}^2)$. In fact, there are functions $f \in \mathrm{L}^2(\mathbb{R}^2)$ for which $\mathcal{R}f(t,\theta)$ does not exists for any $t \in \mathbb{R}$ and $\theta \in [0,\pi)$. In [37], an example of such a function is given by

$$f(x,y) = \big(2 + \|(x,y)\|_{\mathbb{R}^2}\big)^{-1} \Big(\log\big(2 + \|(x,y)\|_{\mathbb{R}^2}\big)\Big)^{-1} \quad \text{for } (x,y) \in \mathbb{R}^2.$$

To restore the continuity of $\mathcal{R}$, [36, Theorem 2.9] suggests to consider, for $\gamma > 0$, the weighted L^2-spaces

$$\mathrm{L}^2(\mathbb{R}^2, (1+x^2+y^2)^\gamma) = \Big\{ f \in \mathrm{L}^2(\mathbb{R}^2) \mid \|f\|_{\mathrm{L}^2(\mathbb{R}^2,(1+x^2+y^2)^\gamma)} < \infty \Big\}$$

with norm

$$\|f\|_{\mathrm{L}^2(\mathbb{R}^2,(1+x^2+y^2)^\gamma)} = \left(\int_{\mathbb{R}} \int_{\mathbb{R}} (1+x^2+y^2)^\gamma \,|f(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y \right)^{1/2}$$

and

$$\mathrm{L}^2(\mathbb{R}\times[0,\pi), (1+t^2)^\gamma) = \Big\{ g \in \mathrm{L}^2(\mathbb{R}\times[0,\pi)) \mid \|g\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi),(1+t^2)^\gamma)} < \infty \Big\}$$

with norm

$$\|g\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi),(1+t^2)^\gamma)} = \left(\int_0^\pi \int_{\mathbb{R}} (1+t^2)^\gamma \,|g(t,\theta)|^2 \,\mathrm{d}t\,\mathrm{d}\theta \right)^{1/2}.$$

Proposition 3.1.3. *The Radon transform $\mathcal{R}$ is a continuous operator from $\mathrm{L}^2(\mathbb{R}^2, (1+x^2+y^2)^\alpha)$ to $\mathrm{L}^2(\mathbb{R}\times[0,\pi), (1+t^2)^{\alpha-1/2})$ for $\alpha > \frac{1}{2}$. In particular, for $f \in \mathrm{L}^2(\mathbb{R}^2, (1+x^2+y^2)^\alpha)$ we have*

$$\|\mathcal{R}f\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi),(1+t^2)^{\alpha-1/2})} \leq \pi^{3/4} \left(\frac{\Gamma(\alpha - 1/2)}{\Gamma(\alpha)} \right)^{1/2} \|f\|_{\mathrm{L}^2(\mathbb{R}^2,(1+x^2+y^2)^\alpha)},$$

where Γ denotes the gamma function.

Proof. Let $f \in \mathrm{L}^2(\mathbb{R}^2, (1+x^2+y^2)^\alpha)$ with $\alpha > \frac{1}{2}$. For the sake of brevity, we set

$$r_\alpha(x,y) = (1+x^2+y^2)^{\alpha/2} \quad \text{for } (x,y) \in \mathbb{R}^2$$

and, for fixed $(t,\theta) \in \mathbb{R}\times[0,\pi)$,

$$x(s) = t\cos(\theta) - s\sin(\theta), \quad y(s) = t\sin(\theta) + s\cos(\theta) \quad \text{for } s \in \mathbb{R}.$$

Then, the definition of the Radon transform $\mathcal{R}$ gives

$$|\mathcal{R}f(t,\theta)|^2 = \left|\int_{\mathbb{R}} f(x(s),y(s))\,\mathrm{d}s\right|^2 = \left|\int_{\mathbb{R}} \frac{r_\alpha(x(s),y(s))}{r_\alpha(x(s),y(s))} f(x(s),y(s))\,\mathrm{d}s\right|^2$$

and, by applying the Cauchy-Schwarz inequality, we obtain

$$|\mathcal{R}f(t,\theta)|^2 \le \int_{\mathbb{R}} \frac{1}{r_\alpha^2(x(s),y(s))}\,\mathrm{d}s \cdot \int_{\mathbb{R}} r_\alpha^2(x(s),y(s))\,|f(x(s),y(s))|^2\,\mathrm{d}s.$$

Using (2.5), we have

$$r_\alpha^2(x(s),y(s)) = (1+t^2+s^2)^\alpha \quad \forall\, s\in\mathbb{R}$$

and the first integral can be calculated as

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{r_\alpha^2(x(s),y(s))}\,\mathrm{d}s &= 2\int_0^\infty (1+t^2+s^2)^{-\alpha}\,\mathrm{d}s = 2(1+t^2)^{-\alpha}\int_0^\infty \left(1+\frac{s^2}{1+t^2}\right)^{-\alpha}\mathrm{d}s\\ &= (1+t^2)^{-\alpha+1/2}\int_0^\infty \sigma^{-1/2}\,(1+\sigma)^{-\alpha}\,\mathrm{d}\sigma,\end{aligned}$$

where we applied the transformation $\sigma = \frac{s^2}{1+t^2}$, which gives $\mathrm{d}s = \frac{1}{2}(1+t^2)^{1/2}\,\sigma^{-1/2}\,\mathrm{d}\sigma$. Note that, for $\alpha > \frac{1}{2}$, we have

$$\int_0^\infty \sigma^{-1/2}\,(1+\sigma)^{-\alpha}\,\mathrm{d}\sigma < \infty.$$

Indeed, substituting $\sigma = \frac{u}{1-u}$ yields $\mathrm{d}\sigma = (1-u)^{-2}\,\mathrm{d}u$ and $\sigma^{-1/2}\,(1+\sigma)^{-\alpha} = u^{-1/2}\,(1-u)^{\alpha+1/2}$, which implies that

$$\begin{aligned}\int_0^\infty \sigma^{-1/2}\,(1+\sigma)^{-\alpha}\,\mathrm{d}\sigma &= \int_0^1 u^{-1/2}\,(1-u)^{\alpha+1/2}\,(1-u)^{-2}\,\mathrm{d}u = \int_0^1 u^{1/2-1}\,(1-u)^{\alpha-1/2-1}\,\mathrm{d}u\\ &= B(1/2,\alpha-1/2) = \frac{\Gamma(1/2)\,\Gamma(\alpha-1/2)}{\Gamma(\alpha)} = \pi^{1/2}\,\frac{\Gamma(\alpha-1/2)}{\Gamma(\alpha)},\end{aligned}$$

where B denotes the beta function and we used its close relation to the gamma function Γ. Hence, for the modulus of $\mathcal{R}f(t,\theta)$ follows that

$$(1+t^2)^{\alpha-1/2}\,|\mathcal{R}f(t,\theta)|^2 \le \pi^{1/2}\,\frac{\Gamma(\alpha-1/2)}{\Gamma(\alpha)}\int_{\mathbb{R}} r_\alpha^2(x(s),y(s))\,|f(x(s),y(s))|^2\,\mathrm{d}s.$$

Therefore, we get the estimate

$$\begin{aligned}\|\mathcal{R}f\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi),(1+t^2)^{\alpha-1/2})} &= \int_0^\pi\int_{\mathbb{R}} (1+t^2)^{\alpha-1/2}\,|\mathcal{R}f(t,\theta)|^2\,\mathrm{d}t\,\mathrm{d}\theta\\ &\le \pi^{1/2}\,\frac{\Gamma(\alpha-1/2)}{\Gamma(\alpha)}\int_0^\pi\int_{\mathbb{R}}\int_{\mathbb{R}} r_\alpha^2(x(s),y(s))\,|f(x(s),y(s))|^2\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}\theta.\end{aligned}$$

Applying the transformation

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta)$$

yields $\mathrm{d}x\,\mathrm{d}y = \mathrm{d}t\,\mathrm{d}s$ and, consequently,

$$\begin{aligned}\int_{\mathbb{R}}\int_{\mathbb{R}} r_\alpha^2(x(s),y(s))\,|f(x(s),y(s))|^2\,\mathrm{d}s\,\mathrm{d}t &= \int_{\mathbb{R}}\int_{\mathbb{R}} (1+x^2+y^2)^\alpha\,|f(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y\\ &= \|f\|^2_{\mathrm{L}^2(\mathbb{R}^2,(1+x^2+y^2)^\alpha)}.\end{aligned}$$

Thus, we finally obtain

$$\|\mathcal{R}f\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi),(1+t^2)^{\alpha-1/2})} \le \pi^{3/2}\,\frac{\Gamma(\alpha-1/2)}{\Gamma(\alpha)}\,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^2,(1+x^2+y^2)^\alpha)},$$

which completes the proof. □

In the next proposition we finally determine the adjoint operator $\mathcal{R}^{\#}$ of the Radon transform as an operator

$$\mathcal{R} : \mathrm{L}^2(\mathbb{R}^2) \supset \mathrm{L}^2_c(\mathbb{R}^2) \longrightarrow \mathrm{L}^2(\mathbb{R} \times [0, \pi)).$$

The result can be found in [117], where no complete proof is given.

Proposition 3.1.4 (Adjoint of the Radon transform, see [117, Theorem 12.3])**.** *The adjoint operator of the Radon transform $\mathcal{R} : \mathrm{L}^2_c(\mathbb{R}^2) \longrightarrow \mathrm{L}^2(\mathbb{R} \times [0, \pi))$ is given by*

$$\mathcal{R}^{\#} g(x, y) = \int_0^{\pi} g(x \cos(\theta) + y \sin(\theta), \theta) \, \mathrm{d}\theta \quad \text{for } (x, y) \in \mathbb{R}^2.$$

For $g \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$, $\mathcal{R}^{\#} g$ is defined almost everywhere on $\mathbb{R}^2$ and satisfies $\mathcal{R}^{\#} g \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$.

Proof. Let $f \in \mathrm{L}^2_c(\mathbb{R}^2)$ and $g \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$. Then, we have $\mathcal{R} f \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$ due to Proposition 3.1.2 and by the definition of the Radon transform follows that

$$\begin{aligned}
(\mathcal{R} f, g)_{\mathrm{L}^2(\mathbb{R} \times [0, \pi))} &= \int_0^{\pi} \int_{\mathbb{R}} \mathcal{R} f(t, \theta) \, g(t, \theta) \, \mathrm{d}t \, \mathrm{d}\theta \\
&= \int_0^{\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t \cos(\theta) - s \sin(\theta), t \sin(\theta) + s \cos(\theta)) \, \mathrm{d}s \right) g(t, \theta) \, \mathrm{d}t \, \mathrm{d}\theta.
\end{aligned}$$

By applying Fubini's theorem and integration by substitution for real-valued functions with

$$x = t \cos(\theta) - s \sin(\theta) \quad \text{and} \quad y = t \sin(\theta) + s \cos(\theta),$$

i.e., $\mathrm{d}x \, \mathrm{d}y = \mathrm{d}t \, \mathrm{d}s$ and

$$t = x \cos(\theta) + y \sin(\theta) \quad \text{and} \quad s = -x \sin(\theta) + y \cos(\theta),$$

first for $f \geq 0$ and $g \geq 0$ and then also for $f \in \mathrm{L}^2_c(\mathbb{R}^2)$ and $g \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$, we get

$$\begin{aligned}
(\mathcal{R} f, g)_{\mathrm{L}^2(\mathbb{R} \times [0, \pi))} &= \int_0^{\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} f(t \cos(\theta) - s \sin(\theta), t \sin(\theta) + s \cos(\theta)) \, g(t, \theta) \, \mathrm{d}s \, \mathrm{d}t \, \mathrm{d}\theta \\
&= \int_0^{\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \, g(x \cos(\theta) + y \sin(\theta), \theta) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}\theta \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \left(\int_0^{\pi} g(x \cos(\theta) + y \sin(\theta), \theta) \, \mathrm{d}\theta \right) \mathrm{d}x \, \mathrm{d}y.
\end{aligned}$$

Thus, we have

$$(\mathcal{R} f, g)_{\mathrm{L}^2(\mathbb{R} \times [0, \pi))} = (f, \mathcal{R}^{\#} g)_{\mathrm{L}^2(\mathbb{R}^2)}$$

with the adjoint operator $\mathcal{R}^{\#}$ of $\mathcal{R}$ given by

$$\mathcal{R}^{\#} g(x, y) = \int_0^{\pi} g(x \cos(\theta) + y \sin(\theta), \theta) \, \mathrm{d}\theta \quad \text{for } (x, y) \in \mathbb{R}^2.$$

For $g \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$, the function $\mathcal{R}^{\#} g$ is defined almost everywhere on $\mathbb{R}^2$ and, further, for all compact subsets $K \subset \mathbb{R}^2$ we have

$$\begin{aligned}
\| \mathcal{R}^{\#} g \|^2_{\mathrm{L}^2(K)} &= \int_K |\mathcal{R}^{\#} g(x, y)|^2 \, \mathrm{d}(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} |\mathcal{R}^{\#} g(x, y)|^2 \, \chi_K(x, y) \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \int_0^{\pi} g(x \cos(\theta) + y \sin(\theta), \theta) \, \mathrm{d}\theta \right|^2 \chi_K(x, y) \, \mathrm{d}x \, \mathrm{d}y.
\end{aligned}$$

An application of the Cauchy-Schwarz inequality yields

$$\| \mathcal{R}^{\#} g \|^2_{\mathrm{L}^2(K)} \leq \pi \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{\pi} |g(x \cos(\theta) + y \sin(\theta), \theta)|^2 \, \mathrm{d}\theta \, \chi_K(x, y) \, \mathrm{d}x \, \mathrm{d}y.$$

By using Fubini's theorem for non-negative functions and again the transformation

$$t = x\cos(\theta) + y\sin(\theta) \quad \text{and} \quad s = -x\sin(\theta) + y\cos(\theta),$$

i.e., $\mathrm{d}x\,\mathrm{d}y = \mathrm{d}s\,\mathrm{d}t$ and

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta),$$

we finally obtain

$$\begin{aligned}
\|\mathcal{R}^{\#}g\|^2_{\mathrm{L}^2(K)} &\leq \pi \int_0^{\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} |g(t,\theta)|^2 \, \chi_K(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) \,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}\theta \\
&= \pi \int_0^{\pi} \int_{\mathbb{R}} |g(t,\theta)|^2 \left(\int_{\ell_{t,\theta}} \chi_K(x,y) \,\mathrm{d}(x,y) \right) \mathrm{d}t\,\mathrm{d}\theta \\
&\leq \pi \, \operatorname{diam}(K) \, \|g\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} < \infty.
\end{aligned}$$

Consequently, $\mathcal{R}^{\#}g \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$ for all $g \in \mathrm{L}^2(\mathbb{R}\times[0,\pi))$. □

The above Proposition 3.1.4 shows that the back projection operator $\mathcal{B}$ is the adjoint operator of the Radon transform $\mathcal{R}$ up to the constant $\frac{1}{\pi}$, i.e.,

$$\mathcal{B} = \frac{1}{\pi}\,\mathcal{R}^{\#}.$$

In particular, for all $g \in \mathrm{L}^2(\mathbb{R}\times[0,\pi))$ the function $\mathcal{B}g$ is defined almost everywhere on $\mathbb{R}^2$ and satisfies

$$\mathcal{B}g \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2).$$

3.2 Approximate reconstruction formula

With the preliminary propositions of the above section we are now prepared to define the approximate FBP reconstruction f_L and derive its representation (3.1) under reasonable assumptions.

Throughout this work, we assume that we are given some bandwidth $L > 0$ and a window function $W : \mathbb{R} \longrightarrow \mathbb{R}$ that satisfies $W \in \mathrm{L}^\infty(\mathbb{R})$ and the following two properties:

(i) W is an even function, i.e., $W(S) = W(-S)$ for all $S \in \mathbb{R}$;

(ii) W is compactly supported with $\operatorname{supp}(W) \subseteq [-1,1]$, i.e., $W(S) = 0$ for all $|S| > 1$.

With that we define the *low-pass filter* $A_L : \mathbb{R} \longrightarrow \mathbb{R}$ via

$$A_L(S) = |S|\, W_L(S) = |S|\, W(S/L) \quad \text{for } S \in \mathbb{R}.$$

Because of the compact support of W we have $A_L \in \mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^2(\mathbb{R})$ for all $L > 0$ and

$$\operatorname{supp}(A_L) \subseteq [-L, L].$$

Now, let the target function f satisfy $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$. Based on the filtered back projection formula (2.9) we define the approximate FBP reconstruction f_L via

$$f_L(x,y) = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[A_L(S)\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \quad \text{for } (x,y) \in \mathbb{R}^2. \tag{3.3}$$

We will see in this section that (3.3) defines a *band-limited* approximation of the target function f.

Definition 3.2.1 (Band-limited function)**.** *Any function f whose Fourier transform $\mathcal{F}f$ has compact support is called a* band-limited *function.*

In the first theorem we show that f_L is defined almost everywhere on $\mathbb{R}^2$ and can be simplified as

$$f_L = \frac{1}{2}\,\mathcal{B}(q_L * \mathcal{R}f), \tag{3.4}$$

where we define the band-limited function $q_L : \mathbb{R} \times [0,\pi) \longrightarrow \mathbb{R}$ via

$$q_L(S,\theta) = \mathcal{F}^{-1}A_L(S) \quad \text{for } (S,\theta) \in \mathbb{R} \times [0,\pi). \tag{3.5}$$

We remark that q_L is well-defined on $\mathbb{R} \times [0,\pi)$ and satisfies $q_L \in \mathrm{L}^\infty(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ due to the Riemann-Lebesgue Lemma A.1.2 and the Rayleigh-Plancherel Theorem A.1.12, since we have $A_L \in \mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^2(\mathbb{R})$ for all $L > 0$. Further, the definition of q_L is independent of the angle $\theta \in [0,\pi)$ and only depends on the radial variable $S \in \mathbb{R}$.

Note that we call any application of the approximate FBP formula (3.4) an *FBP method*.

Theorem 3.2.2 (Simplification of f_L)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with compact support $\operatorname{supp}(W) \subseteq [-1,1]$. Then, for given bandwidth $L > 0$ the approximate FBP reconstruction f_L, given by*

$$f_L(x,y) = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[A_L(S)\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \quad \text{for } (x,y) \in \mathbb{R}^2,$$

is defined almost everywhere on $\mathbb{R}^2$ and satisfies $f_L \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$. Further, f_L can be rewritten as

$$f_L = \frac{1}{2}\,\mathcal{B}(q_L * \mathcal{R}f) \quad \text{a.e. on } \mathbb{R}^2,$$

where the band-limited function $q_L : \mathbb{R} \times [0,\pi) \longrightarrow \mathbb{R}$ is defined via

$$q_L(S,\theta) = \mathcal{F}^{-1}A_L(S) \quad \text{for } (S,\theta) \in \mathbb{R} \times [0,\pi).$$

Proof. Let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with $\operatorname{supp}(W) \subseteq [-1,1]$ and $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ be given. Then, we have $\mathcal{R}f(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R})$ for all $\theta \in [0,\pi)$ according to Proposition 2.2.7 with

$$\|\mathcal{R}f(\cdot,\theta)\|_{\mathrm{L}^1(\mathbb{R})} \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^2)} \quad \forall\, \theta \in [0,\pi)$$

and, thus, $\mathcal{F}(\mathcal{R}f)(\cdot,\theta) \in \mathscr{C}_0(\mathbb{R})$ for all $\theta \in [0,\pi)$ due to the Riemann-Lebesgue Lemma A.1.2. Since $W \in \mathrm{L}^\infty(\mathbb{R})$ has compact support, we further have $A_L \in \mathrm{L}^2(\mathbb{R})$ and for fixed $\theta \in [0,\pi)$ follows that the function

$$S \longmapsto A_L(S)\mathcal{F}(\mathcal{R}f)(S,\theta)$$

is in $\mathrm{L}^2(\mathbb{R})$. Hence, applying the Rayleigh-Plancherel Theorem A.1.12 shows that the function

$$(S,\theta) \longmapsto \mathcal{F}^{-1}[A_L(S)\mathcal{F}(\mathcal{R}f)(S,\theta)]$$

belongs to $\mathrm{L}^2(\mathbb{R} \times [0,\pi))$ and so

$$(x,y) \longmapsto \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[A_L(S)\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) = f_L(x,y)$$

is defined almost everywhere on $\mathbb{R}^2$ by Proposition 3.1.4 and we have $f_L \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$.

Recall that the band-limited function $q_L = \mathcal{F}^{-1}A_L$ satisfies $q_L \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$. Therefore, for all $\theta \in [0,\pi)$ the Fourier inversion formula

$$A_L(S) = \mathcal{F}(\mathcal{F}^{-1}A_L)(S) = \mathcal{F}q_L(S,\theta)$$

holds in L^2-sense due to Corollary A.1.13 and, in particular, for almost all $S \in \mathbb{R}$. Thus, an application of the Fourier convolution theorem, Proposition 3.1.1, shows that

$$A_L(S)\mathcal{F}(\mathcal{R}f)(S,\theta) = \mathcal{F}q_L(S,\theta)\mathcal{F}(\mathcal{R}f)(S,\theta) = \mathcal{F}(q_L * \mathcal{R}f)(S,\theta)$$

holds in L^2-sense for fixed $\theta \in [0,\pi)$, since $\mathcal{R}f(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R})$ and $q_L(\cdot,\theta) \in \mathrm{L}^2(\mathbb{R})$ for all $\theta \in [0,\pi)$.

Furthermore, we have $(q_L * \mathcal{R}f)(\cdot, \theta) \in \mathrm{L}^2(\mathbb{R})$ for all $\theta \in [0, \pi)$ according to Young's inequality, Theorem A.1.16, and, consequently, the Fourier inversion formula holds again in L^2-sense. Hence, for fixed $\theta \in [0, \pi)$ we obtain the representation

$$(q_L * \mathcal{R}f)(S, \theta) = \mathcal{F}^{-1}[\mathcal{F}(q_L * \mathcal{R}f)(S, \theta)] = \mathcal{F}^{-1}[A_L(S)\mathcal{F}(\mathcal{R}f)(S, \theta)]$$

for almost all $S \in \mathbb{R}$. But this already implies the desired representation of the approximate FBP reconstruction f_L as

$$f_L = \frac{1}{2}\,\mathcal{B}(q_L * \mathcal{R}f) \quad \text{a.e. on } \mathbb{R}^2,$$

where we have $f_L \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$ due to Proposition 3.1.4, since $(q_L * \mathcal{R}f) \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$. □

In the following we show that the approximate FBP reconstruction f_L satisfies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ and can be directly written in terms of the target function f. To be more precise, we show that

$$f_L = f * K_L \tag{3.6}$$

with the convolution kernel $K_L : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$K_L(x, y) = \frac{1}{2}\,\mathcal{B}q_L(x, y) \quad \text{for } (x, y) \in \mathbb{R}^2.$$

Since $q_L \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$, the convolution kernel K_L is defined almost everywhere on $\mathbb{R}^2$ and satisfies $K_L \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$ by Proposition 3.1.4.

However, in the next theorem we first show that K_L is actually in $\mathrm{L}^2(\mathbb{R}^2)$ and determine its Fourier transform. To this end, we need to define the bivariate window function $W_L : \mathbb{R}^2 \longrightarrow \mathbb{R}$ as

$$W_L(x, y) = W\left(\frac{r(x, y)}{L}\right) \quad \text{for } (x, y) \in \mathbb{R}^2, \tag{3.7}$$

where we let

$$r(x, y) = \sqrt{x^2 + y^2} \quad \text{for } (x, y) \in \mathbb{R}^2.$$

Theorem 3.2.3 (Convolution kernel K_L)**.** *Let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with $\operatorname{supp}(W) \subseteq [-1, 1]$. Then, for all $L > 0$ the convolution kernel $K_L : \mathbb{R}^2 \longrightarrow \mathbb{R}$, defined as*

$$K_L(x, y) = \frac{1}{2}\,\mathcal{B}q_L(x, y) \quad \textit{for } (x, y) \in \mathbb{R}^2,$$

satisfies $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and its Fourier transform is given by

$$\mathcal{F}K_L(x, y) = W_L(x, y) \quad \textit{for almost all } (x, y) \in \mathbb{R}^2.$$

Proof. Since $W \in \mathrm{L}^\infty(\mathbb{R})$ has compact support, the bivariate window function W_L is compactly supported and satisfies $W_L \in \mathrm{L}^p(\mathbb{R}^2)$ for all $1 \leq p \leq \infty$. Indeed, for $1 \leq p < \infty$ holds that

$$\begin{aligned}\|W_L\|^p_{\mathrm{L}^p(\mathbb{R}^2)} &= \int_{\mathbb{R}}\int_{\mathbb{R}} |W_L(x, y)|^p \,\mathrm{d}x\,\mathrm{d}y = \int_{r(x,y)\leq L} \left|W\left(\frac{r(x, y)}{L}\right)\right|^p \,\mathrm{d}(x, y) \\ &\leq \|W\|^p_{\mathrm{L}^\infty(\mathbb{R})}\,|B_L(0)| = \pi\,L^2\,\|W\|^p_{\mathrm{L}^\infty(\mathbb{R})}.\end{aligned}$$

In particular, we have $W_L \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$. Consequently, with the Riemann-Lebesgue lemma and the Rayleigh-Plancherel theorem follows that $\mathcal{F}^{-1}W_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$. Furthermore, for all $(x, y) \in \mathbb{R}^2$ we obtain

$$\begin{aligned}\mathcal{F}^{-1}W_L(x, y) &= \frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} W_L(X, Y)\,\mathrm{e}^{\mathrm{i}(xX+yY)}\,\mathrm{d}X\,\mathrm{d}Y \\ &= \frac{1}{4\pi^2}\int_0^\pi\int_{\mathbb{R}} W(S/L)\,|S|\,\mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\,\mathrm{d}S\,\mathrm{d}\theta \\ &= \frac{1}{4\pi^2}\int_0^\pi\int_{\mathbb{R}} A_L(S)\,\mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\,\mathrm{d}S\,\mathrm{d}\theta\end{aligned}$$

by transforming $(X, Y) = (S\cos(\theta), S\sin(\theta))$ from Cartesian coordinates to polar coordinates.

Recall that the filter A_L satisfies $A_L \in \mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^2(\mathbb{R})$ and we have $q_L \in \mathrm{L}^2(\mathbb{R} \times [0, \pi))$. Thus, with the definition of the back projection operator $\mathcal{B}$ follows that

$$\mathcal{F}^{-1} W_L(x, y) = \frac{1}{2\pi} \int_0^\pi q_L(x \cos(\theta) + y \sin(\theta), \theta) \, \mathrm{d}\theta = \frac{1}{2} \, \mathcal{B} q_L(x, y) = K_L(x, y) \quad \forall \, (x, y) \in \mathbb{R}^2.$$

Consequently, $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and applying the Fourier inversion formula shows that

$$\mathcal{F} K_L = W_L$$

holds in L^2-sense and, in particular, almost everywhere on $\mathbb{R}^2$. □

Before we proceed, we wish to add one remark concerning the convolution kernel K_L.

Remark 3.2.4. *Since the bivariate window function W_L has compact support and satisfies $W_L \in \mathrm{L}^1(\mathbb{R}^2)$, its inverse Fourier transform is analytic due to the Paley-Wiener Theorem A.1.14. Consequently, the convolution kernel K_L not only satisfies $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$, but also lies in $\mathscr{C}^\infty(\mathbb{R}^2)$. Thus, we have $K_L \in \mathscr{C}_0^\infty(\mathbb{R}^2)$, where*

$$\mathscr{C}_0^\infty(\mathbb{R}^2) = \left\{ f \in \mathscr{C}^\infty(\mathbb{R}^2) \mid f(x, y) \longrightarrow 0 \text{ for } \|(x, y)\|_{\mathbb{R}^2} \longrightarrow \infty \right\},$$

and K_L has compact support in $\mathbb{R}^2$ if and only if $W \equiv 0$. Furthermore, the Riemann-Lebesgue Lemma A.1.2 shows that K_L is typically not *in $\mathrm{L}^1(\mathbb{R}^2)$, since this would imply that the window function W is continuous on $\mathbb{R}$, i.e., $W \in \mathscr{C}(\mathbb{R})$.*

We are now in the position to show the desired representation (3.6) of the approximate FBP reconstruction f_L, i.e.,

$$f_L = f * K_L,$$

in the L^2-sense. In particular, we show that f_L satisfies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ and that its Fourier transform is given by

$$\mathcal{F} f_L = W_L \cdot \mathcal{F} f$$

with the compactly supported and radially symmetric bivariate window function $W_L \in \mathrm{L}^\infty(\mathbb{R}^2)$ defined in (3.7).

Theorem 3.2.5 (Approximate reconstruction f_L)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with $\operatorname{supp}(W) \subseteq [-1, 1]$. Then, for all $L > 0$ the approximate FBP reconstruction f_L satisfies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ and can be written as*

$$f_L = f * K_L$$

in the L^2-sense and, in particular, almost everywhere on $\mathbb{R}^2$.

Proof. Since $f \in \mathrm{L}^1(\mathbb{R}^2)$ by assumption and $K_L \in \mathrm{L}^2(\mathbb{R}^2)$ due to Theorem 3.2.3, applying Young's inequality yields $(f * K_L) \in \mathrm{L}^2(\mathbb{R}^2)$ and, thus, the Fourier inversion formula

$$(f * K_L)(x, y) = \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}(f * K_L)(X, Y) \, \mathrm{e}^{\mathrm{i}(xX + yY)} \, \mathrm{d}X \, \mathrm{d}Y$$

holds in L^2-sense and, in particular, for almost all $(x, y) \in \mathbb{R}^2$. Additionally, from the Fourier convolution theorem, Proposition 3.1.1, and Theorem 3.2.3 we obtain

$$\mathcal{F}(f * K_L) = \mathcal{F} f \cdot \mathcal{F} K_L = W_L \cdot \mathcal{F} f$$

in the L^2-sense and, thus, almost everywhere on $\mathbb{R}^2$. This gives the representation

$$(f * K_L)(x, y) = \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F} f(X, Y) \, W_L(X, Y) \, \mathrm{e}^{\mathrm{i}(xX + yY)} \, \mathrm{d}X \, \mathrm{d}Y.$$

Since $W \in \mathrm{L}^\infty(\mathbb{R})$ has compact support and $f \in \mathrm{L}^1(\mathbb{R}^2)$, we can apply integration by substitution and obtain

$$(f * K_L)(x,y) = \frac{1}{4\pi^2} \int_0^\pi \int_{\mathbb{R}} \mathcal{F}f(S\cos(\theta), S\sin(\theta))\, W(S/L)\, |S|\, \mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\, \mathrm{d}S\, \mathrm{d}\theta$$

by transforming $(X,Y) = (S\cos(\theta), S\sin(\theta))$ from Cartesian coordinates to polar coordinates. Furthermore, for $f \in \mathrm{L}^1(\mathbb{R}^2)$ the Fourier slice Theorem 2.2.11 gives

$$\mathcal{F}f(S\cos(\theta), S\sin(\theta)) = \mathcal{F}(\mathcal{R}f)(S,\theta) \quad \forall\, (S,\theta) \in \mathbb{R} \times [0,\pi),$$

which implies that

$$\begin{aligned}(f * K_L)(x,y) &= \frac{1}{4\pi^2} \int_0^\pi \int_{\mathbb{R}} \mathcal{F}(\mathcal{R}f)(S,\theta)\, W(S/L)\, |S|\, \mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\, \mathrm{d}S\, \mathrm{d}\theta \\ &= \frac{1}{4\pi^2} \int_0^\pi \int_{\mathbb{R}} \mathcal{F}(\mathcal{R}f)(S,\theta)\, A_L(S)\, \mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\, \mathrm{d}S\, \mathrm{d}\theta.\end{aligned}$$

Since the filter A_L is in $\mathrm{L}^1(\mathbb{R}) \cap \mathrm{L}^2(\mathbb{R})$, we have $\mathcal{F}^{-1}A_L \in \mathrm{L}^2(\mathbb{R})$ and the Fourier inversion formula gives

$$A_L(S) = \mathcal{F}(\mathcal{F}^{-1}A_L)(S)$$

in L^2-sense and, particularly, for almost all $S \in \mathbb{R}$. Therefore, by Young's inequality we have $(q_L * \mathcal{R}f)(\cdot,\theta) \in \mathrm{L}^2(\mathbb{R})$ for all $\theta \in [0,\pi)$, where the band-limited function $q_L \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ is given by (3.5), i.e.,

$$q_L(S,\theta) = \mathcal{F}^{-1}A_L(S) \quad \text{for } (S,\theta) \in \mathbb{R} \times [0,\pi).$$

Moreover, applying the Fourier convolution theorem, Proposition 3.1.1, shows that the identity

$$\mathcal{F}(q_L * \mathcal{R}f)(S,\theta) = \mathcal{F}q_L(S,\theta)\, \mathcal{F}(\mathcal{R}f)(S,\theta) = A_L(S)\, \mathcal{F}(\mathcal{R}f)(S,\theta)$$

holds in L^2-sense for fixed $\theta \in [0,\pi)$, since $\mathcal{R}f(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R})$ and $q_L(\cdot,\theta) \in \mathrm{L}^2(\mathbb{R})$ for all $\theta \in [0,\pi)$. Hence, we obtain

$$\begin{aligned}(f * K_L)(x,y) &= \frac{1}{4\pi^2} \int_0^\pi \int_{\mathbb{R}} \mathcal{F}(q_L * \mathcal{R}f)(S,\theta)\, \mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\, \mathrm{d}S\, \mathrm{d}\theta \\ &= \frac{1}{2\pi} \int_0^\pi \mathcal{F}^{-1}[\mathcal{F}(q_L * \mathcal{R}f)](x\cos(\theta) + y\sin(\theta), \theta)\, \mathrm{d}\theta.\end{aligned}$$

Since $(q_L * \mathcal{R}f)(\cdot,\theta) \in \mathrm{L}^2(\mathbb{R})$ for all $\theta \in [0,\pi)$, the Fourier inversion formula holds again in L^2-sense and, in particular, almost everywhere on $\mathbb{R}$. Thus, for fixed $\theta \in [0,\pi)$ we get

$$(q_L * \mathcal{R}f)(S,\theta) = \mathcal{F}^{-1}\big[\mathcal{F}(q_L * \mathcal{R}f)\big](S,\theta)$$

for almost all $S \in \mathbb{R}$ and the definition of the back projection operator $\mathcal{B}$ in combination with Theorem 3.2.2 yields

$$(f * K_L)(x,y) = \frac{1}{2\pi} \int_0^\pi (q_L * \mathcal{R}f)(x\cos(\theta) + y\sin(\theta), \theta)\, \mathrm{d}\theta = \frac{1}{2}\, \mathcal{B}(q_L * \mathcal{R}f)(x,y) = f_L(x,y)$$

for almost all $(x,y) \in \mathbb{R}^2$. This finally implies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ with

$$f_L = \frac{1}{2}\, \mathcal{B}(q_L * \mathcal{R}f) = f * K_L$$

in the L^2-sense and, in particular, almost everywhere on $\mathbb{R}^2$. □

Combining Theorems 3.2.3 and 3.2.5 allows us to determine the Fourier transform $\mathcal{F}f_L$ of the approximate FBP reconstruction f_L.

Corollary 3.2.6 (Fourier transform of f_L)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with $\operatorname{supp}(W) \subseteq [-1,1]$. Then, for all $L > 0$ the Fourier transform $\mathcal{F}f_L$ of the approximate FBP reconstruction f_L is given by*

$$\mathcal{F}f_L = W_L \cdot \mathcal{F}f$$

in L^2-sense and, in particular, almost everywhere on $\mathbb{R}^2$. □

Since the window function W is assumed to be compactly supported, Corollary 3.2.6 shows that the approximate FBP reconstruction formula (3.3) provides a band-limited approximation f_L to the target function f. In particular, the approximation f_L is arbitrarily smooth.

Remark 3.2.7. *Since the target function f and the convolution kernel K_L are both in $\mathrm{L}^2(\mathbb{R}^2)$, the representation* (3.6) *of the approximate FBP reconstruction f_L implies that f_L is bounded and continuous on $\mathbb{R}^2$ due to Theorem A.1.17. In addition, Corollary 3.2.6 and the Paley-Wiener Theorem A.1.14 show that f_L is analytic, since the window function W is compactly supported. Thus, we not only have $f_L \in \mathrm{L}^2(\mathbb{R}^2)$, but also $f_L \in \mathscr{C}_b^\infty(\mathbb{R}^2)$, where*

$$\mathscr{C}_b^\infty(\mathbb{R}^2) = \left\{ f \in \mathscr{C}^\infty(\mathbb{R}^2) \mid \|f\|_\infty < \infty \right\}.$$

Moreover, the FBP approximation f_L of $f \not\equiv 0$ has compact support in $\mathbb{R}^2$ if and only if $W \equiv 0$.

We now list some window functions of suitable low-pass filters, which are commonly use in computerized tomography and can be found in the literature, see, e.g., [18], [83], [96], [112], [113]. To this end, let $\prod_L : \mathbb{R} \longrightarrow \mathbb{R}$ denote the characteristic function of the interval $[-L, L]$ for $L > 0$, i.e.,

$$\prod_L(S) = \chi_{[-L,L]}(S) = \begin{cases} 1 & \text{for } |S| \leq L \\ 0 & \text{for } |S| > L. \end{cases}$$

For the sake of brevity, we set $\prod \equiv \prod_1$.

Example 3.2.8. *The* Ram-Lak filter *is given by the window function*

$$W(S) = \prod(S) \quad \text{for } S \in \mathbb{R}$$

such that

$$A_L(S) = |S| \cdot \prod_L(S) = \begin{cases} |S| & \text{for } |S| \leq L \\ 0 & \text{for } |S| > L. \end{cases}$$

A plot of the Ram-Lak window can be found in Figure 3.1(a).

In the following, let sinc denote the *unnormalized* cardinal sine function.

Example 3.2.9. *The* Shepp-Logan filter *is given by the window function*

$$W(S) = \operatorname{sinc}\left(\frac{\pi S}{2}\right) \cdot \prod(S) \quad \text{for } S \in \mathbb{R}$$

such that

$$A_L(S) = |S| \cdot \operatorname{sinc}\left(\frac{\pi S}{2L}\right) \cdot \prod_L(S) = \begin{cases} \frac{2L}{\pi} \left| \sin\left(\frac{\pi S}{2L}\right) \right| & \text{for } |S| \leq L \\ 0 & \text{for } |S| > L. \end{cases}$$

A plot of the Shepp-Logan window can be found in Figure 3.1(b).

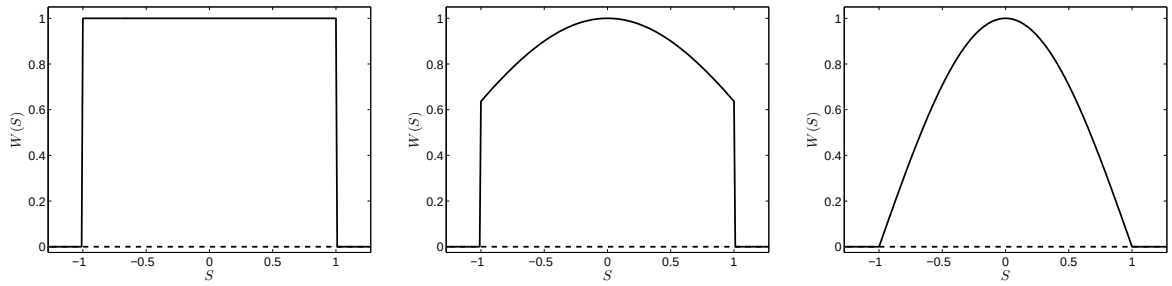
Example 3.2.10. *The* Cosine filter *is given by the window function*

$$W(S) = \cos\left(\frac{\pi S}{2}\right) \cdot \prod(S) \quad \text{for } S \in \mathbb{R}$$

such that

$$A_L(S) = |S| \cdot \cos\left(\frac{\pi S}{2L}\right) \cdot \prod_L(S) = \begin{cases} |S| \cdot \cos\left(\frac{\pi S}{2L}\right) & \text{for } |S| \leq L \\ 0 & \text{for } |S| > L. \end{cases}$$

A plot of the Cosine window can be found in Figure 3.1(c).



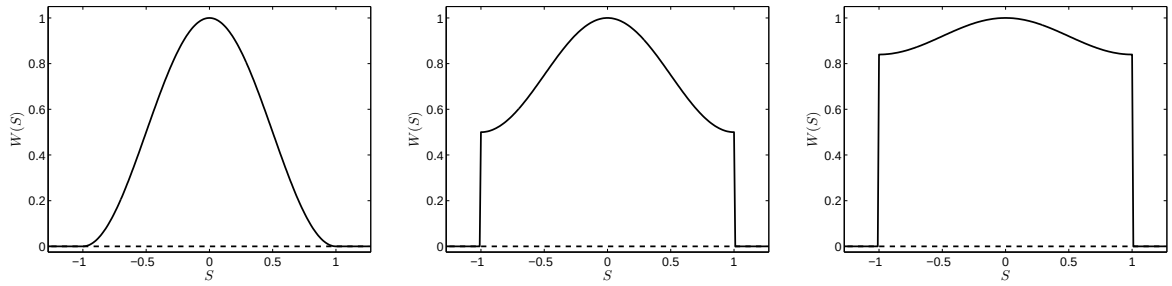
(a) Ram-Lak (b) Shepp-Logan (c) Cosine

Figure 3.1: Window functions of three typical low-pass filters.

Example 3.2.11. *The* Hamming filter *with parameter* $\beta \in \left[\frac{1}{2}, 1\right]$ *is given by the window function*

$$W(S) = (\beta + (1-\beta)\cos(\pi S)) \cdot \prod(S) \quad \text{for } S \in \mathbb{R}.$$

Thus, the Hamming filter is a combination of the Ram-Lak filter and a modified Cosine filter. Plots of the Hamming window with parameters $\beta \in \{0.5, 0.75, 0.92\}$ *can be found in Figure 3.2.*



(a) $\beta = 0.5$ (b) $\beta = 0.75$ (c) $\beta = 0.92$

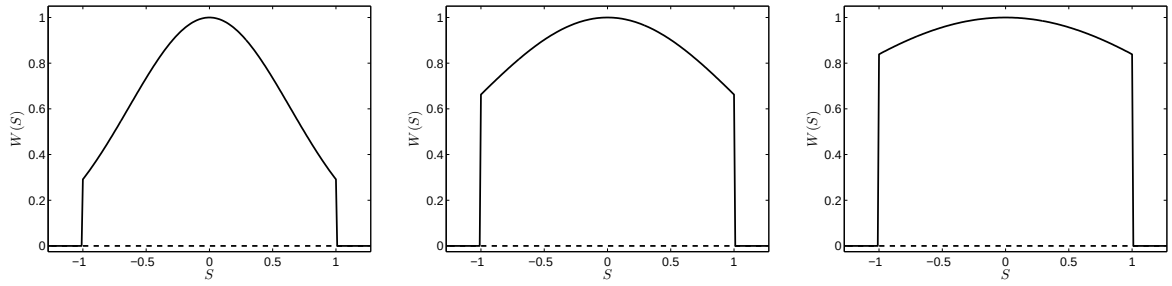
Figure 3.2: Window functions of the Hamming filter with $\beta \in \{0.5, 0.75, 0.92\}$.

Example 3.2.12. *The* Gaussian filter *with parameter* $\beta > 1$ *is given by the window function*

$$W(S) = \exp\left(-\left(\frac{\pi S}{\beta}\right)^2\right) \cdot \prod(S) \quad \text{for } S \in \mathbb{R}.$$

Plots of the Gaussian window with parameters $\beta \in \{2.83, 4.9, 7.5\}$ *can be found in Figure 3.3.*

We summarize the discussion of this section as follows.



(a) $\beta = 2.83$ (b) $\beta = 4.9$ (c) $\beta = 7.5$

Figure 3.3: Window functions of the Gaussian filter with $\beta \in \{2.83, 4.9, 7.5\}$.

Conclusion 3.2.13 (Approximate FBP reconstruction)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and consider the low-pass filter*

$$A_L(S) = |S| \cdot W(S/L) \quad \text{for } S \in \mathbb{R}$$

with bandwidth $L > 0$ and an even window $W \in \mathrm{L}^\infty(\mathbb{R})$ with compact support $\operatorname{supp}(W) \subseteq [-1, 1]$. Then, the approximate FBP reconstruction

$$f_L(x, y) = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[A_L(S)\mathcal{F}(\mathcal{R}f)(S, \theta)])(x, y) \quad \text{for } (x, y) \in \mathbb{R}^2$$

satisfies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ and can be rewritten as

$$f_L = \frac{1}{2}\,\mathcal{B}(q_L * \mathcal{R}f) = f * K_L \quad \text{a.e. on } \mathbb{R}^2$$

with the band-limited function

$$q_L(S, \theta) = \mathcal{F}^{-1}A_L(S) \quad \text{for } (S, \theta) \in \mathbb{R} \times [0, \pi)$$

and the convolution kernel

$$K_L(x, y) = \frac{1}{2}\,\mathcal{B}q_L(x, y) \quad \text{for } (x, y) \in \mathbb{R}^2.$$

Moreover, its Fourier transform $\mathcal{F}f_L \in \mathrm{L}^2(\mathbb{R}^2)$ is given by

$$\mathcal{F}f_L = W_L \cdot \mathcal{F}f \quad \text{a.e. on } \mathbb{R}^2$$

with the radially symmetric bivariate window

$$W_L(x, y) = W\left(\frac{r(x, y)}{L}\right) \quad \text{for } (x, y) \in \mathbb{R}^2,$$

where

$$r(x, y) = \sqrt{x^2 + y^2}.$$

□

3.3 Properties of the convolution kernel

Based on the results of the previous section we now prove some basic properties of the convolution kernel K_L. To be more precise, we show that K_L is radially symmetric, possesses a scaling property with respect to the bandwidth L and is a positive definite function.

The proofs of these features of K_L are based on the following general stretch property of the Fourier transform. Although the result is well-known, we give a proof for completeness.

Proposition 3.3.1 (Fourier stretch theorem)**.** *Let $f \in \mathrm{L}^p(\mathbb{R}^n)$ with $p \in \{1, 2\}$ and let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix. For $1 \le q \le \infty$, we define the stretch operator $\sigma_A : \mathrm{L}^q(\mathbb{R}^n) \longrightarrow \mathrm{L}^q(\mathbb{R}^n)$ via*

$$\sigma_A g(x) = g(Ax) \quad \text{for } x \in \mathbb{R}^n.$$

Then, we have

$$\mathcal{F}(\sigma_A f) = \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f).$$

Proof. We first note that $\sigma_A g$ is again in $\mathrm{L}^q(\mathbb{R}^n)$ for $g \in \mathrm{L}^q(\mathbb{R}^n)$ with $1 \le q \le \infty$. Indeed, for $1 \le q < \infty$, an application of the transformation theorem yields

$$\|\sigma_A g\|^q_{\mathrm{L}^q(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |g(Ax)|^q \,\mathrm{d}x = \frac{1}{|\det(A)|} \int_{\mathbb{R}^n} |g(x)|^q \,\mathrm{d}x = \frac{1}{|\det(A)|}\,\|g\|^q_{\mathrm{L}^p(\mathbb{R}^n)} < \infty \qquad (3.8)$$

and, for $q = \infty$, the invertibility of A gives

$$\|\sigma_A g\|_{\mathrm{L}^\infty(\mathbb{R}^n)} = \operatorname*{ess\,sup}_{x\in\mathbb{R}^n} |\sigma_A g(x)| = \operatorname*{ess\,sup}_{x\in\mathbb{R}^n} |g(Ax)| = \operatorname*{ess\,sup}_{y\in\mathbb{R}^n} |g(y)| = \|g\|_{\mathrm{L}^\infty(\mathbb{R}^n)}.$$

Now, let $f \in \mathrm{L}^1(\mathbb{R}^n)$. Then, for all $\xi \in \mathbb{R}^n$ follows that

$$\begin{aligned}\mathcal{F}(\sigma_A f)(\xi) &= \int_{\mathbb{R}^n} \sigma_A f(x)\, \mathrm{e}^{-i\,x^T\xi}\,\mathrm{d}x = \int_{\mathbb{R}^n} f(Ax)\, \mathrm{e}^{-i\,x^T\xi}\,\mathrm{d}x = \frac{1}{|\det(A)|}\int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-i\,(A^{-1}x)^T\xi}\,\mathrm{d}x \\ &= \frac{1}{|\det(A)|}\int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-i\,x^T(A^{-T}\xi)}\,\mathrm{d}x = \frac{1}{|\det(A)|}\,\mathcal{F}f(A^{-T}\xi) = \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f)(\xi)\end{aligned}$$

by applying again the transformation theorem. Thus, we have

$$\mathcal{F}(\sigma_A f) = \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f) \quad \forall\, f \in \mathrm{L}^1(\mathbb{R}^n).$$

Using density arguments the result also follows for $f \in \mathrm{L}^2(\mathbb{R}^n)$. Indeed, since $\mathrm{L}^1(\mathbb{R}^n)\cap\mathrm{L}^2(\mathbb{R}^n)$ is dense in $\mathrm{L}^2(\mathbb{R}^n)$, for every $f \in \mathrm{L}^2(\mathbb{R}^n)$ there is a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathrm{L}^1(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$ with

$$\|f - f_k\|_{\mathrm{L}^2(\mathbb{R}^n)} \longrightarrow 0 \quad \text{for} \quad k \longrightarrow \infty.$$

We have already proven that

$$\mathcal{F}(\sigma_A f_k) = \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f_k) \quad \forall\, k \in \mathbb{N}.$$

This in combination with the linearity of the involved operators and estimate (3.8) gives

$$\begin{aligned}\left\|\mathcal{F}(\sigma_A f) - \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f)\right\|_{\mathrm{L}^2(\mathbb{R}^n)} &= \left\|\mathcal{F}(\sigma_A f - \sigma_A f_k) + \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f_k - \mathcal{F}f)\right\|_{\mathrm{L}^2(\mathbb{R}^n)} \\ &\le \|\mathcal{F}(\sigma_A(f - f_k))\|_{\mathrm{L}^2(\mathbb{R}^n)} + \frac{1}{|\det(A)|}\,\|\sigma_{A^{-T}}(\mathcal{F}(f - f_k))\|_{\mathrm{L}^2(\mathbb{R}^n)} \\ &\le (2\pi)^{n/2}\,\|\sigma_A(f - f_k)\|_{\mathrm{L}^2(\mathbb{R}^n)} + \frac{|\det(A)|^{1/2}}{|\det(A)|}\,\|\mathcal{F}(f - f_k)\|_{\mathrm{L}^2(\mathbb{R}^n)} \\ &\le \frac{2\,(2\pi)^{n/2}}{|\det(A)|^{1/2}}\,\|f - f_k\|_{\mathrm{L}^2(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.\end{aligned}$$

Consequently, we also have

$$\mathcal{F}(\sigma_A f) = \frac{1}{|\det(A)|}\,\sigma_{A^{-T}}(\mathcal{F}f) \quad \forall\, f \in \mathrm{L}^2(\mathbb{R}^n)$$

and the proof is complete. □

In the first corollary we deduce the radial symmetry of the convolution kernel K_L.

Corollary 3.3.2 (Radial symmetry of K_L)**.** *Let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with* $\operatorname{supp}(W) \subseteq [-1, 1]$. *Then, for all $L > 0$ the convolution kernel $K_L : \mathbb{R}^2 \longrightarrow \mathbb{R}$, defined as*

$$K_L(x, y) = \frac{1}{2}\,\mathcal{B}q_L(x, y) \quad \text{for } (x, y) \in \mathbb{R}^2,$$

is radially symmetric, i.e., there exists a univariate function $k_L : \mathbb{R}_{\ge 0} \longrightarrow \mathbb{R}$ such that

$$K_L(x, y) = k_L(r(x, y)) \quad \forall\,(x, y) \in \mathbb{R}^2,$$

where

$$r(x, y) = \sqrt{x^2 + y^2} \quad \text{for } (x, y) \in \mathbb{R}^2.$$

Proof. In the proof of Theorem 3.2.3 we have seen that the convolution kernel K_L can be represented as

$$K_L(x,y) = \mathcal{F}^{-1}W_L(x,y) \quad \forall\,(x,y) \in \mathbb{R}^2$$

with the compactly supported and radially symmetric bivariate window function $W_L \in \mathrm{L}^\infty(\mathbb{R}^2)$ given by

$$W_L(x,y) = W\left(\frac{r(x,y)}{L}\right) \quad \text{for } (x,y) \in \mathbb{R}^2.$$

Since $W \in \mathrm{L}^\infty(\mathbb{R})$ is compactly supported, we also have $W_L \in \mathrm{L}^1(\mathbb{R}^2)$ and its radial symmetry in combination with Remark A.1.9 implies that we can rewrite K_L as

$$K_L(x,y) = \frac{1}{4\pi^2}\,\mathcal{F}W_L(x,y) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Now, the function K_L is radially symmetric if and only if it is invariant under all rotations that leave the origin fixed, i.e., if and only if

$$\sigma_Q K_L = K_L \quad \forall\, Q \in \mathrm{SO}(2).$$

This is indeed the case, since for all $Q \in \mathrm{SO}(2)$ the radial symmetry of W_L and Proposition 3.3.1 yield

$$\sigma_Q K_L = \frac{1}{4\pi^2}\,\sigma_Q(\mathcal{F}W_L) = \frac{1}{4\pi^2}\,|\det(Q)|\,\mathcal{F}(\sigma_Q W_L) = \frac{1}{4\pi^2}\,\mathcal{F}W_L = K_L,$$

which shows that K_L is radially symmetric. □

Using Proposition 3.3.1 we can also show the following scaling property of the convolution kernel K_L.

Corollary 3.3.3 (Scaling property of K_L)**.** *Let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with $\mathrm{supp}(W) \subseteq [-1,1]$. Then, for all $L > 0$ the convolution kernel $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ satisfies the scaling property*

$$K_L(x,y) = L^2\,K(L\,x, L\,y) \quad \forall\,(x,y) \in \mathbb{R}^2,$$

where we set $K \equiv K_1$ for the sake of brevity.

Proof. As we have seen before, the convolution kernel K_L can be written as

$$K_L(x,y) = \mathcal{F}^{-1}W_L(x,y) = \frac{1}{4\pi^2}\,\mathcal{F}W_L(x,y) \quad \forall\,(x,y) \in \mathbb{R}^2$$

with the radially symmetric bivariate window function $W_L \in \mathrm{L}^1(\mathbb{R}^2)$ satisfying

$$W_L(x,y) = W\left(\frac{r(x,y)}{L}\right) = W\left(r\left(\frac{x}{L},\frac{y}{L}\right)\right) = W\left(\frac{x}{L},\frac{y}{L}\right) \quad \text{for } (x,y) \in \mathbb{R}^2,$$

where $W \equiv W_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Thus, applying Proposition 3.3.1 yields for all $(x,y) \in \mathbb{R}^2$

$$\begin{aligned} K_L(x,y) &= \frac{1}{4\pi^2}\,\mathcal{F}W_L(x,y) = \frac{1}{4\pi^2}\,\mathcal{F}(\sigma_{\mathrm{diag}(L^{-1})}W)(x,y) \\ &= \frac{1}{4\pi^2}\,|\det(\mathrm{diag}(L))|\,\sigma_{\mathrm{diag}(L)}(\mathcal{F}W)(x,y) = L^2\,K(L\,x, L\,y), \end{aligned}$$

as stated. □

We now show that, under suitable conditions on the window function W, the convolution kernel K_L is a *positive definite function*. These play an important role in the multivariate approximation theory and are defined as follows.

Definition 3.3.4 (Positive definite function, see [123, Definition 6.1]). *A continuous function* $\Phi : \mathbb{R}^n \longrightarrow \mathbb{C}$ *is called* positive semi-definite *on* $\mathbb{R}^n$ *if, for all* $N \in \mathbb{N}$ *and all sets of pairwise distinct points* $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$*, the quadratic form*

$$\sum_{j=1}^{N} \sum_{k=1}^{N} \alpha_j \, \overline{\alpha}_k \, \Phi(x_j - x_k) \tag{3.9}$$

is nonnegative for all $\alpha \in \mathbb{C}^N$*. The function* Φ *is called* positive definite *on* $\mathbb{R}^n$ *if the quadratic form* (3.9) *is positive for all* $\alpha \in \mathbb{C}^N \setminus \{0\}$*.*

The following well-known theorem of Bochner provides a characterization of positive semi-definite functions in terms of Fourier transforms.

Theorem 3.3.5 (Bochner's theorem). *A continuous function* $\Phi : \mathbb{R}^n \longrightarrow \mathbb{C}$ *is positive semi-definite if and only if it is the Fourier transform of a finite nonnegative Borel measure* μ *on* $\mathbb{R}^n$*, i.e.,*

$$\Phi(x) = \mathcal{F}\mu(x) = \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\, x^T \xi} \, \mathrm{d}\mu(\xi) \quad \forall\, x \in \mathbb{R}^n.$$

Proof. See, for example, [123, Theorem 6.6]. □

The last auxiliary result deals with a sufficient condition for positive definiteness.

Theorem 3.3.6. *A positive semi-definite function* $\Phi : \mathbb{R}^n \longrightarrow \mathbb{C}$ *is positive definite if the support of the Borel measure* μ *in the representation*

$$\Phi = \mathcal{F}\mu$$

contains an open subset.

Proof. See, for example, [123, Theorem 6.8]. □

We are now prepared to show that the convolution kernel K_L is a positive definite function, provided that the corresponding window function W satisfies reasonable conditions which are stated in the following corollary.

Corollary 3.3.7 (Positive definiteness of K_L). *Let* $W \in \mathrm{L}^\infty(\mathbb{R})$ *be even with* $\operatorname{supp}(W) \subseteq [-1, 1]$*. Further, let* W *be nonnegative, i.e.,*

$$W(S) \geq 0 \quad \forall\, S \in \mathbb{R}.$$

Then, for all $L > 0$ *the convolution kernel* $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ *is positive semi-definite on* $\mathbb{R}^2$*. If, in addition, there exists a non-empty open interval* $I \subset [-1, 1]$ *such that*

$$W(S) > 0 \quad \forall\, S \in I,$$

then the convolution kernel K_L *is positive definite on* $\mathbb{R}^2$*.*

Proof. We consider the compactly supported bivariate window function $W_L \in \mathrm{L}^\infty(\mathbb{R}^2)$ given by

$$W_L(x, y) = W\Big(\frac{r(x, y)}{L}\Big) \quad \text{for } (x, y) \in \mathbb{R}^2.$$

Thus, we have $W_L \in \mathrm{L}^1(\mathbb{R}^2)$ and $W_L(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$. With this we define the measure μ_L for any Borel set $B \subset \mathbb{R}^2$ via

$$\mu_L(B) = \int_B W_L(x, y) \, \mathrm{d}(x, y).$$

Because $W_L \in \mathrm{L}^1(\mathbb{R}^2)$ is nonnegative, μ_L defines a finite nonnegative Borel measure on $\mathbb{R}^2$.

We now consider the convolution kernel $K_L : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$K_L(x,y) = \frac{1}{2}\,\mathcal{B}q_L(x,y) \quad \text{for } (x,y) \in \mathbb{R}^2.$$

In Theorem 3.2.3, we have seen that K_L satisfies $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and, further, for all $(x,y) \in \mathbb{R}^2$ follows that

$$\begin{aligned} K_L(x,y) &= \mathcal{F}^{-1}W_L(x,y) = \frac{1}{4\pi^2}\,\mathcal{F}W_L(x,y) = \frac{1}{4\pi^2}\int_{\mathbb{R}^2} W_L(X,Y)\,\mathrm{e}^{-\mathrm{i}(xX+yY)}\,\mathrm{d}(X,Y) \\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}^2} \mathrm{e}^{-\mathrm{i}(xX+yY)}\,\mathrm{d}\mu_L(X,Y). \end{aligned}$$

This shows that, up to the constant $\frac{1}{4\pi^2}$, K_L is given by the Fourier transform of the finite non-negative Borel measure μ_L. Thus, K_L is positive semi-definite due to Bochner's theorem 3.3.5.

If, in addition, there is a non-empty open interval $I \subset [-1,1]$ such that W is positive on I, then there exists an open subset $O \subset \mathbb{R}^2$ such that

$$W_L(x,y) > 0 \quad \forall\,(x,y) \in O.$$

But this implies that

$$\mu_L(O) = \int_O W_L(x,y)\,\mathrm{d}(x,y) > 0,$$

i.e., the support of the Borel measure μ_L contains the open subset $O \subset \mathbb{R}^2$. Thus, in this case the convolution kernel K_L is positive definite according to Theorem 3.3.6. □

We close this chapter on the method of filtered back projection by studying the example of the Ram-Lak filter, of which an intuitive discussion can also be found in [121].

Example 3.3.8. *Consider the Ram-Lak filter from Example 3.2.8, whose window $W \in \mathrm{L}^\infty(\mathbb{R})$ is given by*

$$W(S) = \prod(S) = \begin{cases} 1 & \text{for } |S| \le 1 \\ 0 & \text{for } |S| > 1, \end{cases}$$

i.e.,

$$A_L(S) = |S| \cdot W(S/L) = \begin{cases} |S| & \text{for } |S| \le L \\ 0 & \text{for } |S| > L. \end{cases}$$

We first compute the corresponding band-limited function $q_L \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ given by

$$q_L(S,\theta) = \mathcal{F}^{-1}A_L(S) \quad \text{for } (S,\theta) \in \mathbb{R} \times [0,\pi).$$

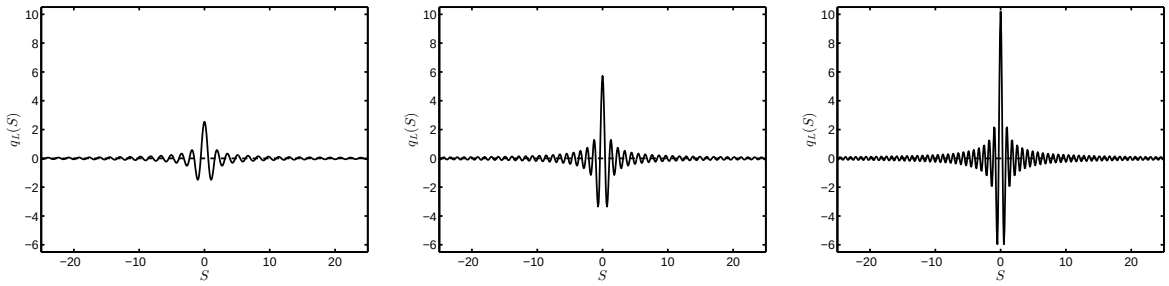
Since A_L is even and compactly supported with $\operatorname{supp}(A_L) \subseteq [-L,L]$, its inverse Fourier transform is given by the inverse cosine transform and for all $(S,\theta) \in \mathbb{R} \times [0,\pi)$ follows that

$$\begin{aligned} q_L(S,\theta) &= \frac{1}{2\pi}\int_{\mathbb{R}} A_L(t)\,\mathrm{e}^{\mathrm{i}St}\,\mathrm{d}t = \frac{1}{\pi}\int_0^L A_L(t)\,\cos(St)\,\mathrm{d}t = \frac{1}{\pi}\int_0^L t\,\cos(St)\,\mathrm{d}t \\ &= \frac{1}{\pi}\left[\frac{\cos(St) + St\sin(St)}{S^2}\right]_{t=0}^{t=L} = \frac{1}{\pi}\left(\frac{\cos(LS) + LS\sin(LS) - 1}{S^2}\right). \end{aligned}$$

Because $\cos(\varphi) = 1 - 2\sin^2(\varphi/2)$ for all $\varphi \in \mathbb{R}$, this can be rewritten as

$$q_L(S,\theta) = \frac{1}{\pi}\left(\frac{LS\sin(LS) - 2\sin^2(LS/2)}{S^2}\right) = \frac{L^2}{2\pi}\left(2\operatorname{sinc}(LS) - \operatorname{sinc}^2(LS/2)\right).$$

Plots of the radially symmetric function q_L can be found in Figure 3.4 for $L \in \{4,6,8\}$.



(a) $L = 4$ (b) $L = 6$ (c) $L = 8$

Figure 3.4: Band-limited function q_L of the Ram-Lak filter for $L \in \{4, 6, 8\}$.

We now come to the calculation of the associated convolution kernel $K_L \in \mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ *given by*

$$K_L(x, y) = \frac{1}{2}\,\mathcal{B}q_L(x, y) \quad \text{for } (x, y) \in \mathbb{R}^2.$$

By considering the radially symmetric bivariate window

$$W_L(x, y) = W\Big(\frac{r(x, y)}{L}\Big) = \begin{cases} 1 & \text{for } x^2 + y^2 \le L^2 \\ 0 & \text{for } x^2 + y^2 > L^2, \end{cases}$$

we can alternatively compute K_L *as the two-dimensional inverse Fourier transform of* W_L*, i.e., for fixed* $(x, y) \in \mathbb{R}^2$ *we have*

$$\begin{aligned} K_L(x, y) = \mathcal{F}^{-1}W_L(x, y) &= \frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} W_L(X, Y)\,\mathrm{e}^{\mathrm{i}(xX+yY)}\,\mathrm{d}X\,\mathrm{d}Y \\ &= \frac{1}{4\pi^2}\int_{r(x,y)\le L} \mathrm{e}^{\mathrm{i}(xX+yY)}\,\mathrm{d}(X, Y). \end{aligned}$$

Transforming from Cartesian coordinates $(X, Y) \in \mathbb{R}^2$ *to polar coordinates* $(S, \theta) \in \mathbb{R}_{\ge 0} \times [0, 2\pi)$ *yields*

$$\begin{aligned} K_L(x, y) &= \frac{1}{4\pi^2}\int_0^L \int_0^{2\pi} S\,\mathrm{e}^{\mathrm{i}S(x\cos(\theta)+y\sin(\theta))}\,\mathrm{d}\theta\,\mathrm{d}S \\ &= \frac{1}{4\pi^2}\int_0^L S\int_0^{2\pi} \cos(S(x\cos(\theta) + y\sin(\theta)))\,\mathrm{d}\theta\,\mathrm{d}S. \end{aligned}$$

Because the mapping $\theta \longmapsto \cos(S(x\cos(\theta) + y\sin(\theta)))$ *is* π*-periodic for all* $S \in \mathbb{R}$*, we further obtain*

$$K_L(x, y) = \frac{1}{2\pi^2}\int_0^L S\int_0^{\pi} \cos(S(x\cos(\theta) + y\sin(\theta)))\,\mathrm{d}\theta\,\mathrm{d}S.$$

Now, if $r(x, y) = 0$*, i.e.,* $x = 0$ *and* $y = 0$*, we have*

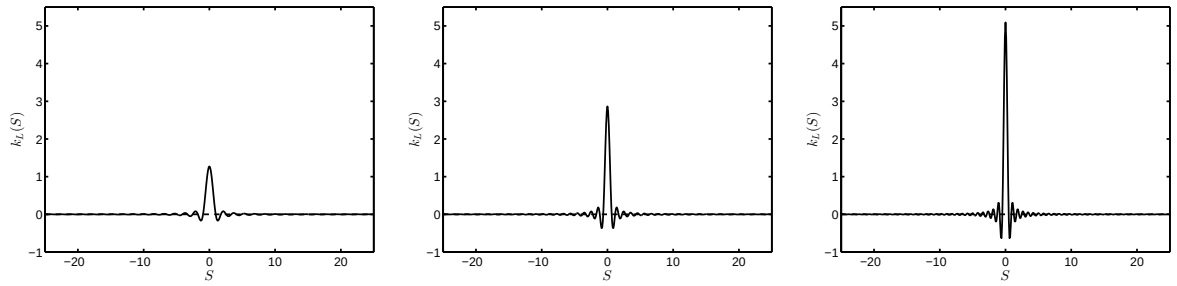
$$K_L(0, 0) = \frac{1}{2\pi^2}\int_0^L S\int_0^{\pi} 1\,\mathrm{d}\theta\,\mathrm{d}S = \frac{L^2}{4\pi}.$$

On the other hand, if $r(x, y) > 0$*, we apply the transformation* $t = S\,r(x, y)$ *to obtain*

$$K_L(x, y) = \frac{1}{2\pi^2}\int_0^{L\,r(x,y)} \frac{t}{r(x, y)^2}\int_0^{\pi} \cos\left(t\,\frac{x\cos(\theta) + y\sin(\theta)}{r(x, y)}\right)\mathrm{d}\theta\,\mathrm{d}t.$$

By defining the angle $\vartheta \equiv \vartheta(x, y) \in [-\pi, \pi)$ *as*

$$\vartheta = \begin{cases} \arctan(-\frac{y}{x}) & \text{for } x > 0 \\ -\frac{\pi}{2}\,\mathrm{sgn}(y) & \text{for } x = 0 \\ \arctan(-\frac{y}{x}) - \pi\,\mathrm{sgn}(y) & \text{for } x < 0, \end{cases}$$



(a) $L = 4$ (b) $L = 6$ (c) $L = 8$

Figure 3.5: Radial part of the convolution kernel K_L of the Ram-Lak filter for $L \in \{4, 6, 8\}$.

we observe that

$$\frac{x\cos(\theta) + y\sin(\theta)}{r(x,y)} = \cos(\theta + \vartheta) \quad \forall\, \theta \in [0, \pi).$$

Consequently, we get

$$\frac{1}{\pi}\int_0^\pi \cos\left(t\,\frac{x\cos(\theta) + y\sin(\theta)}{r(x,y)}\right) \mathrm{d}\theta = \frac{1}{\pi}\int_0^\pi \cos(t\cos(\theta+\vartheta))\,\mathrm{d}\theta = \frac{1}{\pi}\int_\vartheta^{\pi+\vartheta} \cos(t\cos(\varphi))\,\mathrm{d}\varphi$$

and the π-periodicity of the mapping $\varphi \longmapsto \cos(t\cos(\varphi))$ for fixed $t \in \mathbb{R}$ yields

$$\frac{1}{\pi}\int_0^\pi \cos\left(t\,\frac{x\cos(\theta) + y\sin(\theta)}{r(x,y)}\right) \mathrm{d}\theta = \frac{1}{\pi}\int_0^\pi \cos(t\cos(\varphi))\,\mathrm{d}\varphi = J_0(t),$$

where J_0 denotes the Bessel function of the first kind of zero order given by

$$J_0(z) = \frac{1}{\pi}\int_0^\pi \cos(z\cos(\theta))\,\mathrm{d}\theta \quad \text{for } z \in \mathbb{R}.$$

In general, the Bessel function J_k of the first kind of integer order $k \in \mathbb{Z}$ is defined as

$$J_k(z) = \frac{1}{\pi}\int_0^\pi \cos(z\sin(\theta) - k\theta)\,\mathrm{d}\theta \quad \text{for } z \in \mathbb{R},$$

see [1, Chapter 9], and satisfies the reduction formula [1, Formula 11.3.20]

$$\int_0^z t^k\, J_{k-1}(t)\,\mathrm{d}t = z^k\, J_k(z) \quad \forall\, z > 0,\, k > 0.$$

Consequently, for the convolution kernel K_L follows that

$$K_L(x,y) = \frac{1}{2\pi}\int_0^{L\,r(x,y)} \frac{t\,J_0(t)}{r(x,y)^2}\,\mathrm{d}t = \frac{L^2}{2\pi}\,\frac{J_1(L\,r(x,y))}{L\,r(x,y)} \quad \forall\,(x,y) \in \mathbb{R}^2.$$

By defining the univariate function $k : \mathbb{R} \longrightarrow \mathbb{R}$ via

$$k(S) = \frac{1}{2\pi}\,\frac{J_1(S)}{S} \quad \text{for } S \in \mathbb{R},$$

we can finally rewrite K_L as

$$K_L(x,y) = L^2\, k(r(L\,x, L\,y)) \quad \forall\,(x,y) \in \mathbb{R}^2$$

and, thus, we indeed observe its radial symmetry and its scaling property with respect to $L > 0$. Plots of the radial part of the convolution kernel K_L can be found in Figure 3.5 for $L \in \{4, 6, 8\}$.

Chapter 4

Error estimates for filtered back projection reconstructions

In Chapter 3 we have explained how the filtered back projection (FBP) formula (2.9) can be stabilized by incorporating a low-pass filter $A_L : \mathbb{R} \longrightarrow \mathbb{R}$ of the form

$$A_L(S) = |S|\, W_L(S) = |S|\, W(S/L) \quad \text{for } S \in \mathbb{R}$$

with finite bandwidth $L > 0$ and an even window function $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\operatorname{supp}(W) \subseteq [-1, 1]$. This reduces the noise sensitivity of the FBP formula, but only leads to an inexact FBP method. For target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$, we have seen that the resulting approximate FBP reconstruction f_L satisfies $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ and can be expressed in standard form (3.1), i.e.,

$$f_L = \frac{1}{2}\, \mathcal{B}(\mathcal{F}^{-1} A_L * \mathcal{R} f) = f * K_L.$$

Therefore, each FBP method provides one approximation f_L to f, whose quality depends on the choice of the low-pass filter A_L.

In this chapter, we analyse the inherent reconstruction error of the FBP method which is incurred by the application of the low-pass filter A_L, i.e., we wish to analyse the error

$$e_L = f - f_L$$

with respect to the filter's bandwidth L and window function W. To this end, we prove error estimates on e_L for target functions f from Sobolev spaces $\mathrm{H}^\alpha(\mathbb{R}^2)$ of *fractional* order $\alpha > 0$, where

$$\mathrm{H}^\alpha(\mathbb{R}^2) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^2) \mid \|f\|_\alpha < \infty \right\}$$

is equipped with the Sobolev norm

$$\|f\|_\alpha = \left(\frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} (1 + x^2 + y^2)^\alpha \, |\mathcal{F} f(x, y)|^2 \, \mathrm{d}x \, \mathrm{d}y \right)^{1/2}.$$

To be more precise, in the following we analyse the H^σ-norm of the intrinsic FBP reconstruction error e_L for all $0 \leq \sigma \leq \alpha$, where $\alpha > 0$ is the assumed smoothness of the target function $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$ and $\sigma = 0$ yields the classical L^2-case. The obtained error bounds depend on the bandwidth L of the utilized low-pass filter, on the flatness of the filter's window function W at the origin, on the smoothness α of the target function f, and on the order σ of the considered Sobolev norm. Finally, we prove convergence for the approximate FBP reconstruction f_L in the treated Sobolev norms along with asymptotic convergence rates as the bandwidth L goes to infinity. In particular, we show that the rate of convergence saturates at fractional order depending on smoothness properties of the chosen window function.

We remark that some parts of the presented material are already published in the conference proceedings [7], [10], [12] or can be found in the preprints [8], [9], [11].

4.1 Related results

Before we start with the derivation of our error analysis of the filtered back projection method, we describe related techniques and results, which can be found in the literature, and explain the differences to our approach. To this end, if suitable, we adapt the notations to our setting and restrict the results to the two dimensional case, although some references treat arbitrary dimensions.

4.1.1 Summability methods

In [71], Madych describes the reconstruction of functions from Radon data based on summability formulas. The basic idea is to choose a convolution kernel $K : \mathbb{R}^2 \longrightarrow \mathbb{R}$ as an approximation of the identity and to compute the convolution product $f * K_L$ to approximate the target function f. Here, for $L > 0$, the scaled kernel K_L is given by

$$K_L(x, y) = L^2 \, K(L\,x, L\,y) \quad \text{for } (x, y) \in \mathbb{R}^2.$$

If the kernel K is chosen to be a uniform sum of ridge functions, the convolution $f * K_L$ can be expressed in terms of the Radon data $\mathcal{R}f$ as in the approximate FBP formula (3.4), which is proven in [71, Proposition 1]. Further, convolution kernels K that can be represented as uniform sums of ridge functions are characterized in [71, Section 2.2].

For target functions $f \in \mathrm{L}^p(\mathbb{R}^2)$, $1 \le p \le \infty$, the reconstruction error

$$f - f * K_L$$

is then estimated in terms of the L^p-modulus of continuity $\omega_p(f; \delta)$, where, for $\delta > 0$,

$$\omega_p(f; \delta) = \sup_{\|(X,Y)\|_{\mathbb{R}^2} \le \delta} \left(\int_{\mathbb{R}} \int_{\mathbb{R}} |f(x - X, y - Y) - f(x, y)|^p \, \mathrm{d}x \, \mathrm{d}y \right)^{1/p} \quad \text{for } 1 \le p < \infty$$

and

$$\omega_\infty(f; \delta) = \sup_{\|(X,Y)\|_{\mathbb{R}^2} \le \delta} \operatorname*{ess\,sup}_{(x,y) \in \mathbb{R}^2} |f(x - X, y - Y) - f(x, y)|.$$

Under the assumption that K_L, for $L > 0$, forms a family of integrable convolution kernels satisfying

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} K_L(x, y) \, \mathrm{d}x \, \mathrm{d}y &= 1, \\ \int_{\mathbb{R}} \int_{\mathbb{R}} |K_L(x, y)| \, \mathrm{d}x \, \mathrm{d}y &\le c_0, \\ \int_{\mathbb{R}} \int_{\mathbb{R}} (x^2 + y^2)^{1/2} \, |K_L(x, y)| \, \mathrm{d}x \, \mathrm{d}y &\le c_1 \, L^{-1} \end{aligned}$$

for some constants $c_0, c_1 \in \mathbb{R}_{\ge 0}$ independent of L, it is shown in [71, Proposition 7] that

$$\|f - f * K_L\|_{\mathrm{L}^p(\mathbb{R}^2)} \le c \, \omega_p(f; L^{-1}), \tag{4.1}$$

where the constant $c \in \mathbb{R}_{\ge 0}$ is independent of f and L. The proof is based on direct calculations in the case $p \in \{1, \infty\}$, and on the integral variant of Minkowski's inequality for $1 < p < \infty$.

To exploit a higher order moment condition on the convolution kernel K_L of the form

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x^2 + y^2)^{k/2} \, |K_L(x, y)| \, \mathrm{d}x \, \mathrm{d}y \le c_k \, L^{-k}$$

for some integer $k \ge 2$ and a constant $c_k \in \mathbb{R}_{\ge 0}$ independent of L, the modified kernels

$$\tilde{K}_L^k(x, y) = \sum_{j=0}^{k-1} (-1)^{k-j-1} \, \frac{k!}{(k-j)!\, j!} \, (k - j)^{-2} \, K_L\Big(\frac{x}{k-j}, \frac{y}{k-j}\Big) \quad \text{for } (x, y) \in \mathbb{R}^2$$

are introduced and the corresponding reconstruction error

$$f - f * \tilde{K}_L^k$$

is estimated in terms of the k-th order L^p-modulus of smoothness $\omega_p^k(f;\delta)$ of f via

$$\|f - f * \tilde{K}_L^k\|_{\mathrm{L}^p(\mathbb{R}^2)} \leq c\,\omega_p^k(f; L^{-1}), \tag{4.2}$$

where the constant $c \in \mathbb{R}_{\geq 0}$ is again independent of f and L, see [71, Proposition 8].

The constant c in the estimates (4.1) and (4.2) depends on the L^1-norm of the convolution kernel K_L so that the assumption $K_L \in \mathrm{L}^1(\mathbb{R}^2)$ is essential and cannot be omitted. However, the integrability of K_L implies that its Fourier transform $\mathcal{F}K_L$ is continuous on $\mathbb{R}^2$ according to the Riemann-Lebesgue Lemma A.1.2. In our setting this would mean that the univariate window function $W \in \mathrm{L}^\infty(\mathbb{R})$ has to be continuous on $\mathbb{R}$ due to Theorem 3.2.3. As opposed to this, we assume that W has compact support with $\mathrm{supp}(W) \subseteq [-1,1]$ and we want to allow discontinuities of the window at the boundary points of its support. Note that this is the case in most of the examples we consider in this thesis.

Consequently, the assumptions in [71] are not compatible with our setting and, in particular, the approach taken here is essentially different from that in [71].

In [72], Madych considers two particular choices of the convolution kernel K_L, where the first one yields a natural approximation of Radon's classical reconstruction formula from [94] and where the second one leads to an approximation of an alternative inversion formula derived in [71, Corollary 2]. For these two choices of K_L and for target functions $f \in \mathrm{L}^\infty(\mathbb{R}^2)$ that are Hölder continuous of order $\alpha > 0$ at $(x,y) \in \mathbb{R}^2$ the pointwise reconstruction error

$$f(x,y) - (f * K_L)(x,y)$$

is estimated in terms of the parameter L of the scaled kernel K_L and the Hölder exponent α of the target function f. The proof is again substantially based on the integrability of K_L and, therefore, the estimates in [72] are also not applicable to the setting of this work.

4.1.2 Approximate inverse

The method of approximate inverse was developed by Louis and Maass in [65] to solve ill-posed linear operator equations

$$Af = g,$$

where $A : \mathcal{X} \longrightarrow \mathcal{Y}$ is a continuous linear operator between Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$, the data $g \in \mathcal{Y}$ denotes a collection of input measurements and $f \in \mathcal{X}$ is the quantity we search for. Note that in the context of this thesis the operator A is given by the Radon transform $\mathcal{R}$.

Now, the basic idea of approximate inverse is to select a smoothing operator $E_\gamma : \mathcal{X} \longrightarrow \mathcal{X}$, for $\gamma > 0$, and to compute a smoothed version

$$f_\gamma = E_\gamma f$$

of the target function f. If $\mathcal{X}$ is a space of real-valued functions on a domain Ω, this is done by calculating the moments

$$f_\gamma(x) = (f, e_\gamma(x,\cdot))_{\mathcal{X}} \quad \text{for } x \in \Omega$$

with a suitable family of mollifiers $e_\gamma : \Omega \times \Omega \longrightarrow \mathbb{R}$ satisfying

$$\lim_{\gamma \to 0} \|f - f_\gamma\|_{\mathcal{X}} = 0.$$

The computation of f_γ from the given data $g \in \mathcal{Y}$ is achieved by approximating $e_\gamma(x,\cdot)$ in the range of the adjoint operator A^* by the *reconstruction kernel* $v_\gamma(x) \in \mathcal{Y}$ solving

$$\min_{v \in \mathcal{Y}} \|A^* v - e_\gamma(x,\cdot)\|_{\mathcal{X}}$$

so that

$$f_\gamma(x) = (f, e_\gamma(x,\cdot))_{\mathcal{X}} \approx (g, v_\gamma(x))_{\mathcal{Y}} \quad \text{for } x \in \Omega.$$

The mapping $S_\gamma : \mathcal{Y} \longrightarrow \mathcal{X}$, defined as

$$S_\gamma g(x) = (g, v_\gamma(x))_{\mathcal{Y}} \quad \text{for } x \in \Omega,$$

is then called the *approximate inverse* of the operator A. Applied to the Radon transform this yields a reconstruction formula of filtered back projection type similar to (3.1). For detailed investigations on properties of the approximate inverse and its relation to other regularization methods we refer to the papers [60], [61] of Louis.

In [49], Jonas and Louis consider the case where $\mathcal{X}$ and $\mathcal{Y}$ are L^2-spaces, possibly on different domains of n dimensions, and A is an operator with smoothing index $\alpha > 0$ in the Sobolev scale, i.e., there exist constants $c_1, c_2 > 0$ such that

$$c_1 \|f\|_{\mathrm{L}^2} \leq \|Af\|_{\mathrm{H}^\alpha} \leq c_2 \|f\|_{\mathrm{L}^2} \quad \forall\, f \in \mathcal{N}(A)^\perp,$$

where $\mathcal{N}(A)$ denotes the null space of A. For mollifiers e_γ of convolution type,

$$e_\gamma(x,y) = e_\gamma(x-y) \quad \forall\, x, y \in \Omega,$$

sufficient conditions are then derived under which the approximate inverse S_γ yields a regularization method of optimal order, cf. [49, Theorems 3.2 & 5.3]. The proofs are based on similar techniques we use in Section 4.2 for the derivation of our Sobolev error estimate. Furthermore, in the proof of [49, Theorem 5.3], the estimate

$$\|f - f_\gamma\|_{\mathrm{L}^2} \leq c_\beta\, \gamma^{\theta\frac{\beta}{\alpha}} \|f\|_{\mathrm{H}^\beta} \quad \forall\, f \in \mathrm{H}^\beta \tag{4.3}$$

is established for all $0 < \beta \leq \beta^*$ with some constant $c_\beta > 0$. To this end, it is assumed that there are positive constants $\theta, \beta^*, c_{\beta^*} > 0$ such that

$$\sup_\xi \left\{ \left(1 + \|\xi\|_2^2\right)^{-\beta^*/2} \left|(2\pi)^{n/2}\, \mathcal{F}e_\gamma(\xi) - 1\right| \right\} \leq c_{\beta^*}\, \gamma^{\theta\frac{\beta^*}{\alpha}} \quad \forall\, \gamma > 0. \tag{4.4}$$

However, in [49] the authors verify assumption (4.4) only for one example with $n = 1$, where the mollifier e_γ is given by a sinc function. In contrast to that, we aim to develop concrete and easy-to-check conditions on the window function W which guarantee that the inherent FBP reconstruction error $e_L = f - f_L$ behaves in the fashion of estimate (4.3). Moreover, our estimates in Section 4.3 allow for nontrivial statements concerning the behaviour of the reconstruction error in the case $\beta > \beta^*$.

In [68], Louis and Schuster apply the method of approximate inverse to the Radon transform $\mathcal{R}$ to derive inversion formulas for the parallel beam geometry in computerized tomography. They consider both the continuous and discrete setting and explain how to compute the reconstruction kernel for a chosen mollifier. This again leads to inversion formulas of filtered back projection type similar to (3.1), where, in the case of finitely many data, the approach in [68] relies on a truncation of the singular value decomposition of the Radon transform. However, note that [68] contains no results concerning error estimates or convergence rates.

In [102], [103], Rieder and Schuster focus on semi-discrete systems

$$A_n f = g_n,$$

where the semi-discrete operator $A_n : \mathcal{X} \longrightarrow \mathbb{C}^n$ and the measurements $g_n \in \mathbb{C}^n$ are defined via an observation operator $\Psi_n : \mathcal{Y} \longrightarrow \mathbb{C}^n$ by

$$A_n = \Psi_n A \quad \text{and} \quad g_n = \Psi_n g.$$

The authors propose and analyse a technique for approximating the discrete reconstruction kernel for a given mollifier and show convergence of the resulting discrete version of the approximate inverse. Further, they apply their results to the Radon transform and derive convergence rates for the discrete filtered back projection algorithm in parallel beam geometry as the discretization parameters go to zero. Concrete examples of mollifier/reconstruction kernel pairs for the Radon transform are given in [98].

Since the approach in [102], [103] considers the semi-discrete setting, the method parameter $\gamma > 0$ is coupled with the discretization parameters. Therefore, the intrinsic approximation error

$$f - f_\gamma$$

for the continuous approximate inverse reconstruction f_γ of the target function f from complete Radon data is not considered explicitly. Consequently, the results of this work are not covered by the theory of Rieder and Schuster. In addition, in [102], [103] the mollifier is assumed to have compact support, whereas we choose the window function W to be compactly supported with $\operatorname{supp}(W) \subseteq [-1, 1]$. Due to Theorem 3.2.3 and the Paley-Wiener Theorem A.1.14 this implies that the convolution kernel K_L is not compactly supported in the setting of this thesis.

Note that the results in [103] lead to suboptimal convergence rates for the discrete filtered back projection algorithm, as explained in [103] (cf. the paragraph following [103, Corollary 5.6]). As opposed to this, in [100] Rieder and Faridani prove optimal L^2-convergence rates for a semi-discrete filtered back projection algorithm in parallel beam geometry, where no discretization of the back projection operator $\mathcal{B}$ is considered. This is incorporated by Rieder and Schneck in [101] leading to optimal L^2-convergence rates for a fully discrete version of the filtered back projection algorithm, where a sufficiently smooth target function is required.

The proof is based on a new representation of a discretized version of the approximate FBP formula (3.4) utilizing generalized interpolation operators so that techniques from approximation theory can be applied. Based on this, the authors derive asymptotic error estimates with generic constants, that are not given explicitly. Here, the range of Sobolev orders for the assumed smoothness of the target function yielding optimal convergence rates depends on the chosen filter function and interpolation procedure. However, the authors verify the assumptions of their theory only for reconstruction filters which are based on (orthogonalized) B-splines.

We remark that the resulting representation of the discretized approximate FBP formula depends on the utilized filter function, the interpolation procedure and the discretization parameters. In particular, the inherent FBP reconstruction error

$$e_L = f - f_L,$$

which is incurred by a low-pass filter of finite bandwidth L and which we aim to analyse, is not estimated explicitly in [100], [101].

4.1.3 Other approaches

In [97], in the context of the numerical approximation of solutions of partial differential equations by particle methods, Raviart analyses the reconstruction error

$$f - f * K_L$$

for target function f from Sobolev spaces of integer order, i.e., $f \in \mathrm{H}^{m,p}(\mathbb{R}^2)$ for some $m \in \mathbb{N}_0$ and $1 \leq p \leq \infty$, where

$$\mathrm{H}^{m,p}(\mathbb{R}^2) = \left\{ f \in \mathrm{L}^p(\mathbb{R}^2) \mid \|f\|_{m,p} < \infty \right\}$$

with the norm

$$\|f\|_{m,p} = \begin{cases} \left(\sum_{\alpha+\beta \leq m} \left\| \frac{\partial^\alpha}{\partial x^\alpha} \frac{\partial^\beta}{\partial y^\beta} f \right\|_{\mathrm{L}^p(\mathbb{R}^2)}^p \right)^{1/p} & \text{for } 1 \leq p < \infty \\ \max_{\alpha+\beta \leq m} \left\| \frac{\partial^\alpha}{\partial x^\alpha} \frac{\partial^\beta}{\partial y^\beta} f \right\|_{\mathrm{L}^\infty(\mathbb{R}^2)} & \text{for } p = \infty. \end{cases}$$

Furthermore, the convolution kernel K is required to satisfy $K \in \mathscr{C}(\mathbb{R}^2) \cap \mathrm{L}^1(\mathbb{R}^2)$ and, for $L > 0$, the scaled kernel K_L is again defined as

$$K_L(x, y) = L^2 \, K(L\,x, L\,y) \quad \text{for } (x, y) \in \mathbb{R}^2.$$

In addition, it is assumed that there exists an integer $k \geq 1$ such that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} K(x, y) \, \mathrm{d}x \, \mathrm{d}y = 1,$$

$$\int_{\mathbb{R}} \int_{\mathbb{R}} x^\alpha \, y^\beta \, K(x, y) \, \mathrm{d}x \, \mathrm{d}y = 0 \quad \forall\, \alpha, \beta \in \mathbb{N}_0 : \; 1 \leq \alpha + \beta \leq k - 1,$$

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x^2 + y^2)^{k/2} \, |K(x, y)| \, \mathrm{d}x \, \mathrm{d}y < \infty.$$

For target functions $f \in \mathrm{H}^{k,p}(\mathbb{R}^2)$, with $1 \leq p \leq \infty$, [97, Lemma I.4.4] then yields error estimates of the form

$$\| f - f * K_L \|_{\mathrm{L}^p(\mathbb{R}^2)} \leq C \, L^{-k} \, |f|_{k,p} \tag{4.5}$$

with some constant $C > 0$ and

$$|f|_{k,p} = \begin{cases} \left(\sum_{\alpha+\beta=k} \left\| \frac{\partial^\alpha}{\partial x^\alpha} \frac{\partial^\beta}{\partial y^\beta} f \right\|_{\mathrm{L}^p(\mathbb{R}^2)}^p \right)^{1/p} & \text{for } 1 \leq p < \infty \\ \max_{\alpha+\beta=k} \left\| \frac{\partial^\alpha}{\partial x^\alpha} \frac{\partial^\beta}{\partial y^\beta} f \right\|_{\mathrm{L}^\infty(\mathbb{R}^2)} & \text{for } p = \infty. \end{cases}$$

We remark that the proof of (4.5) is based on a Taylor expansion of the target function f and the vanishing moment conditions on K. Further, the required differentiability order of f is coupled with the order k of the moment conditions posed on K and the constant $C > 0$ in (4.5) is not given explicitly. However, the k-th order moment condition

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x^2 + y^2)^{k/2} \, |K(x, y)| \, \mathrm{d}x \, \mathrm{d}y < \infty$$

in combination with the integrability of the kernel K implies that its Fourier transform $\mathcal{F}K$ is k-times continuously differentiable on $\mathbb{R}^2$. In the setting of this thesis this would mean that the univariate window function $W \in \mathrm{L}^\infty(\mathbb{R})$ has to be k-times continuously differentiable on $\mathbb{R}$ due to Theorem 3.2.3. Thus, as explained before, the assumptions in [97] are not compatible with our setting.

In Section 4.3 we prove error estimates on e_L for target functions from fractional Sobolev spaces, where we assume that the window function W is k-times continuously differentiable only on the interval $[-1, 1]$ where it is supported and allow the window to jump to zero at the boundary of its support. Moreover, the smoothness $\alpha > 0$ of the target function f is not coupled with the differentiability order k of W and the constants appearing in the estimates are given explicitly. Finally, we observe saturation of the decay rate of the error bounds at order k. However, in the saturation case the constant in the error bounds is strictly monotonically decreasing in $\alpha > k$ such that a smoother target function still allows for a better approximation. This behaviour cannot be observed based on the estimates proven in [97].

In [110], Schomburg analyses the convergence rates of certain delta sequences in Sobolev spaces of negative fractional order, cf. Appendix A.3. To be more precise, for a tempered distribution $\phi \in \mathrm{H}^{-\alpha}(\mathbb{R}^2)$ with $\alpha > 1$, which satisfies further assumptions specified in [110, Theorem 1], asymptotic estimates of the error

$$\phi_n - \delta$$

are derived in the $\mathrm{H}^{-\alpha}$-norm. Here, $\delta \in \mathrm{H}^{-\alpha}(\mathbb{R}^2)$ denotes the Dirac delta distribution, given by

$$\langle \delta, \psi \rangle = \psi(0) \quad \text{for } \psi \in \mathcal{S}(\mathbb{R}^2)$$

with the duality pairing $\langle \cdot, \cdot \rangle$ on $\mathrm{H}^{-\alpha}(\mathbb{R}^2) \times \mathrm{H}^{\alpha}(\mathbb{R}^2)$, and the sequence $(\phi_n)_{n\in\mathbb{N}}$ in $\mathrm{H}^{-\alpha}(\mathbb{R}^2)$ is defined via

$$\langle \phi_n, \psi \rangle = \langle \phi, \psi(\cdot/n, \cdot/n) \rangle \quad \text{for } \psi \in \mathcal{S}(\mathbb{R}^2).$$

Based on the observation that for an even convolution kernel $K \in \mathscr{C}(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ the scaled kernels K_L with $L > 0$, given by

$$K_L(x, y) = L^2 \, K(L\,x, L\,y) \quad \text{for } (x, y) \in \mathbb{R}^2,$$

can be considered as tempered distributions in $\mathrm{H}^{-\alpha}(\mathbb{R}^2)$ with

$$\langle K_L, f \rangle = \int_{\mathbb{R}} \int_{\mathbb{R}} K_L(X, Y) \, f(X, Y) \, \mathrm{d}X \, \mathrm{d}Y = \int_{\mathbb{R}} \int_{\mathbb{R}} K(X, Y) \, f(X/L, Y/L) \, \mathrm{d}X \, \mathrm{d}Y = \langle K, f(\cdot/L, \cdot/L) \rangle,$$

the results from [110] can be used to prove asymptotic pointwise error estimates on

$$f - f * K_L$$

for target functions $f \in \mathrm{H}^{\alpha}(\mathbb{R}^2)$ with $\alpha > 1$ in terms of the H^{α}-norm of f. Indeed, for fixed $(x, y) \in \mathbb{R}^2$, we have

$$\langle K_L(x - \cdot, y - \cdot), f \rangle = \int_{\mathbb{R}} \int_{\mathbb{R}} K_L(x - X, y - Y) \, f(X, Y) \, \mathrm{d}X \, \mathrm{d}Y = (f * K_L)(x, y)$$

and

$$\| K_L(x - \cdot, y - \cdot) - \delta(x - \cdot, y - \cdot) \|_{\mathrm{H}^{-\alpha}(\mathbb{R}^2)} = \| K_L - \delta \|_{\mathrm{H}^{-\alpha}(\mathbb{R}^2)}$$

so that

$$|f(x, y) - (f * K_L)(x, y)| \leq \| K_L - \delta \|_{\mathrm{H}^{-\alpha}(\mathbb{R}^2)} \, \| f \|_{\mathrm{H}^{\alpha}(\mathbb{R}^2)} \quad \forall \, (x, y) \in \mathbb{R}^2.$$

Note that the constants in the asymptotic estimates of [110] are generic and not given explicitly. Moreover, we are interested in error estimates on $f - f * K_L$ for target functions $f \in \mathrm{H}^{\alpha}(\mathbb{R}^2)$, where the smoothness α of f is only assumed to be positive. Especially the case $0 < \alpha \leq 1$ is of particular interest, as explained in [81], so that the assumption $\alpha > 1$ is too restrictive for the setting of this work.

We finally remark that pointwise and L^{∞}-error estimates on $e_L = f - f_L$, along with asymptotic pointwise error formulas, are discussed by Munshi in [76] and by Munshi et al. in [78]. Their theoretical results were further supported by numerical experiments in [22], [77], [79], [80], [122]. In their derivations, the authors assume certain moment conditions on the convolution kernel K and differentiability of the target function f in a strict sense, what we can avoid in this thesis. To illustrate this, we present the following special case, which is based on the results from [76].

Under the assumption that $f \in \mathscr{C}_b^2(\mathbb{R}^2)$ is bounded and twice continuously differentiable on $\mathbb{R}^2$, that the window function W is twice continuously differentiable in a neighbourhood around 0 and that the convolution kernel K_L satisfies $K_L \in \mathrm{L}^1(\mathbb{R}^2)$ with

$$\int_{\mathbb{R}} \int_{\mathbb{R}} K_L(x, y) \, \mathrm{d}x \, \mathrm{d}y = 1$$

and

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x^2 + y^2) \, |K_L(x, y)| \, \mathrm{d}x \, \mathrm{d}y < \infty,$$

the asymptotic pointwise error formula

$$e_L(x, y) = \frac{1}{2} \, L^{-2} \, \Delta f(x, y) \, W''(0) + o(L^{-2}) \quad \forall \, (x, y) \in \mathbb{R}^2 \tag{4.6}$$

can be established, where Δf denotes the Laplacian of f, i.e.

$$\Delta f = \frac{\partial^2}{\partial x^2} f + \frac{\partial^2}{\partial y^2} f.$$

The basic idea for (4.6) comes from Munshi in [76], but has been adopted to our setting in [6]. In this thesis, however, we avoid assuming classical differentiability of the target function f and pose our assumptions directly on the window function W instead of the convolution kernel K_L.

4.2 Error analysis in fractional Sobolev spaces

In this section we now analyse the reconstruction error of the approximate FBP formula (3.1), i.e.,

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f) = f * K_L \in \mathrm{L}^2(\mathbb{R}^2),$$

for the reconstruction of an unknown target function $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ from the knowledge of its Radon transform

$$\mathcal{R}f(t,\theta) = \int_{\ell_{t,\theta}} f(x,y)\,\mathrm{d}(x,y) \quad \text{for } (t,\theta) \in \mathbb{R}\times[0,\pi).$$

Note that in practical applications, however, the Radon data $g = \mathcal{R}f \in \mathrm{L}^2(\mathbb{R}\times[0,\pi))$ is usually not known precisely, but only up to a noise level $\delta > 0$, and we have to reconstruct f from given noisy data $g^\delta \in \mathrm{L}^2(\mathbb{R}\times[0,\pi))$ with

$$\|g - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \le \delta.$$

Applying the FBP method (3.4) to the noisy measurements g^δ yields the reconstruction

$$f_L^\delta = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * g^\delta)$$

and, using standard concepts from the theory of inverse problems and regularization (cf. [29]), we observe that the overall FBP reconstruction error

$$e_L^\delta = f - f_L^\delta$$

can be split into an *approximation error* term and a *data error* term,

$$e_L^\delta = \underbrace{f - f_L}_{\substack{\text{approximation}\\ \text{error}}} + \underbrace{f_L - f_L^\delta}_{\substack{\text{data}\\ \text{error}}}.$$

In this thesis, we focus on the approximation error term, i.e., we analyse the inherent reconstruction error of the FBP method which is incurred by the application of the low-pass filter A_L. More precisely, we wish to analyse the inherent FBP reconstruction error

$$e_L = f - f_L \tag{4.7}$$

with respect to the filter's window function W and bandwidth L.

To this end, we prove error estimates on e_L for target functions f from Sobolev spaces $\mathrm{H}^\alpha(\mathbb{R}^2)$ of fractional order $\alpha > 0$. Let us recall that the *Sobolev space* $\mathrm{H}^\alpha(\mathbb{R}^2)$ of order $\alpha > 0$ is defined as

$$\mathrm{H}^\alpha(\mathbb{R}^2) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^2) \mid \|f\|_\alpha < \infty \right\},$$

where

$$\|f\|_\alpha^2 = \frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} (1 + x^2 + y^2)^\alpha\, |\mathcal{F}f(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y.$$

In relevant applications of (medical) image processing, Sobolev spaces of compactly supported functions,

$$\mathrm{H}_0^\alpha(\Omega) = \left\{ f \in \mathrm{H}^\alpha(\mathbb{R}^2) \mid \mathrm{supp}(f) \subseteq \overline{\Omega} \right\},$$

on an open and bounded domain $\Omega \subset \mathbb{R}^2$, and of fractional order $\alpha > 0$ play an important role (cf. [81]). In fact, the density function of an image in $\Omega \subset \mathbb{R}^2$ has usually jumps along smooth curves, but is smooth off these curve singularities. Such functions belong to any Sobolev space $\mathrm{H}_0^\alpha(\Omega)$ with $\alpha < \frac{1}{2}$. Consequently, we can consider the density of an image as a function in a Sobolev space $\mathrm{H}_0^\alpha(\Omega)$ whose order α is close to $\frac{1}{2}$.

4.2.1 L^2-error estimates

We start with proving L^2-error estimates for the intrinsic FBP reconstruction error e_L in (4.7), where our upper bound on the L^2-norm of e_L will be split into one error term depending on the filter's window function W and another one depending on its bandwidth $L > 0$.

Let us first discuss the special case of the Ram-Lak filter from Example 3.2.8, i.e.,

$$A_L(S) = \begin{cases} |S| & \text{for } |S| \leq L \\ 0 & \text{for } |S| > L. \end{cases}$$

Note that the Ram-Lak filter's window function W is given by the characteristic function $\chi_{[-1,1]}$ of the interval $[-1, 1]$, so that the scaled window W_L of bandwidth $L > 0$ is the characteristic function $\chi_{[-L,L]}$ of the interval $[-L, L]$. Based on this observation, we see that the reconstruction error $e_L = f - f_L$ vanishes identically, $e_L \equiv 0$, for target functions f with band-limited Radon transform $\mathcal{R}f$, provided that the filter's bandwidth L is at least as large as the largest frequency contained in $\mathcal{R}f$. Indeed, in this case the approximate Ram-Lak FBP method (3.4) coincides with the exact FBP formula (2.9), so that $f \equiv f_L$.

Yet it remains to discuss how reasonable it is to assume that the target function f has a band-limited Radon transform $\mathcal{R}f$. To further elaborate this, let us recall the classical Fourier slice Theorem 2.2.11, which states that for any $f \in \mathrm{L}^1(\mathbb{R}^2)$ the identity

$$\mathcal{F}(\mathcal{R}f)(S, \theta) = \mathcal{F}f(S\cos(\theta), S\sin(\theta))$$

holds for all $(S, \theta) \in \mathbb{R} \times [0, \pi)$. Hence, the (univariate) Fourier transform of $\mathcal{R}f$ is entirely determined by the (bivariate) Fourier transform of f, and vice versa.

Now, $\mathcal{R}f$ is band-limited if and only if $\mathcal{R}f$ has a compactly supported Fourier transform $\mathcal{F}(\mathcal{R}f)$. However, in applications of medical image reconstruction, it is usually assumed that the target function $f \not\equiv 0$ is compactly supported. But in this case, its Fourier transform $\mathcal{F}f$ is analytic, due to the Paley-Wiener Theorem A.1.14, and so is $\mathcal{F}(\mathcal{R}f)$ an analytic function, according to the Fourier slice theorem. Hence, $\mathcal{F}(\mathcal{R}f)$ cannot have compact support, i.e., $\mathcal{R}f$ cannot be band-limited.

To conclude our discussion on the special case of the Ram-Lak filter, we see that, for compactly supported $f \not\equiv 0$, the error e_L of the Ram-Lak FBP method cannot be zero for finite bandwidth $L > 0$. But if we let L go to infinity, the Ram-Lak FBP method (3.4) will coincide, in the limit, with the exact, but numerically unstable filtered back projection formula (2.9), i.e., $e_\infty \equiv 0$. In other words, any admissible target function f can be approximated arbitrarily well by f_L for sufficiently large L.

Let us now turn to the analysis of the reconstruction error $e_L = f - f_L$ in $\mathrm{L}^2(\mathbb{R}^2)$ for general low-pass filters. To start with, we consider $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$. Further, throughout this section we assume that the low-pass filter's window $W \in \mathrm{L}^\infty(\mathbb{R})$ is even and compactly supported with $\operatorname{supp}(W) \subseteq [-1, 1]$. Due to Theorem 3.2.5 we then have $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ for the approximate FBP reconstruction, along with the representation (3.6), i.e.,

$$f_L = f * K_L.$$

But this immediately implies that

$$\|e_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \|f - f_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \|f - f * K_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \frac{1}{4\pi^2} \|\mathcal{F}f - \mathcal{F}f \cdot \mathcal{F}K_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)}$$

by applying the Fourier convolution Theorem 3.1.1 and the Rayleigh-Plancherel Theorem A.1.12. Moreover, by Theorem 3.2.3, the Fourier transform $\mathcal{F}K_L$ of the convolution kernel K_L is given by the bivariate window function W_L in (3.7), i.e., $\mathcal{F}K_L = W_L$, so that

$$\|e_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \frac{1}{4\pi^2} \|\mathcal{F}f - W_L \cdot \mathcal{F}f\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} |(\mathcal{F}f - W_L \cdot \mathcal{F}f)(x, y)|^2 \, \mathrm{d}x \, \mathrm{d}y.$$

Using the compact support of the window W, we now split this representation for $\|e_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)}$ into the sum of two integrals,

$$\|e_L\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \underbrace{\frac{1}{4\pi^2}\int_{r(x,y)\le L} |(\mathcal{F}f - W_L\cdot\mathcal{F}f)(x,y)|^2\,\mathrm{d}(x,y)}_{=I_1} + \underbrace{\frac{1}{4\pi^2}\int_{r(x,y)>L} |\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y)}_{=I_2},$$

where we let

$$r(x,y) = \sqrt{x^2+y^2} \quad \text{for } (x,y)\in\mathbb{R}^2.$$

Note that the first error term I_1 occurs if the chosen window function W is not constant 1, $W \not\equiv 1$, on $[-1,1]$, whereas the second error term I_2 occurs if the target function f is not band-limited, which is usually the case in applications of medical image reconstruction.

In the following, we analyse the two error terms separately. Since the window W is essentially bounded on $[-1,1]$, i.e., $W\in\mathrm{L}^\infty([-1,1])$, the first integral I_1 can be bounded above by

$$I_1 \le \|1-W_L\|^2_{\mathrm{L}^\infty([-L,L])}\,\frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} |\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y) = \|1-W\|^2_{\mathrm{L}^\infty([-1,1])}\,\|f\|^2_{\mathrm{L}^2(\mathbb{R}^2)}.$$

To estimate the second integral I_2, we now assume that the target function f belongs to a Sobolev space of positive fractional order. In fact, for $f\in\mathrm{H}^\alpha(\mathbb{R}^2)$, with $\alpha>0$, the integral I_2 can be bounded above by

$$\begin{aligned} I_2 &= \frac{1}{4\pi^2}\int_{r(x,y)>L} (1+x^2+y^2)^\alpha\,(1+x^2+y^2)^{-\alpha}\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y)\\ &\le \frac{1}{4\pi^2}\int_{r(x,y)>L} (1+x^2+y^2)^\alpha\,L^{-2\alpha}\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y) \le L^{-2\alpha}\,\|f\|^2_\alpha. \end{aligned}$$

We can summarize our discussion as follows.

Theorem 4.2.1 (L^2-error estimate)**.** *Let $f\in\mathrm{L}^1(\mathbb{R}^2)\cap\mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha>0$ and let $W\in\mathrm{L}^\infty(\mathbb{R})$ be even and compactly supported with $\operatorname{supp}(W)\subseteq[-1,1]$. Then, the L^2-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by*

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \|1-W\|_{\mathrm{L}^\infty([-1,1])}\,\|f\|_{\mathrm{L}^2(\mathbb{R}^2)} + L^{-\alpha}\,\|f\|_\alpha \tag{4.8}$$

for all $L>0$. □

We make the following remark about the result in Theorem 4.2.1.

Remark 4.2.2. *The first term on the right hand side of* (4.8) *depends on the choice of W, but not on L. Therefore, to obtain convergence $\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)}\longrightarrow 0$ for $L\longrightarrow\infty$ from* (4.8)*, we have to require*

$$\|1-W\|_{\mathrm{L}^\infty([-1,1])} = 0,$$

which is only satisfied by the window function $W=\chi_{[-1,1]}$ of the Ram-Lak filter*. However, in this case, the smoothness α of the target function f determines the decay rate in* (4.8) *by*

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le L^{-\alpha}\,\|f\|_\alpha = \mathcal{O}(L^{-\alpha}) \quad \textit{for} \quad L\longrightarrow\infty.$$

4.2.2 H^σ-error estimates

In this section we now prove H^σ-error estimates, for $0\le\sigma\le\alpha$, for the FBP reconstruction error $e_L = f - f_L$ for target functions $f\in\mathrm{L}^1(\mathbb{R}^2)\cap\mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha>0$, where we again assume that the filter's window $W\in\mathrm{L}^\infty(\mathbb{R})$ is even and compactly supported with $\operatorname{supp}(W)\subseteq[-1,1]$. As a special case, choosing $\sigma=0$ yields a refined L^2-error estimate which allows us to show convergence of e_L for a larger class of window functions than based on Theorem 4.2.1.

We first show that the approximate FBP reconstruction f_L belongs to the Sobolev space $\mathrm{H}^\sigma(\mathbb{R}^2)$ for $0 \le \sigma \le \alpha$. In Theorem 3.2.3 we have proven that the convolution kernel K_L belongs to $\mathscr{C}_0(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and, moreover, that its Fourier transform is given by $\mathcal{F}K_L = W_L$ with the compactly supported bivariate window function $W_L \in \mathrm{L}^\infty(\mathbb{R}^2)$ in (3.7). This in combination with representation (3.6) for f_L and the Fourier convolution Theorem 3.1.1 yields the estimate

$$\begin{aligned}\|f_L\|_\sigma^2 = \|f * K_L\|_\sigma^2 &= \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} (1 + x^2 + y^2)^\sigma \, |(W_L \cdot \mathcal{F}f)(x,y)|^2 \, \mathrm{d}x \, \mathrm{d}y \\ &\le \left(\operatorname*{ess\,sup}_{r(x,y) \le L} |W_L(x,y)|^2 \right) \|f\|_\alpha^2 = \|W\|_{\mathrm{L}^\infty(\mathbb{R})}^2 \, \|f\|_\alpha^2,\end{aligned}$$

where we again let

$$r(x,y) = \sqrt{x^2 + y^2} \quad \text{for } (x,y) \in \mathbb{R}^2.$$

Thus, for $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$, the approximate FBP reconstruction f_L belongs to the Sobolev space $\mathrm{H}^\sigma(\mathbb{R}^2)$ for all $0 \le \sigma \le \alpha$.

Let us now turn to the analysis of the FBP reconstruction error $e_L = f - f_L$ with respect to the H^σ-norm. For $\gamma \ge 0$, we define

$$r_\gamma(x,y) = (1 + r(x,y)^2)^\gamma = (1 + x^2 + y^2)^\gamma \quad \text{for } (x,y) \in \mathbb{R}^2$$

so that the H^σ-norm of e_L can be expressed as

$$\begin{aligned}\|e_L\|_\sigma^2 &= \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} (1 + x^2 + y^2)^\sigma \, |\mathcal{F}(f - f_L)(x,y)|^2 \, \mathrm{d}x \, \mathrm{d}y \\ &= \frac{1}{4\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} r_\sigma(x,y) \, |(\mathcal{F}f - W_L \cdot \mathcal{F}f)(x,y)|^2 \, \mathrm{d}x \, \mathrm{d}y = I_1 + I_2,\end{aligned}$$

where

$$I_1 = \frac{1}{4\pi^2} \int_{r(x,y) \le L} r_\sigma(x,y) \, |1 - W_L(x,y)|^2 \, |\mathcal{F}f(x,y)|^2 \, \mathrm{d}(x,y)$$

and

$$I_2 = \frac{1}{4\pi^2} \int_{r(x,y) > L} r_\sigma(x,y) \, |\mathcal{F}f(x,y)|^2 \, \mathrm{d}(x,y).$$

For $\gamma \ge 0$, we define

$$\Phi_{\gamma,W}(L) = \operatorname*{ess\,sup}_{S \in [-1,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\gamma} \quad \text{for } L > 0 \tag{4.9}$$

so that we can bound I_1 from above by

$$I_1 \le \left(\operatorname*{ess\,sup}_{r(x,y) \le L} \frac{(1 - W_L(x,y))^2}{r_{\alpha-\sigma}(x,y)} \right) \|f\|_\alpha^2 = \Phi_{\alpha-\sigma,W}(L) \|f\|_\alpha^2,$$

since

$$\operatorname*{ess\,sup}_{r(x,y) \le L} \frac{(1 - W_L(x,y))^2}{r_{\alpha-\sigma}(x,y)} = \operatorname*{ess\,sup}_{S \in [-L,L]} \frac{(1 - W(S/L))^2}{(1 + S^2)^{\alpha-\sigma}} = \operatorname*{ess\,sup}_{S \in [-1,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^{\alpha-\sigma}}.$$

In addition, for $0 \le \sigma \le \alpha$, we can bound I_2 by

$$I_2 \le L^{2(\sigma-\alpha)} \, \frac{1}{4\pi^2} \int_{r(x,y) > L} r_\alpha(x,y) \, |\mathcal{F}f(x,y)|^2 \, \mathrm{d}(x,y) \le L^{2(\sigma-\alpha)} \, \|f\|_\alpha^2.$$

Combining the estimates for I_1 and I_2, we finally obtain

$$\|e_L\|_\sigma^2 \le \left(\Phi_{\alpha-\sigma,W}(L) + L^{2(\sigma-\alpha)} \right) \|f\|_\alpha^2.$$

In summary, we have just established the following result.

Theorem 4.2.3 (H^σ-error estimate)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha > 0$ and let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even and compactly supported with $\operatorname{supp}(W) \subseteq [-1,1]$. Then, for $0 \leq \sigma \leq \alpha$, the H^σ-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by*

$$\|e_L\|_\sigma \leq \left(\Phi^{1/2}_{\alpha-\sigma,W}(L) + L^{\sigma-\alpha}\right) \|f\|_\alpha, \tag{4.10}$$

where

$$\Phi_{\alpha-\sigma,W}(L) = \operatorname*{ess\,sup}_{S\in[-1,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^{\alpha-\sigma}} \quad \text{for } L > 0. \qquad \square$$

For $\sigma = 0$ we obtain an L^2-error estimate which is more refined than that in Theorem 4.2.1.

Corollary 4.2.4 (Refined L^2-error estimate)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha > 0$ and let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even and compactly supported with $\operatorname{supp}(W) \subseteq [-1,1]$. Then, the L^2-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by*

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \left(\Phi^{1/2}_{\alpha,W}(L) + L^{-\alpha}\right) \|f\|_\alpha \tag{4.11}$$

for all $L > 0$. $\square$

In the next paragraph we will see that the refined L^2-error estimate (4.11) allows us to conclude convergence of the FBP reconstructions f_L to the target function f in the L^2-norm for a larger class of window functions W than based on Theorem 4.2.1.

4.2.3 Convergence

We now prove that, under suitable assumptions on the window function W, the approximate FBP reconstruction f_L converges to the target function f in the H^σ-norm for all $0 \leq \sigma < \alpha$. To this end, we first show that the error term $\Phi_{\gamma,W}(L)$ in (4.9) tends to zero as the bandwidth L goes to infinity.

Theorem 4.2.5 (Convergence of $\Phi_{\gamma,W}$)**.** *Let $W \in \mathscr{C}([-1,1])$ be even and continuous on $[-1,1]$ with $W(0) = 1$. Then, for any $\gamma > 0$,*

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \longrightarrow 0 \quad \text{for} \quad L \longrightarrow \infty.$$

Note that we require continuity of the compactly supported window function W only on the interval $[-1,1]$, but we allow W to have discontinuities at the boundary points of $[-1,1]$.

Proof. For the sake of brevity, we define the auxiliary function $\Phi_{\gamma,W,L} : [-1,1] \longrightarrow \mathbb{R}$ via

$$\Phi_{\gamma,W,L}(S) = \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } S \in [-1,1].$$

Because W is even and continuous on $[-1,1]$, $\Phi_{\gamma,W,L}$ is also even and continuous on $[-1,1]$. Thus, it attains its maximum on $[0,1]$ and we have

$$\Phi_{\gamma,W}(L) = \sup_{S\in[-1,1]} \Phi_{\gamma,W,L}(S) = \max_{S\in[-1,1]} \Phi_{\gamma,W,L}(S) = \max_{S\in[0,1]} \Phi_{\gamma,W,L}(S).$$

In the following, let $S^*_{\gamma,W,L} \in [0,1]$ denote the smallest maximizer of $\Phi_{\gamma,W,L}$ on $[0,1]$. For all $0 < L_1 < L_2$ now follows that

$$\begin{aligned}\Phi_{\gamma,W}(L_1) = \Phi_{\gamma,W,L_1}(S^*_{\gamma,W,L_1}) \geq \Phi_{\gamma,W,L_1}(S^*_{\gamma,W,L_2}) &= \frac{\left(1-W(S^*_{\gamma,W,L_2})\right)^2}{\left(1+L_1^2(S^*_{\gamma,W,L_2})^2\right)^\gamma} \\ \geq \frac{\left(1-W(S^*_{\gamma,W,L_2})\right)^2}{\left(1+L_2^2(S^*_{\gamma,W,L_2})^2\right)^\gamma} &= \Phi_{\gamma,W,L_2}(S^*_{\gamma,W,L_2}) = \Phi_{\gamma,W}(L_2).\end{aligned}$$

Furthermore,

$$0 \leq \Phi_{\gamma,W}(L) \leq \|1 - W\|^2_{\mathrm{L}^\infty([-1,1])} \quad \forall\, L > 0.$$

Therefore, $\Phi_{\gamma,W}(L)$ is bounded and monotonically decreasing in $L > 0$ so that $\Phi_{\gamma,W}(L)$ is convergent for $L \longrightarrow \infty$, i.e.,

$$\exists\, \Phi^*_{\gamma,W} \geq 0 : \ \Phi_{\gamma,W}(L) \longrightarrow \Phi^*_{\gamma,W} \quad \text{for} \quad L \longrightarrow \infty.$$

To determine the limit $\Phi^*_{\gamma,W}$, we distinguish two cases.

Case 1: $S^*_{\gamma,W,L}$ is uniformly bounded away from 0, i.e.,

$$\exists\, c \equiv c(\gamma, W) > 0 \,\forall\, L > 0 : \ S^*_{\gamma,W,L} \geq c.$$

In this case we get

$$0 \leq \Phi_{\gamma,W,L}(S^*_{\gamma,W,L}) = \frac{\left(1 - W(S^*_{\gamma,W,L})\right)^2}{\left(1 + L^2 (S^*_{\gamma,W,L})^2\right)^\gamma} \leq \frac{\|1 - W\|^2_{\mathrm{L}^\infty([-1,1])}}{(1 + L^2 c^2)^\gamma} \xrightarrow{L \to \infty} 0.$$

Case 2: $S^*_{\gamma,W,L}$ tends to 0 as L goes to ∞, i.e.,

$$S^*_{\gamma,W,L} \longrightarrow 0 \quad \text{for} \quad L \longrightarrow \infty.$$

Because W is continuous on $[-1, 1]$ and satisfies $W(0) = 1$, we have

$$W(S^*_{\gamma,W,L}) \longrightarrow W(0) = 1 \quad \text{for} \quad L \longrightarrow \infty$$

and, consequently,

$$0 \leq \Phi_{\gamma,W,L}(S^*_{\gamma,W,L}) = \frac{\left(1 - W(S^*_{\gamma,W,L})\right)^2}{\left(1 + L^2 (S^*_{\gamma,W,L})^2\right)^\gamma} \leq \left(1 - W(S^*_{\gamma,W,L})\right)^2 \xrightarrow{L \to \infty} 0.$$

Hence, in both cases we have

$$\Phi_{\gamma,W}(L) = \Phi_{\gamma,W,L}(S^*_{\gamma,W,L}) \longrightarrow 0 = \Phi^*_{\gamma,W} \quad \text{for} \quad L \longrightarrow \infty$$

and the proof is complete. □

By combining Theorems 4.2.3 and 4.2.5, we can now conclude convergence of the approximate FBP reconstruction f_L in the H^σ-norm for all $0 \leq \sigma < \alpha$.

Corollary 4.2.6 (Convergence of $\|e_L\|_\sigma$)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$, for some $\alpha > 0$, and let $W \in \mathscr{C}([-1, 1])$ be even with $W(0) = 1$. Then, for $0 \leq \sigma < \alpha$, the H^σ-norm of the FBP reconstruction error $e_L = f - f_L$ converges to 0 as L goes to ∞, i.e.,*

$$\|e_L\|_\sigma = o(1) \quad \textit{for} \quad L \longrightarrow \infty. \qquad \square$$

In Corollary 4.2.6 we have seen that for target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$ the H^σ-norm of the FBP reconstruction error $e_L = f - f_L$ converges to 0 for L going to ∞, as long as $0 \leq \sigma < \alpha$. Additionally, we had to assume that the even window function W is continuous on the interval $[-1, 1]$ and satisfies $W(0) = 1$. In the following, we relax these assumptions and proof convergence for the H^σ-norm of e_L for all $0 \leq \sigma \leq \alpha$, where especially the case $\sigma = \alpha$ is included. However, the proof technique is not suitable for determining the rate of convergence.

Theorem 4.2.7 (Convergence in the H^σ-norm)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for some $\alpha \geq 0$ and $W \in \mathrm{L}^\infty(\mathbb{R})$ be even and compactly supported with $\operatorname{supp}(W) \subseteq [-1, 1]$. Further, let W be continuous at 0 with $W(0) = 1$. Then, for $0 \leq \sigma \leq \alpha$, the H^σ-norm of the FBP reconstruction error $e_L = f - f_L$ converges to 0 as L goes to ∞, i.e.,*

$$\|e_L\|_\sigma = o(1) \quad \textit{for} \quad L \longrightarrow \infty.$$

Proof. For $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha \geq 0$ and an even window function $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\mathrm{supp}(W) \subseteq [-1,1]$ we have seen that the approximate FBP reconstruction f_L belongs to the Sobolev space $\mathrm{H}^\sigma(\mathbb{R}^2)$ for all $0 \leq \sigma \leq \alpha$ and that its Fourier transform is given by

$$\mathcal{F}f_L = W_L \cdot \mathcal{F}f$$

with the radially symmetric bivariate window function $W_L \in \mathrm{L}^\infty(\mathbb{R}^2)$ in (3.7), i.e.,

$$W_L(x,y) = W\Big(\frac{r(x,y)}{L}\Big) \quad \text{for } (x,y) \in \mathbb{R}^2,$$

where

$$r(x,y) = \sqrt{x^2+y^2} \quad \text{for } (x,y) \in \mathbb{R}^2.$$

Thus, for the H^σ-norm of the FBP reconstruction error $e_L = f - f_L$ follows that

$$\begin{aligned}\|e_L\|_\sigma^2 &= \frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} (1+x^2+y^2)^\sigma \,|\mathcal{F}(f-f_L)(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} (1+r(x,y)^2)^\sigma \,|1-W_L(x,y)|^2\,|\mathcal{F}f(x,y)|^2 \,\mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

For all $L > 0$ and almost all $(x,y) \in \mathbb{R}^2$ holds that

$$(1+r(x,y)^2)^\sigma \,|1-W_L(x,y)|^2\,|\mathcal{F}f(x,y)|^2 \leq 2\left(1+\|W\|^2_{\mathrm{L}^\infty(\mathbb{R})}\right)(1+r(x,y)^2)^\alpha\,|\mathcal{F}f(x,y)|^2,$$

where

$$\int_{\mathbb{R}}\int_{\mathbb{R}} (1+r(x,y)^2)^\alpha\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y = 4\pi^2\,\|f\|_\alpha^2 < \infty.$$

Further,

$$|1-W_L(x,y)|^2 = \left|1-W\Big(\frac{r(x,y)}{L}\Big)\right|^2 \xrightarrow{L\to\infty} 0 \quad \forall\,(x,y) \in \mathbb{R}^2,$$

since the window W is continuous at the origin and satisfies $W(0) = 1$. Therefore, we can apply Lebesgue's theorem on dominated convergence and obtain

$$\int_{\mathbb{R}}\int_{\mathbb{R}} (1+r(x,y)^2)^\sigma\,|1-W_L(x,y)|^2\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y \longrightarrow 0 \quad \text{for} \quad L \longrightarrow \infty.$$

Consequently,

$$\lim_{L\to\infty} \|e_L\|_\sigma = 0,$$

i.e., the approximate FBP reconstruction f_L converges to f in the H^σ-norm for all $0 \leq \sigma \leq \alpha$. □

As a corollary we obtain the convergence of the FBP reconstruction f_L in $\mathrm{L}^2(\mathbb{R}^2)$, which is also proven in [92] under stronger assumptions.

Corollary 4.2.8 (Convergence in the L^2-norm)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with* $\mathrm{supp}(W) \subseteq [-1,1]$*. Further, let W be continuous at 0 with $W(0) = 1$. For $L > 0$, define the low-pass filter $A_L \in \mathrm{L}^\infty(\mathbb{R})$ as*

$$A_L(S) = |S|\,W(S/L) \quad \textit{for } S \in \mathbb{R}.$$

Then, the approximate FBP reconstruction

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f) \in \mathrm{L}^2(\mathbb{R}^2)$$

converges to f in L^2-sense as L goes to ∞, i.e.,

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \longrightarrow 0 \quad \textit{for} \quad L \longrightarrow \infty.$$ □

We wish to make the following remark concerning the assumptions on the target function f.

Remark 4.2.9. *In practical situations, the target function $f \in \mathrm{L}^2(\mathbb{R}^2)$ is usually compactly supported in an open and bounded domain $\Omega \subset \mathbb{R}^2$. In this case, the assumption $f \in \mathrm{L}^1(\mathbb{R}^2)$ is automatically fulfilled and can be omitted in our theory, since Hölder's inequality yields*

$$\|f\|_{\mathrm{L}^1(\mathbb{R}^2)} = \int_\Omega |f(x,y)| \,\mathrm{d}(x,y) \leq |\Omega| \, \|f\|_{\mathrm{L}^2(\mathbb{R}^2)}$$

for all $f \in \mathrm{L}^2_c(\mathbb{R}^2)$ with $\operatorname{supp}(f) \subseteq \overline{\Omega}$. Consequently, the H^σ-error estimate in Theorem 4.2.3 as well as the convergence result in Corollary 4.2.6 remain valid if we only assume $f \in \mathrm{H}^\alpha_0(\Omega)$ for some $\alpha > 0$. In the language of inverse problems such an assumption is called source condition *for the target function f.*

To close this section we state the following negative result concerning the approximation of Sobolev functions $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$ with smoothness $\alpha \geq 0$ by band-limited functions $f_L \in \mathrm{H}^\alpha(\mathbb{R}^2)$ of bandwidth $L > 0$. More precisely, we show that the convergence of the approximation error can be arbitrarily slow as the bandwidth L goes to ∞. Note that this result is independent of the concrete approximation scheme.

Theorem 4.2.10. *Let $(\eta_k)_{k\in\mathbb{N}_0}$ be a monotonically decreasing sequence of real numbers $\eta_k > 0$ with*

$$\eta_k \searrow 0 \quad \text{for} \quad k \longrightarrow \infty$$

and let $(L_k)_{k\in\mathbb{N}_0}$ be a monotonically increasing sequence of bandwidths $L_k > 0$ satisfying

$$L_k \nearrow \infty \quad \text{for} \quad k \longrightarrow \infty.$$

Then, for every given smoothness $\alpha \geq 0$ there exists a function $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$ with

$$\|f - f_{L_k}\|_\alpha \geq \eta_k \quad \forall\, k \in \mathbb{N}_0.$$

Proof. For $k \in \mathbb{N}$ we define $\nu_k = \eta_{k-1}^2 - \eta_k^2$. By construction we have $\nu_k > 0$ and $(\nu_k)_{k\in\mathbb{N}} \in \ell^1(\mathbb{N})$ with

$$\sum_{k\in\mathbb{N}} \nu_k = \sum_{k\in\mathbb{N}} (\eta_{k-1}^2 - \eta_k^2) = \eta_0^2 < \infty.$$

Furthermore, for given smoothness $\alpha \geq 0$ we define the constant $\mu_k > 0$ via

$$\begin{aligned}\mu_k &= \frac{1}{4\pi^2} \int_{B_k \setminus B_{k-1}} (1 + x^2 + y^2)^\alpha \,\mathrm{d}(x,y) = \frac{1}{2\pi} \int_{L_{k-1}}^{L_k} (1 + S^2)^\alpha \, S \,\mathrm{d}S \\ &= \frac{1}{4\pi} \int_{L_{k-1}^2}^{L_k^2} (1 + S)^\alpha \,\mathrm{d}S = \frac{(1 + L_k^2)^{\alpha+1} - (1 + L_{k-1}^2)^{\alpha+1}}{4\pi\,(\alpha+1)},\end{aligned}$$

where $(B_k)_{k\in\mathbb{N}_0}$ is the nested sequence of the balls around zero with increasing radius L_k, i.e.,

$$B_k = \left\{ (x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq L_k^2 \right\}.$$

With this we now construct the requested function $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$ in Fourier domain via

$$\mathcal{F}f\Big|_{B_0} = 0 \quad \text{and} \quad \mathcal{F}f\Big|_{B_k \setminus B_{k-1}} = \sqrt{\frac{\nu_k}{\mu_k}}\, \chi_{B_k \setminus B_{k-1}} \quad \text{for } k \in \mathbb{N}.$$

Then, we indeed have

$$\begin{aligned}\|f\|_\alpha^2 &= \frac{1}{4\pi^2} \int_{\mathbb{R}^2} (1 + x^2 + y^2)^\alpha \, |\mathcal{F}f(x,y)|^2 \,\mathrm{d}(x,y) \\ &= \sum_{k\in\mathbb{N}} \frac{\nu_k}{\mu_k} \frac{1}{4\pi^2} \int_{B_k \setminus B_{k-1}} (1 + x^2 + y^2)^\alpha \,\mathrm{d}(x,y) = \sum_{k\in\mathbb{N}} \nu_k = \eta_0^2 < \infty\end{aligned}$$

and for the band-limited approximation $f_{L_k} \in \mathrm{H}^\alpha(\mathbb{R}^2)$ with bandwidth $L_k > 0$ follows that

$$\begin{aligned}\|f - f_{L_k}\|_\alpha^2 &= \frac{1}{4\pi^2} \int_{\mathbb{R}^2} (1 + x^2 + y^2)^\alpha \, |\mathcal{F}(f - f_{L_k})(x,y)|^2 \, \mathrm{d}(x,y) \\ &\geq \frac{1}{4\pi^2} \int_{\mathbb{R}^2 \setminus B_k} (1 + x^2 + y^2)^\alpha \, |\mathcal{F}f(x,y)|^2 \, \mathrm{d}(x,y) \\ &= \sum_{n>k} \frac{\nu_n}{\mu_n} \frac{1}{4\pi^2} \int_{B_n \setminus B_{n-1}} (1 + x^2 + y^2)^\alpha \, \mathrm{d}(x,y) = \sum_{n \geq k+1} \nu_n = \eta_k^2,\end{aligned}$$

which completes the proof. □

4.3 Rate of convergence

In the previous section we have proven that, under suitable assumptions on the window function W, the inherent FBP reconstruction error $e_L = f - f_L$ in (4.7) converges to zero in the H^σ-norm, for all $0 \leq \sigma \leq \alpha$, as the bandwidth L goes to infinity. Based on the H^σ-error estimate (4.10) in Theorem 4.2.3, i.e.,

$$\|e_L\|_\sigma \leq \left(\Phi_{\alpha-\sigma,W}^{1/2}(L) + L^{\sigma-\alpha} \right) \|f\|_\alpha,$$

where, for $\gamma \geq 0$,

$$\Phi_{\gamma,W}(L) = \operatorname*{ess\,sup}_{S \in [-1,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\gamma} \quad \text{for } L > 0,$$

we now analyse the convergence rate of $\|e_L\|_\sigma$ as L goes to ∞.

Note that the error term $\Phi_{\gamma,W}(L)$ only vanishes if the window W with $\operatorname{supp}(W) \subseteq [-1,1]$ satisfies

$$W = 1 \quad \text{a.e. on } [-1,1].$$

In this sense, the Ram-Lak filter from Example 3.2.8 is the *optimal* low-pass filter and we obtain

$$\|e_L\|_\sigma \leq L^{\sigma-\alpha} \|f\|_\alpha = \mathcal{O}(L^{\sigma-\alpha}) \quad \text{for} \quad L \longrightarrow \infty.$$

In particular, the convergence rate of the FBP reconstruction error with the Ram-Lak filter is determined by the difference between the smoothness α of the target function $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$ and the order σ of the Sobolev norm in which the reconstruction error e_L is measured.

Let us now analyse the decay rate of the FBP reconstruction error e_L in the H^σ-norm for arbitrary even window functions $W \in \mathrm{L}^\infty([-1,1])$ with $\operatorname{supp}(W) \subseteq [-1,1]$. To this end, let $S^*_{\gamma,W,L} \in [0,1]$, for $\gamma \geq 0$, denote the smallest maximizer in $[0,1]$ of the even function

$$\Phi_{\gamma,W,L}(S) = \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\gamma} \quad \text{for } S \in [-1,1],$$

i.e.,

$$S^*_{\gamma,W,L} = \sup \left\{ S \in [0,1] \mid \|\Phi_{\gamma,W,L}\|_{\mathrm{L}^\infty([0,S])} < \|\Phi_{\gamma,W,L}\|_{\mathrm{L}^\infty([0,1])} = \Phi_{\gamma,W}(L) \right\}.$$

To determine the rate of convergence for $\|e_L\|_\sigma$, we assume that $S^*_{\alpha-\sigma,W,L}$ is uniformly bounded away from 0 for all $0 \leq \sigma \leq \alpha$, i.e., there exists a constant $c_{\alpha-\sigma,W} > 0$ such that

$$S^*_{\alpha-\sigma,W,L} \geq c_{\alpha-\sigma,W} \quad \forall\, L > 0. \tag{A}$$

Then, the error term $\Phi_{\alpha-\sigma,W}(L)$ is bounded above by

$$\Phi_{\alpha-\sigma,W}(L) \leq \frac{\|1 - W\|^2_{\mathrm{L}^\infty([0,1])}}{\left(1 + L^2 (S^*_{\alpha-\sigma,W,L})^2\right)^{\alpha-\sigma}} \leq c_{\alpha-\sigma,W}^{2(\sigma-\alpha)} \, \|1 - W\|^2_{\mathrm{L}^\infty([0,1])} \, L^{2(\sigma-\alpha)},$$

since, for $\gamma = \alpha - \sigma$,

$$\Phi_{\gamma,W}(L) = \lim_{\varepsilon \searrow 0} \|\Phi_{\gamma,W,L}\|_{\mathrm{L}^\infty([S^*_{\gamma,W,L} - \varepsilon, 1])} \leq \lim_{\varepsilon \searrow 0} \frac{\|1 - W\|^2_{\mathrm{L}^\infty([0,1])}}{\left(1 + L^2 (S^*_{\gamma,W,L} - \varepsilon)^2\right)^\gamma} = \frac{\|1 - W\|^2_{\mathrm{L}^\infty([0,1])}}{\left(1 + L^2 (S^*_{\gamma,W,L})^2\right)^\gamma}.$$

Consequently, under Assumption (A) we obtain

$$\|e_L\|_\sigma \leq \left(c_{\alpha-\sigma,W}^{\sigma-\alpha}\,\|1-W\|_{\mathrm{L}^\infty([0,1])}+1\right) L^{\sigma-\alpha}\,\|f\|_\alpha,$$

i.e.,

$$\|e_L\|_\sigma = \mathcal{O}(L^{\sigma-\alpha}) \quad \text{for} \quad L \longrightarrow \infty.$$

In summary, we have just established the following result.

Theorem 4.3.1 (Convergence rate of $\|e_L\|_\sigma$)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for some $\alpha > 0$ and let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even and compactly supported with $\mathrm{supp}(W) \subseteq [-1,1]$. Further, let Assumption (A) be satisfied. Then, for $0 \leq \sigma \leq \alpha$, the H^σ-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by*

$$\|e_L\|_\sigma \leq \left(c_{\alpha-\sigma,W}^{\sigma-\alpha}\,\|1-W\|_{\mathrm{L}^\infty([0,1])}+1\right) L^{\sigma-\alpha}\,\|f\|_\alpha. \tag{4.12}$$

In particular,

$$\|e_L\|_\sigma = \mathcal{O}(L^{\sigma-\alpha}) \quad \textit{for} \quad L \longrightarrow \infty. \qquad \square$$

Note that the decay rate $\alpha - \sigma$ in (4.12) is determined by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error e_L is measured. Moreover, the bound on the inherent FBP reconstruction error in (4.12) is affine-linear with respect to $\|1-W\|_{\mathrm{L}^\infty([0,1])}$ and this quantity can be used to evaluate the approximation quality of the chosen window W satisfying Assumption (A).

Finally, we remark that Assumption (A) is fulfilled for a large class of window functions. For instance, let the window $W \in \mathrm{L}^\infty([-1,1])$ satisfy

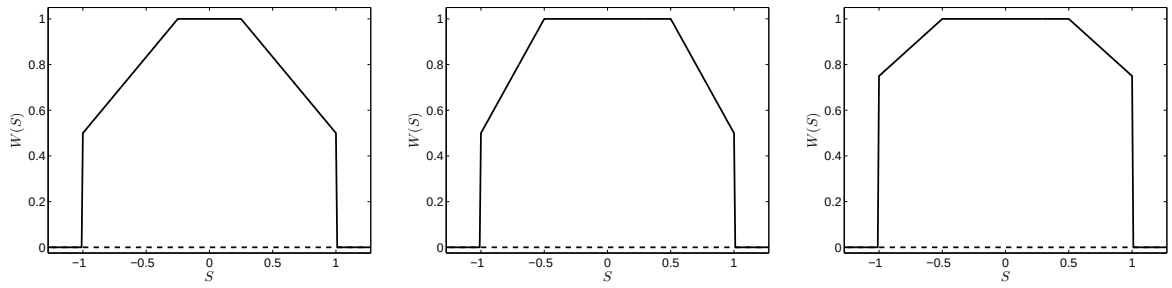
$$W(S) = 1 \quad \forall\, S \in (-\varepsilon, \varepsilon) \tag{B}$$

with $\varepsilon \in (0,1)$. Then, Assumption (A) is fulfilled with the constant $c_{\alpha-\sigma,W} = \varepsilon$ for all $0 \leq \sigma \leq \alpha$. To illustrate this, we consider the following concrete example.

Example 4.3.2. *The* generalized Ramp filter *of width $\beta \in [0,1)$ and with jump height $\lambda \in [0,1]$ is given by the window function*

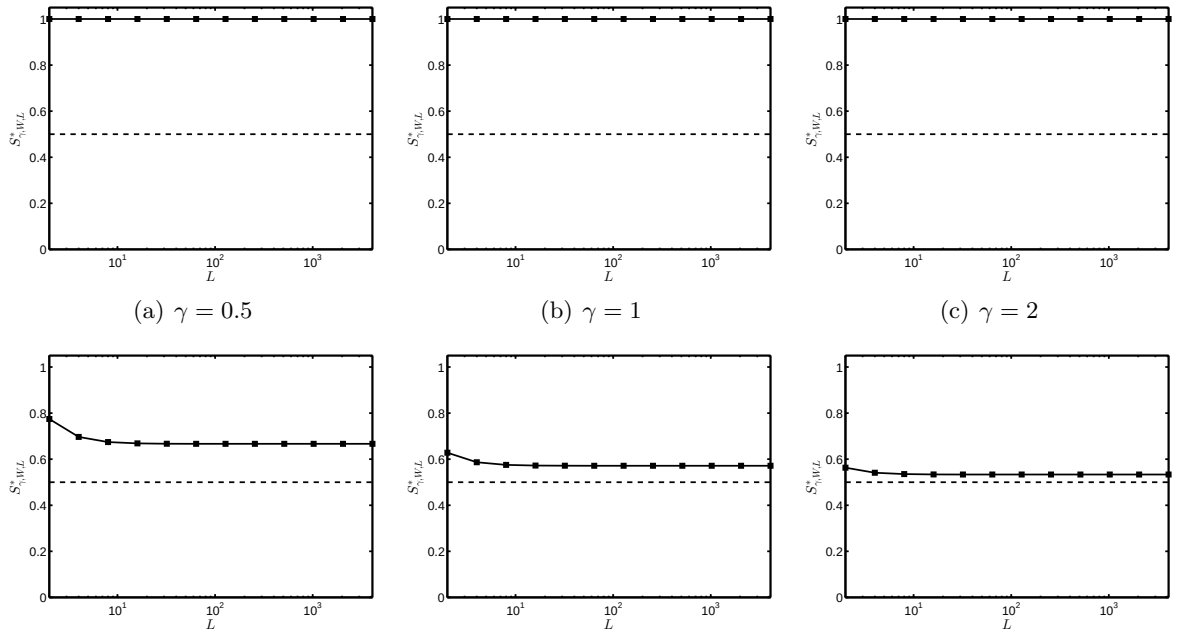
$$W(S) = \begin{cases} 1 & \textit{for } |S| \leq \beta \\ \frac{1-\beta\lambda}{1-\beta} - \frac{1-\lambda}{1-\beta}|S| & \textit{for } \beta < |S| \leq 1 \\ 0 & \textit{for } |S| > 1. \end{cases}$$

We remark that choosing $\lambda = 1$ results in the classical Ram-Lak filter from Example 3.2.8. For further plots of the generalized Ramp window with different parameters we refer to Figure 4.1.



(a) $\beta = 0.25$, $\lambda = 0.5$ (b) $\beta = 0.5$, $\lambda = 0.5$ (c) $\beta = 0.5$, $\lambda = 0.75$

Figure 4.1: Window functions of the generalized Ramp filter with different parameters.



(a) $\gamma = 0.5$ (b) $\gamma = 1$ (c) $\gamma = 2$

(d) $\gamma = 4$ (e) $\gamma = 8$ (f) $\gamma = 16$

Figure 4.2: Maximizer of $\Phi_{\gamma,W}(L)$ for the generalized Ramp filter with $\beta = 0.5$, $\lambda = 0.75$.

We observe that for all parameters $\beta \in (0,1)$ and $\lambda \in [0,1)$ the above window function satisfies Assumption (B)*. Consequently, our theory predicts that Assumption* (A) *is satisfied as well with $c_{\gamma,W} = \beta$ for all parameters $\gamma > 0$ and that the error term $\Phi_{\gamma,W}(L)$ is of order*

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty.$$

In fact, numerical calculations of the maximizer $S^*_{\gamma,W,L} \in [0,1]$ and the error term $\Phi_{\gamma,W}(L)$ show that Assumption (A) *is indeed satisfied for all choices of $\gamma > 0$ and all parameters $\beta \in (0,1)$ and $\lambda \in [0,1)$ with*

$$S^*_{\gamma,W,L} \geq \beta \quad \forall\, L > 0.$$

Further, $\Phi_{\gamma,W}(L)$ exactly behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty.$$

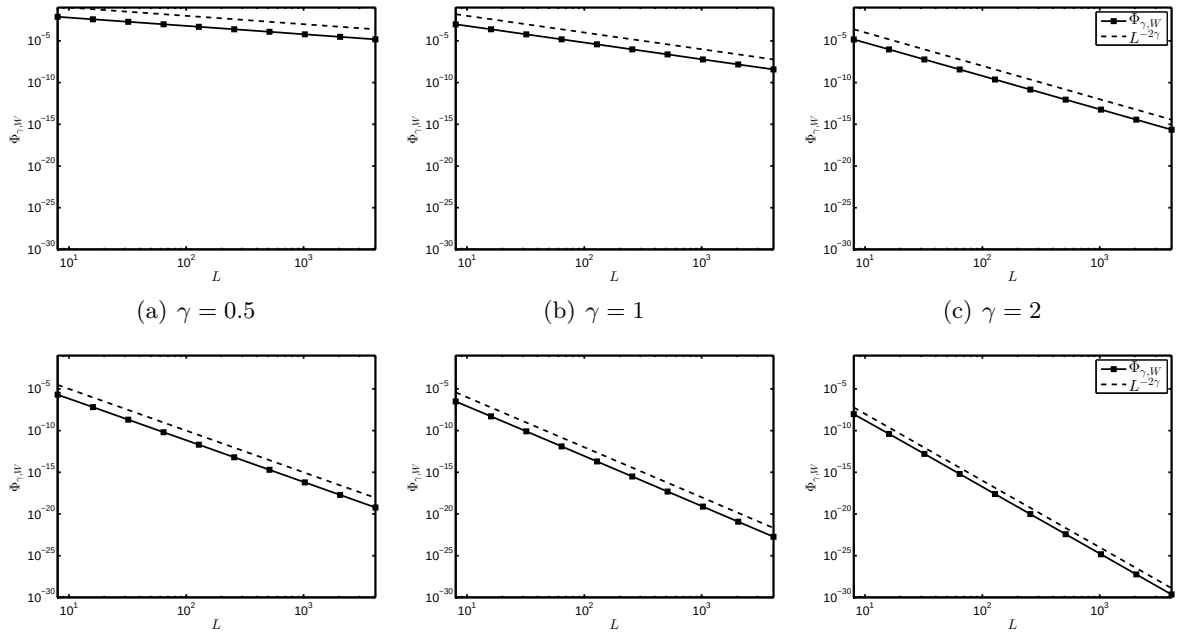
To illustrate this, Figure 4.2 shows $S^*_{\gamma,W,L}$ for $\gamma \in \{0.5, 1, 2, 4, 8, 16\}$ as a function of the bandwidth $L > 0$ in semi-logarithmic scales exemplary for the parameters $\beta = 0.5$ and $\lambda = 0.75$. In addition, Figure 4.3 shows $\Phi_{\gamma,W}(L)$ for $\gamma \in \{0.5, 1, 2, 2.5, 3, 4\}$ in logarithmic scales. We remark that the same behaviour of $S^*_{\gamma,W,L}$ and $\Phi_{\gamma,W}(L)$ was observed for any other choices of β and λ.

Note that Assumption (A), however, is not always satisfied for typical choices of the low-pass filter. To further explain this, we now investigate the behaviour of $S^*_{\gamma,W,L} \in [0,1]$ and $\Phi_{\gamma,W}(L)$ numerically for the following commonly used filter functions $A_L(S) = |S|\, W(S/L)$:

Name	$W(S)$ for $\lvert S\rvert \leq 1$	Parameter
Shepp-Logan	$\operatorname{sinc}(\pi S/2)$	-
Cosine	$\cos(\pi S/2)$	-
Hamming	$\beta + (1-\beta)\cos(\pi S)$	$\beta \in [1/2, 1]$
Gaussian	$\exp(-(\pi S/\beta)^2)$	$\beta > 1$

Note that each of these windows $W \in \mathrm{L}^\infty(\mathbb{R})$ is even with compact support $\operatorname{supp}(W) = [-1,1]$.

In our numerical experiments, we calculated $S^*_{\gamma,W,L}$ and $\Phi_{\gamma,W}(L)$ as functions of the bandwidth $L > 0$ for the above mentioned window functions W and for different parameters $\gamma > 0$.



(a) $\gamma = 0.5$ (b) $\gamma = 1$ (c) $\gamma = 2$

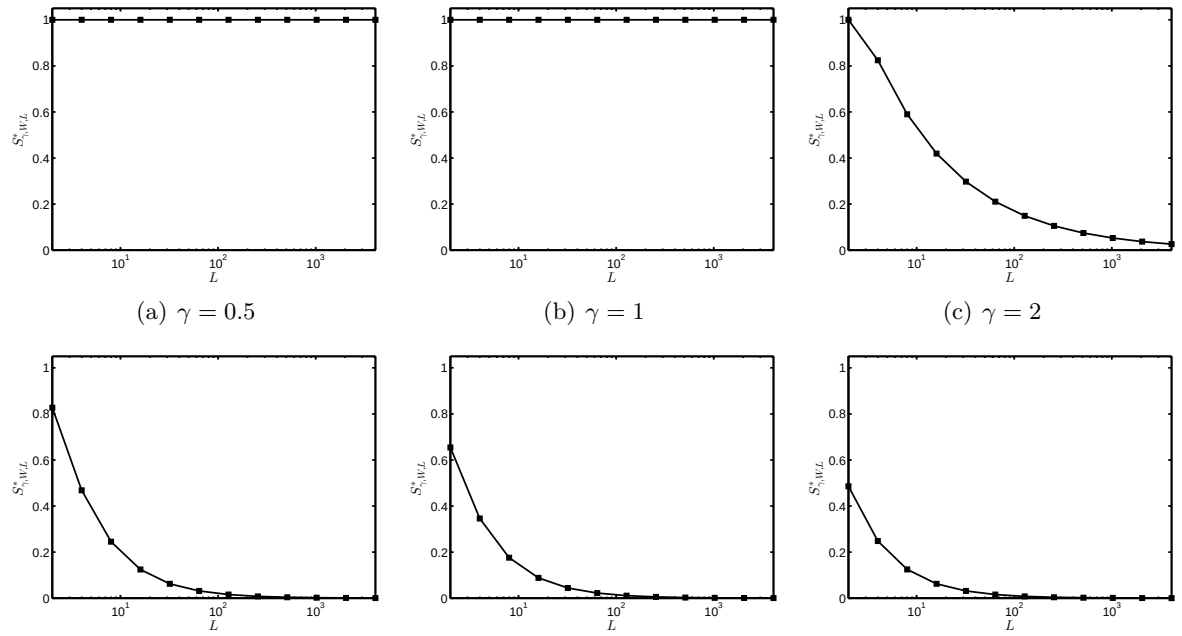
(d) $\gamma = 2.5$ (e) $\gamma = 3$ (f) $\gamma = 4$

Figure 4.3: Decay rate of $\Phi_{\gamma,W}(L)$ for the generalized Ramp filter with $\beta = 0.5$, $\lambda = 0.75$.

Figure 4.4 shows the behaviour of the maximizer $S^*_{\gamma,W,L}$ in semi-logarithmic scales for the Shepp-Logan filter and for the parameters $\gamma \in \{0.5, 1, 2, 2.5, 3, 4\}$. For $\gamma \in \{0.5, 1\}$ we observe that

$$S^*_{\gamma,W,L} = 1 \quad \forall\, L > 0$$

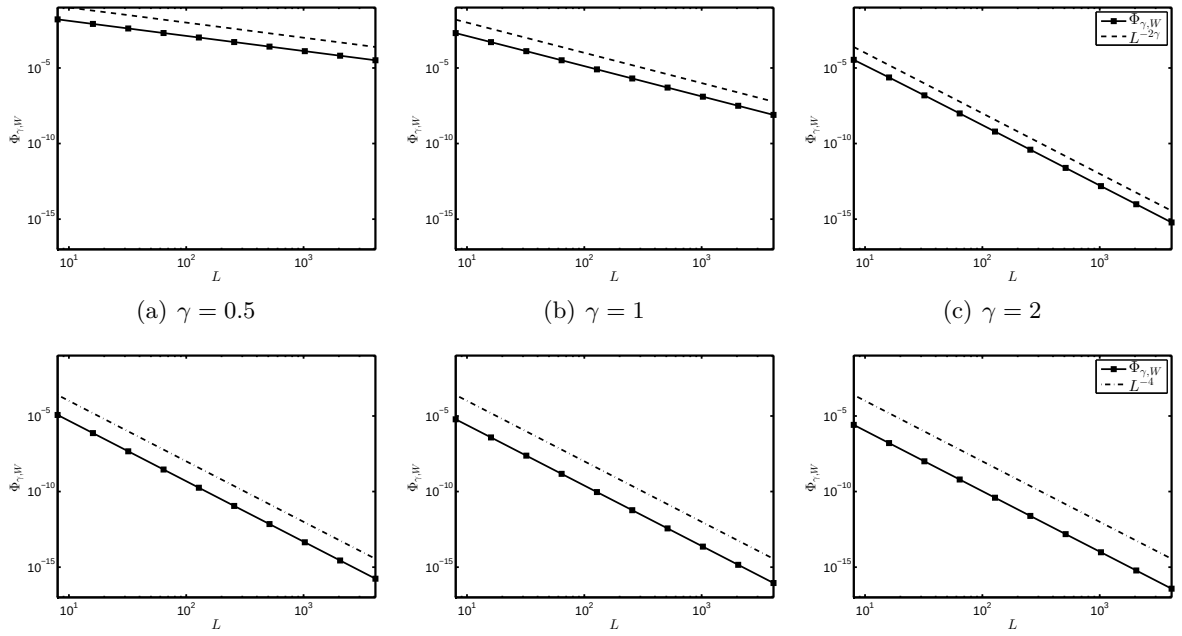
and, in particular, $S^*_{\gamma,W,L}$ is uniformly bounded away from 0 for all $L > 0$ so that Assumption (A) is satisfied. In contrast to that, for $\gamma \in \{2, 2.5, 3, 4\}$ the maximizer $S^*_{\gamma,W,L}$ goes to 0 as the bandwidth L tends to ∞ and, thus, Assumption (A) is not fulfilled in this case.



(a) $\gamma = 0.5$ (b) $\gamma = 1$ (c) $\gamma = 2$

(d) $\gamma = 2.5$ (e) $\gamma = 3$ (f) $\gamma = 4$

Figure 4.4: Maximizer of $\Phi_{\gamma,W}(L)$ for the Shepp-Logan filter.



(a) $\gamma = 0.5$ (b) $\gamma = 1$ (c) $\gamma = 2$

(d) $\gamma = 2.5$ (e) $\gamma = 3$ (f) $\gamma = 4$

Figure 4.5: Decay rate of $\Phi_{\gamma,W}(L)$ for the Shepp-Logan filter.

Additionally, Figure 4.5 shows the behaviour of $\Phi_{\gamma,W}(L)$ in logarithmic scales for the same filter and parameters. For $\gamma \in \{0.5, 1, 2\}$ we observe that $\Phi_{\gamma,W}(L)$ behaves exactly as $L^{-2\gamma}$, whereas for $\gamma \in \{2.5, 3, 4\}$ the behaviour of $\Phi_{\gamma,W}(L)$ corresponds to L^{-4}. In the latter case, however, $\Phi_{\gamma,W}(L)$ decreases at increasing values $\gamma > 2$. We remark that the same behaviour was observed in our numerical experiments for the other window functions W mentioned above.

We summarize our numerical experiments (for all window functions W from above) as follows. For $\gamma < 2$, we see that Assumption (A), i.e.,

$$\exists\, c_{\gamma,W} > 0 : \; S^*_{\gamma,W,L} \geq c_{\gamma,W} \quad \forall\, L > 0,$$

is fulfilled, where, in particular,

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty.$$

For $\gamma \geq 2$, however, we have

$$S^*_{\gamma,W,L} \longrightarrow 0 \quad \text{for} \quad L \longrightarrow \infty$$

and the convergence rate of $\Phi_{\gamma,W}$ stagnates at

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-4}) \quad \text{for} \quad L \longrightarrow \infty.$$

Note that all above windows W are twice continuously differentiable on $[-1, 1]$, $W \in \mathscr{C}^2([-1, 1])$, and satisfy

$$W(0) = 1 \quad \text{and} \quad W'(0) = 0.$$

Based on this, the observed behaviour of the error term $\Phi_{\gamma,W}(L)$ can also be proven in theory, as explained in the subsequent paragraph.

4.3.1 Convergence rates for $\mathscr{C}^k$-windows

In this section we consider the special case of even window functions W with compact support in the interval $[-1, 1]$ that additionally satisfy $W \in \mathscr{C}^k([-1, 1])$, for some $k \in \mathbb{N}$, with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k - 1.$$

We start with analysing the convergence behaviour of the error term $\Phi_{\gamma,W}(L)$ in (4.9), i.e.,

$$\Phi_{\gamma,W}(L) = \sup_{S\in[-1,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } L>0.$$

To this end, for parameters $\gamma \geq 0$, $L>0$ and $\nu>0$ we first consider the auxiliary function

$$\phi_{\gamma,L,\nu}(S) = \frac{S^{2\nu}}{(1+L^2S^2)^\gamma} \quad \text{for } S\in\mathbb{R}. \tag{4.13}$$

Lemma 4.3.3. *The maximum of the function $\phi_{\gamma,L,\nu}$ in (4.13) on $[0,1]$ is bounded above by*

$$\max_{S\in[0,1]} \phi_{\gamma,L,\nu}(S) \leq \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq \nu \vee \left(\gamma > \nu \wedge L < \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}}\right) \\ c_{\gamma,\nu}^2\, L^{-2\nu} & \text{for } \gamma > \nu \wedge L \geq \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}}, \end{cases}$$

where the constant

$$c_{\gamma,\nu} = \left(\frac{\nu}{\gamma-\nu}\right)^{\nu/2} \left(\frac{\gamma-\nu}{\gamma}\right)^{\gamma/2} \quad \text{for } \gamma>\nu$$

is strictly monotonically decreasing in $\gamma>\nu$ and satisfies

$$c_{\gamma,\nu} \longrightarrow 1 \quad \text{for} \quad \gamma \longrightarrow \nu \qquad \text{and} \qquad c_{\gamma,\nu} \longrightarrow 0 \quad \text{for} \quad \gamma \longrightarrow \infty$$

with decay rate

$$c_{\gamma,\nu} \leq \nu^{\nu/2}\, \gamma^{-\nu/2} = \mathcal{O}(\gamma^{-\nu/2}) \quad \text{for} \quad \gamma \longrightarrow \infty.$$

Proof. The even and continuous function $\phi_{\gamma,L,\nu}$ in (4.13),

$$\phi_{\gamma,L,\nu}(S) = \frac{S^{2\nu}}{(1+L^2\,S^2)^\gamma} \quad \text{for } S\in\mathbb{R},$$

attains its maximum on $[0,1]$ in the half-open interval $(0,1]$, since we have

$$\phi_{\gamma,L,\nu}(S) \geq 0 \quad \forall\, S\in[0,1]$$

and

$$\phi_{\gamma,L,\nu}(S) = 0 \quad \Longleftrightarrow \quad S=0.$$

For $S\in(0,1)$ the first derivative of $\phi_{\gamma,L,\nu}(S)$ is given by

$$\phi'_{\gamma,L,\nu}(S) = \frac{2\nu\, S^{2\nu-1}\,(1+L^2\,S^2) - 2\gamma\, L^2\, S^{2\nu+1}}{(1+L^2\,S^2)^{\gamma+1}} = \frac{2\,S^{2\nu-1}\,(\nu+(\nu-\gamma)\,L^2\,S^2)}{(1+L^2\,S^2)^{\gamma+1}}$$

so that the necessary condition for a maximum of $\phi_{\gamma,L,\nu}$ in $(0,1)$ reads

$$\phi'_{\gamma,L,\nu}(S) = 0 \quad \overset{S>0}{\Longleftrightarrow} \quad (\gamma-\nu)\,L^2\,S^2 = \nu.$$

Case 1: For $0\leq\gamma\leq\nu$ the equation

$$(\gamma-\nu)\,L^2\,S^2 = \nu$$

has no solution in $[0,1]$ and, actually,

$$\phi'_{\gamma,L,\nu}(S) > 0 \quad \forall\, S\in(0,1].$$

This shows that $\phi_{\gamma,L,\nu}$ is strictly monotonically increasing in $(0,1]$ and, thus, maximal at $S^*=1$, i.e.,

$$\max_{S\in[0,1]} \phi_{\gamma,L,\nu}(S) = \phi_{\gamma,L,\nu}(1) = \frac{1}{(1+L^2)^\gamma} \leq L^{-2\gamma}.$$

Case 2: For $\gamma > \nu$ the unique positive solution of the equation

$$(\gamma - \nu)\, L^2\, S^2 = \nu$$

is given by

$$S^* = \frac{\sqrt{\nu}}{L\sqrt{\gamma - \nu}},$$

where

$$S^* \in [0,1] \quad \Longleftrightarrow \quad L \geq \frac{\sqrt{\nu}}{\sqrt{\gamma - \nu}}.$$

For convenience, we define the function $g_{\gamma,L,\nu} : \mathbb{R} \longrightarrow \mathbb{R}$ via

$$g_{\gamma,L,\nu}(S) = \nu + (\nu - \gamma)\, L^2\, S^2 \quad \text{for } S \in \mathbb{R}.$$

Since $g_{\gamma,L,\nu}$ is a parabola opening downwards with vertex in 0, we obtain

$$g_{\gamma,L,\nu}(s) > g_{\gamma,L,\nu}(S) \quad \forall\, 0 < s < S$$

such that

$$\phi'_{\gamma,L,\nu}(S) < \phi'_{\gamma,L,\nu}(S^*) = 0 < \phi'_{\gamma,L,\nu}(s) \quad \forall\, 0 < s < S^* < S.$$

Consequently, $\phi_{\gamma,L,\nu}$ is strictly monotonically increasing on $(0, S^*)$ and strictly monotonically decreasing on (S^*, ∞). Therefore, S^* is the unique maximizer of $\phi_{\gamma,L,\nu}$ on $\mathbb{R}_{\geq 0}$ and it follows that

$$\operatorname*{arg\,max}_{S \in [0,1]} \phi_{\gamma,L,\nu}(S) = \begin{cases} 1 & \text{for } L < \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}} \\ S^* & \text{for } L \geq \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}} \end{cases} = \begin{cases} 1 & \text{for } L < \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}} \\ \frac{\sqrt{\nu}}{L\sqrt{\gamma-\nu}} & \text{for } L \geq \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}}. \end{cases}$$

Because of

$$\phi_{\gamma,L,\nu}(S^*) = \Big(\frac{\nu}{\gamma - \nu}\Big)^{\nu} \Big(\frac{\gamma - \nu}{\gamma}\Big)^{\gamma} L^{-2\nu} = c_{\gamma,\nu}^2\, L^{-2\nu}$$

we finally get

$$\max_{S \in [0,1]} \phi_{\gamma,L,\nu}(S) = \begin{cases} \phi_{\gamma,L,\nu}(1) & \text{for } L < \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}} \\ \phi_{\gamma,L,\nu}(S^*) & \text{for } L \geq \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}} \end{cases} \leq \begin{cases} L^{-2\gamma} & \text{for } L < \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}} \\ c_{\gamma,\nu}^2\, L^{-2\nu} & \text{for } L \geq \frac{\sqrt{\nu}}{\sqrt{\gamma-\nu}}. \end{cases}$$

Let us now regard the constant $c_{\gamma,\nu}$ as a function $c_\nu \equiv c_\nu(\gamma)$ of the parameter $\gamma > \nu$, i.e.,

$$c_\nu(\gamma) = \Big(\frac{\nu}{\gamma - \nu}\Big)^{\nu/2} \Big(\frac{\gamma - \nu}{\gamma}\Big)^{\gamma/2} \quad \text{for } \gamma > \nu.$$

By using the representation

$$\Big(\frac{\gamma - \nu}{\gamma}\Big)^{\gamma/2} = \exp\left(\frac{\gamma}{2}\, \log\Big(1 - \frac{\nu}{\gamma}\Big)\right),$$

for the derivative of c_ν follows that

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} c_\nu(\gamma) = \frac{1}{2} \Big(\frac{\nu}{\gamma - \nu}\Big)^{\nu/2} \Big(\frac{\gamma - \nu}{\gamma}\Big)^{\gamma/2} \log\Big(1 - \frac{\nu}{\gamma}\Big) < 0 \quad \forall\, \gamma > \nu$$

and, consequently, the constant $c_{\gamma,\nu}$ is strictly monotonically decreasing in $\gamma > \nu$. In addition, we have

$$\Big(\frac{\gamma - \nu}{\gamma}\Big)^{\gamma/2} = \Big(1 - \frac{\nu}{\gamma}\Big)^{\gamma/2} \longrightarrow \mathrm{e}^{-\nu/2} \quad \text{for} \quad \gamma \longrightarrow \infty$$

and

$$\Big(\frac{\nu}{\gamma - \nu}\Big)^{\nu/2} \longrightarrow 0 \quad \text{for} \quad \gamma \longrightarrow \infty.$$

This already implies that

$$c_{\gamma,\nu} \longrightarrow 0 \quad \text{for} \quad \gamma \longrightarrow \infty$$

with decay rate

$$c_{\gamma,\nu} = \nu^{\nu/2}\,(\gamma-\nu)^{(\gamma-\nu)/2}\,\gamma^{-\gamma/2} \leq \nu^{\nu/2}\,\gamma^{(\gamma-\nu)/2}\,\gamma^{-\gamma/2} = \nu^{\nu/2}\,\gamma^{-\nu/2} = \mathcal{O}(\gamma^{-\nu/2}) \quad \text{for} \quad \gamma \longrightarrow \infty.$$

Since

$$(\gamma-\nu)^{(\gamma-\nu)/2} = \exp\Big(\frac{\gamma-\nu}{2}\,\log(\gamma-\nu)\Big) \longrightarrow \mathrm{e}^0 = 1 \quad \text{for} \quad \gamma \longrightarrow \nu$$

and

$$\gamma^{-\gamma/2} \longrightarrow \nu^{-\nu/2} \quad \text{for} \quad \gamma \longrightarrow \nu,$$

we further obtain

$$c_{\gamma,\nu} = \nu^{\nu/2}\,(\gamma-\nu)^{(\gamma-\nu)/2}\,\gamma^{-\gamma/2} \longrightarrow \nu^{\nu/2}\,\nu^{-\nu/2} = 1 \quad \text{for} \quad \gamma \longrightarrow \nu$$

and the proof is complete. □

With Lemma 4.3.3 we are now prepared to analyse the convergence rate of the error term $\Phi_{\gamma,W}(L)$ for $\mathscr{C}^k$-window functions.

Theorem 4.3.4 (Convergence rate of $\Phi_{\gamma,W}$ for $\mathscr{C}^k$-windows)**.** *Let the window function W satisfy $W \in \mathscr{C}^k([-1,1])$, for $k \in \mathbb{N}$, with*

$$W(0) = 1 \quad \textit{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Then, for $0 \leq \gamma \leq k$, we have

$$\Phi_{\gamma,W}(L) \leq \frac{1}{(k!)^2}\,\|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\,L^{-2\gamma}$$

and, for $\gamma > k$,

$$\Phi_{\gamma,W}(L) \leq \begin{cases} \frac{1}{(k!)^2}\,\|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\,L^{-2\gamma} & \textit{for } L < L^* \\ \frac{c^2_{\gamma,k}}{(k!)^2}\,\|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\,L^{-2k} & \textit{for } L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\gamma-k}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k} = \Big(\frac{k}{\gamma-k}\Big)^{k/2}\,\Big(\frac{\gamma-k}{\gamma}\Big)^{\gamma/2} \qquad \textit{for } \gamma > k.$$

In particular,

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k,\gamma\}}\Big) \quad \textit{for} \quad L \longrightarrow \infty.$$

Note that we require differentiability of the window W only on the interval $[-1,1]$ where it is supported. But we allow discontinuities of W at the boundary points of $[-1,1]$.

Proof. Since the window W is assumed to be continuous on $[-1,1]$, we have

$$\Phi_{\gamma,W}(L) = \max_{S\in[-1,1]} \Phi_{\gamma,W,L}(S)$$

with

$$\Phi_{\gamma,W,L}(S) = \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } S \in [-1,1].$$

Let $S \in [-1,1]$ be fixed. By assumption, W is even and satisfies $W \in \mathscr{C}^k([-1,1])$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Thus, we can apply Taylor's Theorem A.4.2 and with some ξ between 0 and S follows that

$$W(S) = W(0) + \sum_{j=1}^{k-1} \frac{1}{j!} W^{(j)}(0)\, S^j + \frac{1}{k!} W^{(k)}(\xi)\, S^k = 1 + \frac{1}{k!} W^{(k)}(\xi)\, S^k.$$

This leads to

$$\Phi_{\gamma,W,L}(S) = \frac{W^{(k)}(\xi)^2}{(k!)^2} \frac{S^{2k}}{(1+L^2S^2)^\gamma} \leq \frac{\|W^{(k)}\|^2_{\mathrm{L}^\infty([-1,1])}}{(k!)^2} \frac{S^{2k}}{(1+L^2S^2)^\gamma}.$$

Hence,

$$\Phi_{\gamma,W}(L) \leq \frac{\|W^{(k)}\|^2_{\mathrm{L}^\infty([-1,1])}}{(k!)^2} \max_{S\in[-1,1]} \frac{S^{2k}}{(1+L^2S^2)^\gamma} = \frac{1}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])} \max_{S\in[0,1]} \phi_{\gamma,L,k}(S)$$

and we now need to analyse

$$\phi_{\gamma,L,k}(S) = \frac{S^{2k}}{(1+L^2S^2)^\gamma} \quad \text{for } S \in [0,1].$$

Due to Lemma 4.3.3 the maximum of $\phi_{\gamma,L,k}$ on $[0,1]$ is bounded above by

$$\max_{S\in[0,1]} \phi_{\gamma,L,k}(S) \leq \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq k \vee (\gamma > k \wedge L < L^*) \\ c^2_{\gamma,k}\, L^{-2k} & \text{for } \gamma > k \wedge L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\gamma-k}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k} = \Big(\frac{k}{\gamma-k}\Big)^{k/2} \Big(\frac{\gamma-k}{\gamma}\Big)^{\gamma/2} \quad \text{for } \gamma > k.$$

Combining our results yields

$$\begin{aligned} \Phi_{\gamma,W}(L) &\leq \frac{1}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])} \max_{S\in[0,1]} \phi_{\gamma,L,k}(S) \\ &\leq \begin{cases} \frac{1}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\, L^{-2\gamma} & \text{for } \gamma \leq k \vee (\gamma > k \wedge L < L^*) \\ \frac{c^2_{\gamma,k}}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\, L^{-2k} & \text{for } \gamma > k \wedge L \geq L^*, \end{cases} \end{aligned}$$

which completes the proof. □

We remark that the results of Theorem 4.3.4 comply with our numerical observations from the previous section, where we have $k = 2$. In particular, we have observed saturation of the convergence order of $\Phi_{\gamma,W}(L)$ for $\gamma > 2$ at rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-4}) \quad \text{for} \quad L \longrightarrow \infty$$

through our numerical experiments. Further, we have seen that $\Phi_{\gamma,W}(L)$ continues to decrease at increasing values $\gamma > 2$, which corresponds to the monotonic decrease of the constant $c_{\gamma,2}$.

Combining the Theorems 4.2.3 and 4.3.4 now yields the following H^σ-error estimate for the approximate FBP reconstruction with $\mathscr{C}^k$-windows.

Corollary 4.3.5 (H^σ-error estimate for $\mathscr{C}^k$-windows)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$, for $\alpha > 0$, and let $W \in \mathscr{C}^k([-1,1])$, for $k \in \mathbb{N}$, with*

$$W(0) = 1 \quad \textit{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Then, for $0 \leq \sigma \leq \alpha$, *the* H^σ*-norm of the inherent FBP reconstruction error* $e_L = f - f_L$ *is bounded above by*

$$\|e_L\|_\sigma \leq \left(\frac{1}{k!}\,\|W^{(k)}\|_{\mathrm{L}^\infty([0,1])} + 1\right) L^{\sigma-\alpha}\,\|f\|_\alpha \tag{4.14}$$

if $\alpha - \sigma \leq k$ *and, if* $\alpha - \sigma > k$, *by*

$$\|e_L\|_\sigma \leq \begin{cases} \left(\frac{1}{k!}\,\|W^{(k)}\|_{\mathrm{L}^\infty([0,1])} + 1\right) L^{\sigma-\alpha}\,\|f\|_\alpha & \text{for } L < L^* \\ \left(\frac{c_{\alpha-\sigma,k}}{k!}\,\|W^{(k)}\|_{\mathrm{L}^\infty([0,1])}\,L^{-k} + L^{\sigma-\alpha}\right)\|f\|_\alpha & \text{for } L \geq L^* \end{cases} \tag{4.15}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\alpha-\sigma-k}}$ *and the strictly monotonically decreasing constant*

$$c_{\alpha-\sigma,k} = \Big(\frac{k}{\alpha-\sigma-k}\Big)^{k/2}\Big(\frac{\alpha-\sigma-k}{\alpha-\sigma}\Big)^{(\alpha-\sigma)/2} \qquad \text{for } \alpha-\sigma > k.$$

In particular,

$$\|e_L\|_\sigma = \mathcal{O}\Big(L^{-\min\{k,\alpha-\sigma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty. \qquad \square$$

Note that for $\alpha - \sigma \leq k$ the convergence order in Corollary 4.3.5 is determined by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error e_L is measured, whereas for $\alpha - \sigma > k$ the convergence order saturates at rate $\mathcal{O}(L^{-k})$. But in this case the error bound still decreases at increasing smoothness α, since the involved constant $c_{\alpha-\sigma,k}$ is strictly monotonically decreasing in $\alpha-\sigma > k$. Thus, a smoother target function allows for a better approximation, as expected. Nevertheless, the attainable convergence rate is limited by the differentiability order k of the filter's window W.

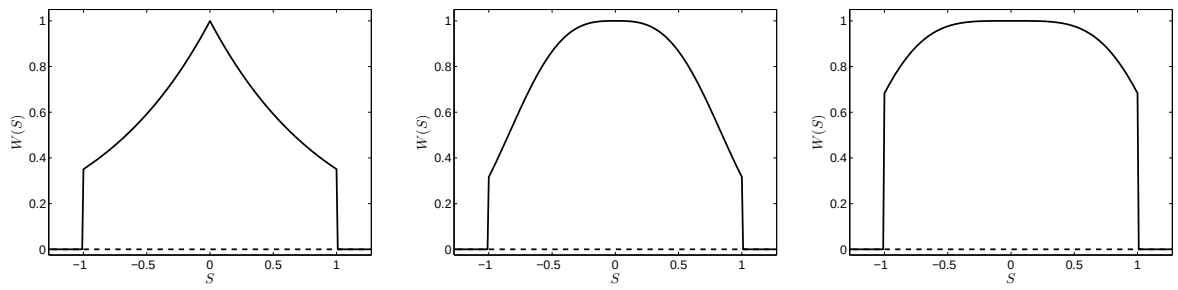
Finally, note that the bound on the inherent FBP reconstruction error in Corollary 4.3.5 is affine-linear with respect to $\|W^{(k)}\|_{\mathrm{L}^\infty([0,1])}$ and that this quantity can be used to evaluate the approximation quality of the chosen $\mathscr{C}^k$-window function W.

We close this section with the following numerical example, which illustrates that the proven decay rate of $\Phi_{\gamma,W}(L)$ is optimal for $\mathscr{C}^k$-windows.

Example 4.3.6. *We investigate the behaviour of the error term* $\Phi_{\gamma,W}(L)$ *numerically for the* generalized Gaussian filter $A_L(S) = |S|\,W(S/L)$ *with the window function*

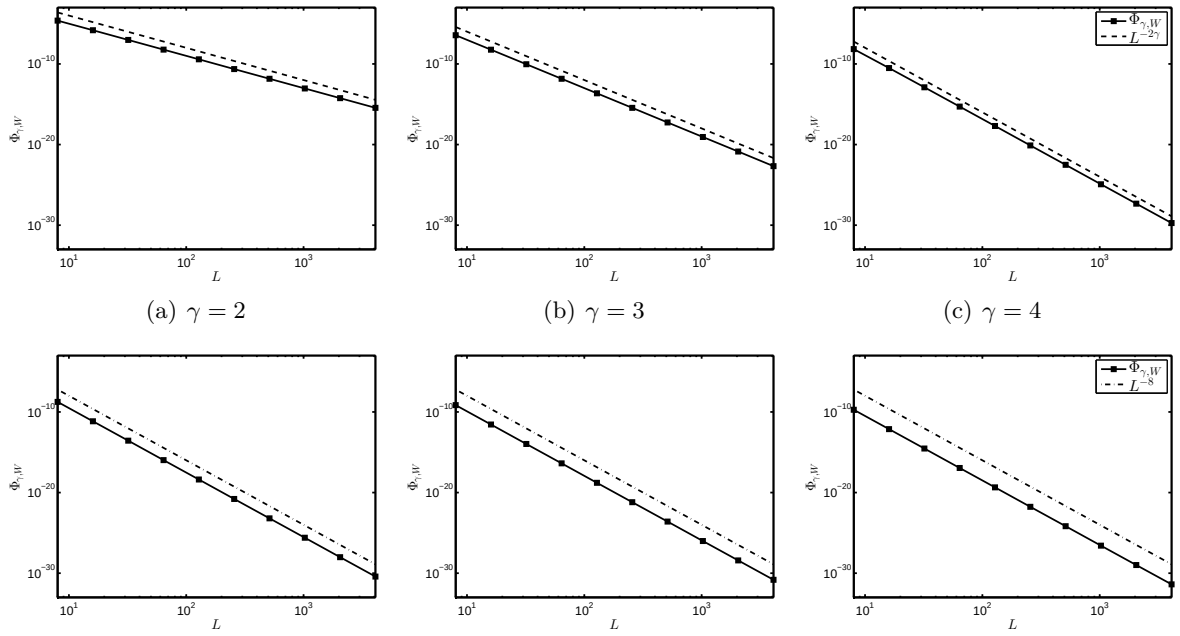
$$W(S) = \exp\Big(-\Big(\frac{\pi S}{\beta}\Big)^k\Big) \qquad \text{for } S \in [-1,1]$$

for some $k \in \mathbb{N}$ *and* $\beta > 1$. *Note that* $W \in \mathscr{C}([-1,1])$ *is even and compactly supported in* $[-1,1]$ *and that choosing* $k = 2$ *gives the classical Gaussian filter from Example 3.2.12. For further plots of the generalized Gaussian window with different parameters we refer to Figure 4.6.*



(a) $k = 1$, $\beta = 3$ (b) $k = 3$, $\beta = 3$ (c) $k = 4$, $\beta = 4$

Figure 4.6: Window functions of the generalized Gaussian filter with different parameters.



(a) $\gamma = 2$ (b) $\gamma = 3$ (c) $\gamma = 4$

(d) $\gamma = 4.5$ (e) $\gamma = 5$ (f) $\gamma = 6$

Figure 4.7: Decay rate of $\Phi_{\gamma,W}(L)$ for the generalized Gaussian filter with $k = 4$, $\beta = 4$.

For even $k \in \mathbb{N}$ we have $W \in \mathscr{C}^k([-1,1])$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1 \quad \text{and} \quad W^{(k)}(0) = -k! \left(\frac{\pi}{\beta}\right)^k \neq 0.$$

Consequently, our theory in Theorem 4.3.4 predicts that the error term $\Phi_{\gamma,W}(L)$ behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty.$$

On the other hand, for odd $k \in \mathbb{N}$ we only have $W \in \mathscr{C}^{k-1}([-1,1])$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1,$$

but $W^{(k-1)}$ is not differentiable at 0, as this can be seen in Figure 4.6(a) for the case $k = 1$. Therefore, our theory now predicts that $\Phi_{\gamma,W}(L)$ behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k-1,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty.$$

In our numerical experiments, we evaluated $\Phi_{\gamma,W}(L)$ as a function of the bandwidth $L > 0$ for the Gaussian window W using various combinations of parameters $k \in \mathbb{N}$, $\beta > 1$ and $\gamma > 0$.

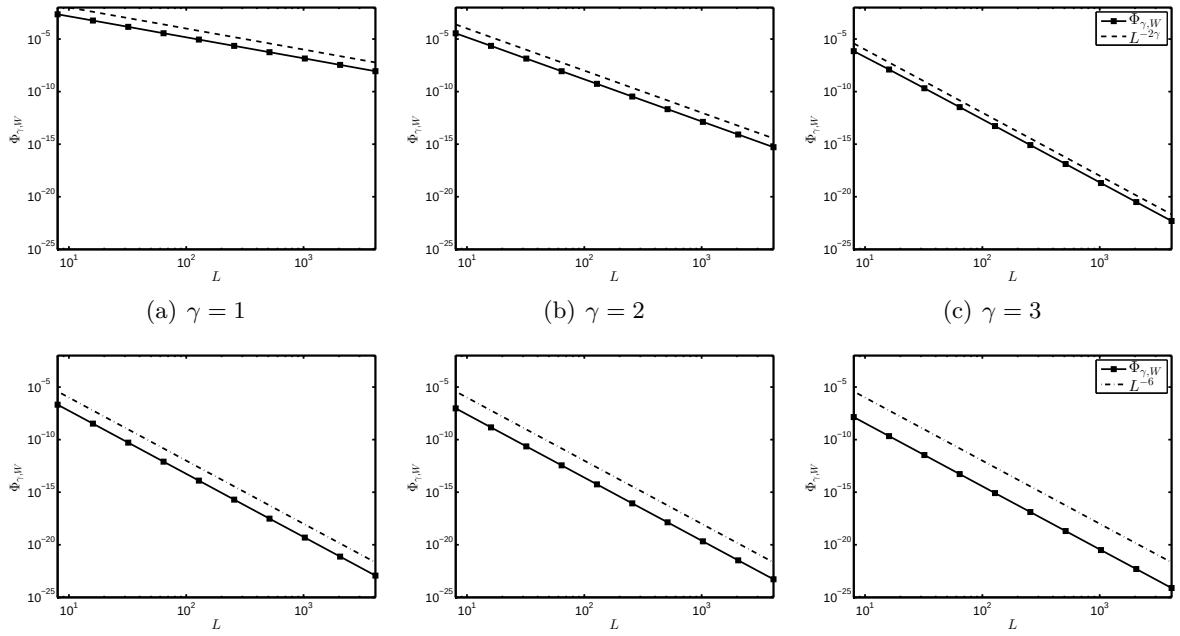
Figure 4.7 shows the behaviour of $\Phi_{\gamma,W}(L)$ in logarithmic scales for the generalized Gaussian filter with $k = 4$, $\beta = 4$ and for the parameters $\gamma \in \{2, 3, 4, 4.5, 5, 6\}$. For $\gamma \in \{2, 3, 4\}$ we observe that $\Phi_{\gamma,W}(L)$ exactly behaves as $L^{-2\gamma}$, whereas for $\gamma \in \{4.5, 5, 6\}$ the behaviour of $\Phi_{\gamma,W}(L)$ corresponds to L^{-8}. But $\Phi_{\gamma,W}(L)$ continues to decrease at increasing $\gamma > k$.

We can summarize the results of our numerical experiments for any even $k \in \mathbb{N}$ as follows. For $\gamma \le k$ we observe

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty,$$

whereas for $\gamma > k$ the convergence order of $\Phi_{\gamma,W}(L)$ saturates at rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2k}) \quad \text{for} \quad L \longrightarrow \infty.$$



(a) $\gamma = 1$ (b) $\gamma = 2$ (c) $\gamma = 3$

(d) $\gamma = 3.5$ (e) $\gamma = 4$ (f) $\gamma = 6$

Figure 4.8: Decay rate of $\Phi_{\gamma,W}(L)$ for the generalized Gaussian filter with $k = 3$, $\beta = 4$.

Note that the results of Theorem 4.3.4 entirely comply with our numerical observations for the generalized Gaussian filters with even $k \in \mathbb{N}$. So have we, in particular, observed the saturation of the convergence rate of $\Phi_{\gamma,W}(L)$ for $\gamma > k$ at

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2k}) \quad \text{for} \quad L \longrightarrow \infty.$$

Thus, our numerical results show that the proven convergence order of the error term $\Phi_{\gamma,W}(L)$ is optimal for $\mathscr{C}^k$-window functions.

Figure 4.8 now shows the behaviour of $\Phi_{\gamma,W}(L)$ in logarithmic scales for the generalized Gaussian filter with $k = 3$, $\beta = 4$ and for the parameters $\gamma \in \{1, 2, 3, 3.5, 4, 6\}$. For $\gamma \in \{1, 2, 3\}$ we observe that $\Phi_{\gamma,W}(L)$ behaves as $L^{-2\gamma}$, whereas for $\gamma \in \{3.5, 4, 6\}$ the behaviour of $\Phi_{\gamma,W}(L)$ corresponds to L^{-6}. But $\Phi_{\gamma,W}(L)$ continues to decrease at increasing $\gamma > k$.

We can summarize the results of our numerical experiments also for odd $k \in \mathbb{N}$ as follows. For $\gamma \leq k$ we observe

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty,$$

while for $\gamma > k$ the convergence order of $\Phi_{\gamma,W}(L)$ saturates at rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2k}) \quad \text{for} \quad L \longrightarrow \infty.$$

We remark that our numerical results for the generalized Gaussian filters with odd $k \in \mathbb{N}$ show a better convergence behaviour of the error term $\Phi_{\gamma,W}(L)$ than predicted by Theorem 4.3.4. Indeed, the above observations correspond to the theoretical results for $\mathscr{C}^k$-window functions, although $W^{(k-1)}$ is not differentiable at 0. However, $W^{(k-1)}$ is Lipschitz continuous on $[-1, 1]$ so that its pointwise derivative $W^{(k)}$ exists almost everywhere and satisfies $W^{(k)} \in \mathrm{L}^\infty([-1, 1])$.

4.3.2 Convergence rates for Lipschitz-windows

In the previous Section 4.3.1 we have analysed the convergence behaviour of the inherent FBP reconstruction error $e_L = f - f_L$ in (4.7) by considering the error term

$$\Phi_{\gamma,W}(L) = \sup_{S \in [-1,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\gamma} \quad \text{for } L > 0$$

for the special case of $\mathscr{C}^k$-window functions. In particular, for $W \in \mathscr{C}^k([-1,1])$ satisfying

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1 \quad \text{and} \quad W^{(k)}(0) \neq 0$$

we have observed that the convergence rate of the error bound saturates at $\mathcal{O}(L^{-2k})$ for $\gamma > k$.

Note that if $W^{(k-1)} \in \mathscr{C}([-1,1])$ is not continuously differentiable on $[-1,1]$ but satisfies

$$W^{(k-1)}(0) = 0,$$

the theory predicts saturation of the order of convergence at rate $\mathcal{O}(L^{-2(k-1)})$ for $\gamma > k-1$. However, in our numerical experiments in Example 4.3.6 for the generalized Gaussian filter with odd $k \in \mathbb{N}$ we observe saturation at rate $\mathcal{O}(L^{-2k})$ for $\gamma > k$, although $W^{(k-1)} \in \mathscr{C}([-1,1])$ is not differentiable at the origin.

Therefore, in the following we relax the assumption of our error theory from Section 4.3.1 so that we can predict the convergence behaviour of $\Phi_{\gamma,W}(L)$ correctly for a larger class of low-pass filters. To this end, we now consider even window functions $W \in \mathrm{L}^\infty(\mathbb{R})$ with compact support $\operatorname{supp}(W) \subseteq [-1,1]$ that are absolutely continuous on $[-1,1]$, i.e., $W \in \mathcal{AC}([-1,1])$. We refer to Appendix A.4 for the definition and some basic properties. In particular, this implies that W is pointwise differentiable almost everywhere on $[-1,1]$ with an integrable derivative $W' \in \mathrm{L}^1([-1,1])$ and the fundamental theorem of calculus for $\mathcal{AC}$-functions, Theorem A.4.4, yields

$$W(S) = W(-1) + \int_{-1}^{S} W'(t)\,\mathrm{d}t \quad \forall\, S \in [-1,1].$$

Furthermore, for some $k \in \mathbb{N}$ we assume that W satisfies $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $1 \le j \le k-1$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1$$

and that $W^{(k)} \in \mathrm{L}^\infty([-1,1])$ is essentially bounded on $[-1,1]$. Equivalently, we can also assume that $W^{(k-1)}$ is Lipschitz continuous on $[-1,1]$, $W^{(k-1)} \in \mathscr{C}^{0,1}([-1,1])$, as explained in the following remark, see also Remark A.4.6. That is why we henceforth speak of Lipschitz-windows.

Remark 4.3.7. *If $W \in \mathcal{AC}([-1,1])$ is absolutely continuous on $[-1,1]$ with $W' \in \mathrm{L}^\infty([-1,1])$, then W is in fact Lipschitz continuous on $[-1,1]$, $W \in \mathscr{C}^{0,1}([-1,1])$, with Lipschitz constant*

$$|W|_{\mathscr{C}^{0,1}([-1,1])} \le \|W'\|_{\mathrm{L}^\infty([-1,1])}.$$

Conversely, if $W \in \mathscr{C}^{0,1}([-1,1])$ is Lipschitz continuous on the interval $[-1,1]$, then W is also absolutely continuous on $[-1,1]$, $W \in \mathcal{AC}([-1,1])$, and its derivative W', which exists almost everywhere on $[-1,1]$, is essentially bounded on $[-1,1]$, i.e., $W' \in \mathrm{L}^\infty([-1,1])$, with

$$\|W'\|_{\mathrm{L}^\infty([-1,1])} \le |W|_{\mathscr{C}^{0,1}([-1,1])}.$$

We now prove that Theorem 4.3.4 still holds true under the relaxed assumptions listed above.

Theorem 4.3.8 (Convergence rate of $\Phi_{\gamma,W}$ for Lipschitz-windows)**.** *For $k \in \mathbb{N}$, let the window function W satisfy $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \le j \le k-1$ and $W^{(k)} \in \mathrm{L}^\infty([-1,1])$ with*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1.$$

Then, for $0 \le \gamma \le k$, we have

$$\Phi_{\gamma,W}(L) \le \frac{1}{(k!)^2}\,\|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\,L^{-2\gamma}$$

and, for $\gamma > k$,

$$\Phi_{\gamma,W}(L) \le \begin{cases} \frac{1}{(k!)^2}\,\|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\,L^{-2\gamma} & \text{for } L < L^* \\ \frac{c^2_{\gamma,k}}{(k!)^2}\,\|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\,L^{-2k} & \text{for } L \ge L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\gamma - k}}$ *and the strictly monotonically decreasing constant*

$$c_{\gamma,k} = \Big(\frac{k}{\gamma - k}\Big)^{k/2} \Big(\frac{\gamma - k}{\gamma}\Big)^{\gamma/2} \quad \text{for } \gamma > k.$$

In particular,

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty.$$

Note that any $\mathscr{C}^k$-window $W \in \mathscr{C}^k([-1,1])$ satisfies $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \le j \le k-1$ and $W^{(k)} \in \mathrm{L}^\infty([-1,1])$. In particular, the assumptions of the above Theorem 4.3.8 are weaker than the assumptions of Theorem 4.3.4.

Proof. Since the window W is assumed to be even and continuous on $[-1,1]$, we obtain

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\gamma} \quad \text{for } L > 0.$$

By assumption we have $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \le j \le k-1$ so that $W^{(j+1)}$ exists almost everywhere on $[-1,1]$ with $W^{(j+1)} \in \mathrm{L}^1([-1,1])$ and

$$W^{(j)}(S) = W^{(j)}(0) + \int_0^S W^{(j+1)}(t)\,\mathrm{d}t \quad \forall\, S \in [0,1].$$

If $k = 1$, for $S \in [0,1]$ follows that

$$W(S) = \underbrace{W(0)}_{=1} + \int_0^S W'(t)\,\mathrm{d}t = 1 + \int_0^S W'(t)\,\mathrm{d}t$$

and the assumption $W' \in \mathrm{L}^\infty([-1,1])$ yields

$$|1 - W(S)| = \left|\int_0^S W'(t)\,\mathrm{d}t\right| \le \|W'\|_{\mathrm{L}^\infty([0,1])}\, S \quad \forall\, S \in [0,1].$$

For $k \ge 2$, we can apply integration by parts, Theorem A.4.5, and obtain, for all $S \in [0,1]$,

$$\begin{aligned} W(S) &= W(0) + \int_0^S W'(t)\,\mathrm{d}t = 1 + \int_0^S W'(t)\,\mathrm{d}t = 1 + W'(S)\,S - \int_0^S t\,W''(t)\,\mathrm{d}t \\ &= 1 + S\Big(\underbrace{W'(0)}_{=0} + \int_0^S W''(t)\,\mathrm{d}t\Big) - \int_0^S t\,W''(t)\,\mathrm{d}t = 1 + \int_0^S (S-t)\,W''(t)\,\mathrm{d}t. \end{aligned}$$

By iteratively applying integration by parts follows that

$$\begin{aligned} W(S) &= 1 + \Big[-\frac{1}{2}\,W''(t)\,(S-t)^2\Big]_{t=0}^{t=S} + \frac{1}{2}\int_0^S (S-t)^2\,W'''(t)\,\mathrm{d}t \\ &= 1 + \frac{1}{2}\,\underbrace{W''(0)}_{=0}\,S^2 + \frac{1}{2}\int_0^S (S-t)^2\,W'''(t)\,\mathrm{d}t = 1 + \frac{1}{2}\int_0^S (S-t)^2\,W'''(t)\,\mathrm{d}t \\ &= \dots \\ &= 1 + \frac{1}{(k-1)!}\int_0^S (S-t)^{k-1}\,W^{(k)}(t)\,\mathrm{d}t. \end{aligned}$$

Consequently, the assumption $W^{(k)} \in \mathrm{L}^\infty([-1,1])$ implies that, for all $S \in [0,1]$,

$$\begin{aligned} |1 - W(S)| &= \left|\frac{1}{(k-1)!}\int_0^S (S-t)^{k-1}\,W^{(k)}(t)\,\mathrm{d}t\right| \le \frac{1}{(k-1)!}\,\|W^{(k)}\|_{\mathrm{L}^\infty([0,1])}\int_0^S (S-t)^{k-1}\,\mathrm{d}t \\ &= \frac{1}{k!}\,\|W^{(k)}\|_{\mathrm{L}^\infty([0,1])}\,S^k. \end{aligned}$$

Hence, for any $k \in \mathbb{N}$ we have

$$|1 - W(S)| \leq \frac{1}{k!} \|W^{(k)}\|_{\mathrm{L}^\infty([0,1])} S^k \quad \forall\, S \in [0,1]$$

and, thus, the error term $\Phi_{\gamma,W}(L)$ can be bounded above by

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \leq \frac{1}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])} \max_{S\in[0,1]} \frac{S^{2k}}{(1+L^2S^2)^\gamma}.$$

In Lemma 4.3.3 we have shown that

$$\max_{S\in[0,1]} \frac{S^{2k}}{(1+L^2S^2)^\gamma} \leq \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq k \vee (\gamma > k \wedge L < L^*) \\ c^2_{\gamma,k}\, L^{-2k} & \text{for } \gamma > k \wedge L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\gamma-k}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k} = \Big(\frac{k}{\gamma-k}\Big)^{k/2} \Big(\frac{\gamma-k}{\gamma}\Big)^{\gamma/2} \qquad \text{for } \gamma > k.$$

Combining our estimates yields

$$\begin{aligned} \Phi_{\gamma,W}(L) &\leq \frac{1}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])} \max_{S\in[-1,1]} \frac{S^{2k}}{(1+L^2S^2)^\gamma} \\ &\leq \begin{cases} \frac{1}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\, L^{-2\gamma} & \text{for } \gamma \leq k \vee (\gamma > k \wedge L < L^*) \\ \frac{c^2_{\gamma,k}}{(k!)^2} \|W^{(k)}\|^2_{\mathrm{L}^\infty([0,1])}\, L^{-2k} & \text{for } \gamma > k \wedge L \geq L^* \end{cases} \end{aligned}$$

and the proof is complete. □

We remark that the generalized Gaussian filter from Example 4.3.6 satisfies the assumptions of Theorem 4.3.8 for all choices of the parameters $k \in \mathbb{N}$ and $\beta > 1$. In particular, our error theory now predicts that the error term $\Phi_{\gamma,W}(L)$ behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty$$

for both even and odd $k \in \mathbb{N}$, as observed numerically in Example 4.3.6. Thus, the theoretical results of Theorem 4.3.8 now totally comply with our numerical observations.

Finally, combining the Theorems 4.2.3 and 4.3.8 shows that the H^σ-error estimate of the inherent FBP reconstruction error in Corollary 4.3.5 is still valid under the relaxed assumptions from above.

Corollary 4.3.9 (H^σ-error estimate for Lipschitz-windows)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha > 0$ and, for $k \in \mathbb{N}$, let the window W satisfy $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \leq j \leq k-1$ and $W^{(k)} \in \mathrm{L}^\infty([-1,1])$ with*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Then, for $0 \leq \sigma \leq \alpha$, the H^σ-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by

$$\|e_L\|_\sigma \leq \left(\frac{1}{k!} \|W^{(k)}\|_{\mathrm{L}^\infty([0,1])} + 1\right) L^{\sigma-\alpha} \|f\|_\alpha$$

if $\alpha - \sigma \leq k$ and, if $\alpha - \sigma > k$, by

$$\|e_L\|_\sigma \leq \begin{cases} \Big(\frac{1}{k!} \|W^{(k)}\|_{\mathrm{L}^\infty([0,1])} + 1\Big) L^{\sigma-\alpha} \|f\|_\alpha & \text{for } L < L^* \\ \Big(\frac{c_{\alpha-\sigma,k}}{k!} \|W^{(k)}\|_{\mathrm{L}^\infty([0,1])}\, L^{-k} + L^{\sigma-\alpha}\Big) \|f\|_\alpha & \text{for } L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\alpha-\sigma-k}}$ *and the strictly monotonically decreasing constant*

$$c_{\alpha-\sigma,k} = \Big(\frac{k}{\alpha-\sigma-k}\Big)^{k/2} \Big(\frac{\alpha-\sigma-k}{\alpha-\sigma}\Big)^{(\alpha-\sigma)/2} \qquad \text{for } \alpha-\sigma > k.$$

In particular,

$$\|e_L\|_\sigma = \mathcal{O}\Big(L^{-\min\{k,\alpha-\sigma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty. \qquad \square$$

As in Corollary 4.3.5, the convergence rate in Corollary 4.3.9 is determined by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error e_L is measured, as long as $\alpha - \sigma \leq k$. But for $\alpha - \sigma > k$ the order of convergence saturates at rate $\mathcal{O}(L^{-k})$. However, in this case the error bound still decreases at increasing α, since the involved constant $c_{\alpha-\sigma,k}$ is strictly monotonically decreasing in $\alpha-\sigma > k$.

We close this section by remarking that with Theorem 4.3.8 we are now able to treat the generalized Ramp filter from Example 4.3.2 also for the width $\beta = 0$, in which case its window function W is given by

$$W(S) = \begin{cases} 1 - (1-\lambda)\,|S| & \text{for } |S| \leq 1 \\ 0 & \text{for } |S| > 1 \end{cases}$$

with jump height $\lambda \in [0,1)$. Figure 4.9 shows some plots of the window for $\lambda \in \{0.25, 0.5, 0.75\}$.

We observe that the generalized Ramp window of width $\beta = 0$ is Lipschitz continuous on the interval $[-1,1]$ and satisfies $W(0) = 1$, but it is not differentiable at $S = 0$. Consequently, the assumptions of Theorem 4.3.8 are satisfies with $k = 1$ and our error theory predicts that the error term $\Phi_{\gamma,W}(L)$ behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{1,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty.$$

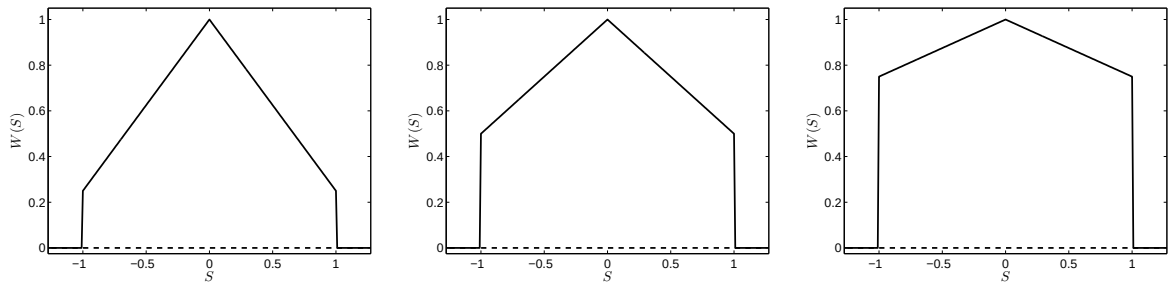
In particular, the convergence order of $\Phi_{\gamma,W}(L)$ is prognosticated to stagnate for $\gamma > 1$ at rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2}) \quad \text{for} \quad L \longrightarrow \infty.$$

In contrast to that, recall that for positive width $\beta \in (0,1)$ and any jump height $\lambda \in [0,1)$ Assumption (A) is satisfied for all parameters $\gamma > 0$ so that $\Phi_{\gamma,W}(L)$ is always of order

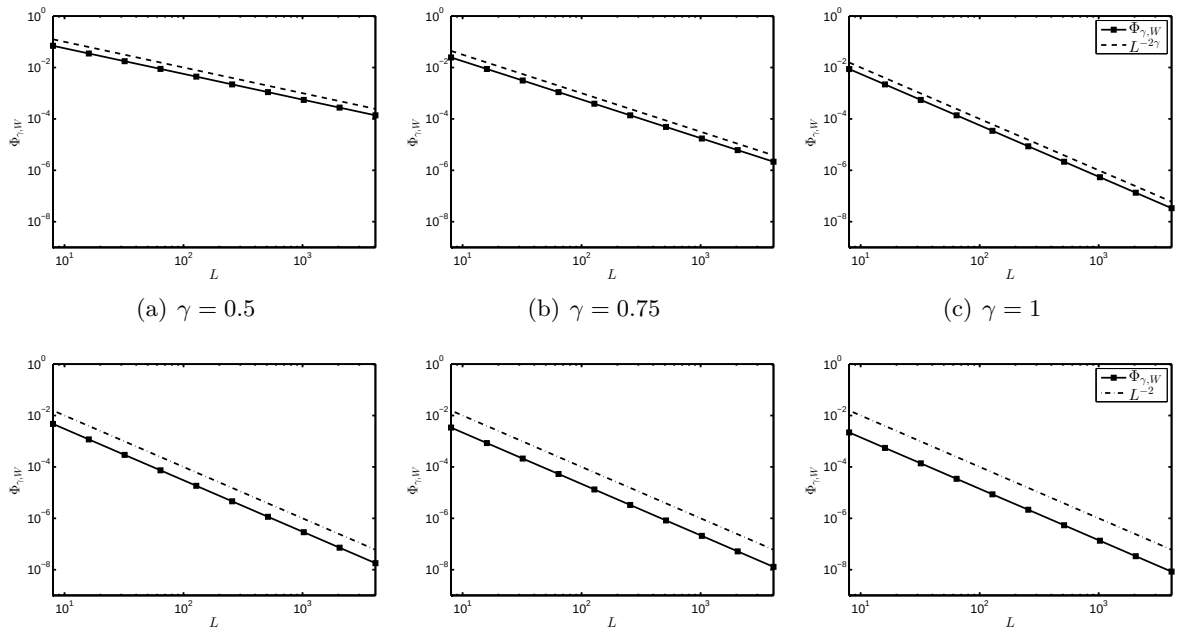
$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty.$$

We calculated $\Phi_{\gamma,W}(L)$ numerically for the above window function W with different jump heights $\lambda \in [0,1)$ and for different parameters $\gamma > 0$. As an example, Figure 4.10 shows the behaviour of $\Phi_{\gamma,W}(L)$ in logarithmic scales for $\lambda = 0.25$ and $\gamma \in \{0.5, 0.75, 1, 1.25, 1.5, 2\}$. We observe that $\Phi_{\gamma,W}(L)$ exactly behaves as $L^{-2\gamma}$ for $\gamma \in \{0.5, 0.75, 1\}$, whereas for $\gamma \in \{1.25, 1.5, 2\}$ the behaviour of $\Phi_{\gamma,W}(L)$ corresponds to L^{-2}. But $\Phi_{\gamma,W}(L)$ still decreases at increasing $\gamma > 1$.



(a) $\lambda = 0.25$ (b) $\lambda = 0.5$ (c) $\lambda = 0.75$

Figure 4.9: Window functions of the generalized Ramp filter with $\beta = 0$ and different $\lambda \in [0,1)$.



(a) $\gamma = 0.5$ (b) $\gamma = 0.75$ (c) $\gamma = 1$

(d) $\gamma = 1.25$ (e) $\gamma = 1.5$ (f) $\gamma = 2$

Figure 4.10: Decay rate of $\Phi_{\gamma,W}(L)$ for the generalized Ramp filter with $\beta = 0$, $\lambda = 0.25$.

We can summarize the results of our numerical experiments for any jump $\lambda \in [0, 1)$ as follows. For $\gamma \leq 1$ we observe

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty,$$

while for $\gamma > 1$ the convergence order of $\Phi_{\gamma,W}(L)$ saturates at rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2}) \quad \text{for} \quad L \longrightarrow \infty.$$

In conclusion, our numerical observations for the generalized Ramp filter of width $\beta = 0$ totally comply with our theoretical results in Theorem 4.3.8 for Lipschitz-window functions with $k = 1$.

4.3.3 Convergence rates for $\mathcal{AC}$-windows with L^p-derivatives

We now generalize the results of the previous Section 4.3.2 by considering even window functions $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\mathrm{supp}(W) \subseteq [-1, 1]$ that satisfy $W^{(j)} \in \mathcal{AC}([-1, 1])$ for all $0 \leq j \leq k-1$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1$$

as well as $W^{(k)} \in \mathrm{L}^p([-1, 1])$ for some $1 < p < \infty$ and $k \in \mathbb{N}$. In this case, we will show that the error term

$$\Phi_{\gamma,W}(L) = \sup_{S \in [-1,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\gamma} \quad \text{for } L > 0$$

behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k-1/p,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty$$

so that its order of convergence saturates for $\gamma > k - \frac{1}{p}$ at possibly fractional rate $\mathcal{O}(L^{-2(k-1/p)})$.

Theorem 4.3.10 (Convergence rate of $\Phi_{\gamma,W}$ for $\mathcal{AC}$-windows with L^p-derivatives)**.** *For $k \in \mathbb{N}$, let the window function W satisfy $W^{(j)} \in \mathcal{AC}([-1, 1])$ for all $0 \leq j \leq k-1$ and $W^{(k)} \in \mathrm{L}^p([-1, 1])$ for some $1 < p < \infty$ with*

$$W(0) = 1 \quad \textit{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Then, for $\gamma \leq k - \frac{1}{p}$*, we have*

$$\Phi_{\gamma,W}(L) \leq \frac{1}{((k-1)!)^2} \left(\frac{1-1/p}{k-1/p}\right)^{2(1-1/p)} \|W^{(k)}\|^2_{\mathrm{L}^p([0,1])}\, L^{-2\gamma}$$

and, for $\gamma > k - \frac{1}{p}$*,*

$$\Phi_{\gamma,W}(L) \leq \begin{cases} \frac{1}{((k-1)!)^2} \left(\frac{1-1/p}{k-1/p}\right)^{2(1-1/p)} \|W^{(k)}\|^2_{\mathrm{L}^p([0,1])}\, L^{-2\gamma} & \text{for } L < L^* \\ \frac{c^2_{\gamma,k,p}}{((k-1)!)^2} \left(\frac{1-1/p}{k-1/p}\right)^{2(1-1/p)} \|W^{(k)}\|^2_{\mathrm{L}^p([0,1])}\, L^{-2(k-1/p)} & \text{for } L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k-1/p}}{\sqrt{\gamma-k+1/p}}$ *and the strictly monotonically decreasing constant*

$$c_{\gamma,k,p} = \left(\frac{k-1/p}{\gamma-k+1/p}\right)^{(k-1/p)/2} \left(\frac{\gamma-k+1/p}{\gamma}\right)^{\gamma/2} \quad \text{for } \gamma > k - \frac{1}{p}.$$

In particular,

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k-1/p,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty.$$

Proof. Since the window function W is assumed to be even and continuous on $[-1,1]$, the error term $\Phi_{\gamma,W}(L)$ can be expressed as

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } L > 0.$$

By assumption we further have $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \leq j \leq k-1$ implying that $W^{(j+1)}$ exists almost everywhere on $[-1,1]$ with $W^{(j+1)} \in \mathrm{L}^1([-1,1])$ and

$$W^{(j)}(S) = W^{(j)}(0) + \int_0^S W^{(j+1)}(t)\,\mathrm{d}t \quad \forall\, S \in [0,1].$$

As in the proof of Theorem 4.3.8, we (iteratively) apply integration by parts and obtain

$$W(S) = 1 + \frac{1}{(k-1)!} \int_0^S (S-t)^{k-1}\, W^{(k)}(t)\,\mathrm{d}t \quad \forall\, S \in [0,1]$$

by using

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Since $W^{(k)} \in \mathrm{L}^p([-1,1])$ for some $1 < p < \infty$, Hölder's inequality, Theorem A.4.1, gives

$$\begin{aligned} |1 - W(S)| &= \frac{1}{(k-1)!} \left| \int_0^S (S-t)^{k-1}\, W^{(k)}(t)\,\mathrm{d}t \right| \leq \frac{1}{(k-1)!} \int_0^S (S-t)^{k-1}\, |W^{(k)}(t)|\,\mathrm{d}t \\ &\leq \frac{1}{(k-1)!} \left(\int_0^S (S-t)^{q(k-1)}\,\mathrm{d}t \right)^{1/q} \left(\int_0^S |W^{(k)}(t)|^p\,\mathrm{d}t \right)^{1/p}, \end{aligned}$$

where $1 < q < \infty$ is the Hölder conjugate of p satisfying

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \Longleftrightarrow \quad q = \frac{p}{p-1}.$$

Furthermore, for $S \in [0,1]$ we have

$$\begin{aligned} \int_0^S (S-t)^{q(k-1)}\,\mathrm{d}t &= -\frac{1}{q(k-1)+1} \Big[(S-t)^{q(k-1)+1}\Big]_{t=0}^{t=S} = \frac{1}{q(k-1)+1}\, S^{q(k-1)+1} \\ &= \frac{p-1}{pk-1}\, S^{q(k-1)+1} = \frac{1-1/p}{k-1/p}\, S^{q(k-1)+1}. \end{aligned}$$

Hence, it follows that

$$|1 - W(S)| \leq \frac{1}{(k-1)!} \left(\frac{1 - 1/p}{k - 1/p}\right)^{1-1/p} S^{k-1/p} \, \|W^{(k)}\|_{\mathrm{L}^p([0,1])} \quad \forall\, S \in [0,1]$$

and, thus, the error term $\Phi_{\gamma,W}(L)$ can be bounded above by

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1 - W(S))^2}{(1 + L^2S^2)^\gamma} \leq \frac{1}{((k-1)!)^2} \left(\frac{1 - 1/p}{k - 1/p}\right)^{2-2/p} \|W^{(k)}\|^2_{\mathrm{L}^p([0,1])} \max_{S\in[0,1]} \frac{S^{2(k-1/p)}}{(1 + L^2S^2)^\gamma}.$$

Finally, applying Lemma 4.3.3 shows that

$$\max_{S\in[0,1]} \frac{S^{2(k-1/p)}}{(1 + L^2S^2)^\gamma} \leq \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq k - \frac{1}{p} \vee \left(\gamma > k - \frac{1}{p} \wedge L < L^*\right) \\ c^2_{\gamma,k,p}\, L^{-2(k-1/p)} & \text{for } \gamma > k - \frac{1}{p} \wedge L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k-1/p}}{\sqrt{\gamma-k+1/p}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k,p} = \left(\frac{k - 1/p}{\gamma - k + 1/p}\right)^{(k-1/p)/2} \left(\frac{\gamma - k + 1/p}{\gamma}\right)^{\gamma/2} \quad \text{for } \gamma > k - \frac{1}{p}.$$

Combining the estimates completes the proof. □

We remark that the error estimates in Theorem 4.3.10, where $W^{(k)} \in \mathrm{L}^p([-1,1])$ for some $1 < p < \infty$, are consistent with the error estimates in Theorem 4.3.8, where $W^{(k)} \in \mathrm{L}^\infty([-1,1])$. Indeed, since

$$c_{\gamma,k,p} = \left(\frac{k - 1/p}{\gamma - k + 1/p}\right)^{(k-1/p)/2} \left(\frac{\gamma - k + 1/p}{\gamma}\right)^{\gamma/2} = (k - 1/p)^{(k-1/p)/2} \, (\gamma - k + 1/p)^{(\gamma-k+1/p)/2} \, \gamma^{-\gamma/2}$$

with

$$(k - 1/p)^{k-1/p} \longrightarrow k^k \quad \text{for} \quad p \longrightarrow \infty$$

and

$$(\gamma - k + 1/p)^{\gamma-k+1/p} \longrightarrow (\gamma - k)^{\gamma-k} \quad \text{for} \quad p \longrightarrow \infty,$$

we have

$$c_{\gamma,k,p} = \left(\frac{k - 1/p}{\gamma - k + 1/p}\right)^{(k-1/p)/2} \left(\frac{\gamma - k + 1/p}{\gamma}\right)^{\gamma/2} \xrightarrow{p\to\infty} \left(\frac{k}{\gamma - k}\right)^{k/2} \left(\frac{\gamma - k}{\gamma}\right)^{\gamma/2} = c_{\gamma,k}.$$

Moreover,

$$\left(\frac{1 - 1/p}{k - 1/p}\right)^{1-1/p} \longrightarrow \frac{1}{k} \quad \text{for} \quad p \longrightarrow \infty$$

so that in total

$$\frac{1}{((k-1)!)^2} \left(\frac{1 - 1/p}{k - 1/p}\right)^{2(1-1/p)} c^2_{\gamma,k,p} \xrightarrow{p\to\infty} \frac{1}{(k!)^2}\, c^2_{\gamma,k}.$$

Finally, combining the Theorems 4.2.3 and 4.3.10 now yields the following H^σ-error estimates for $\mathcal{AC}$-windows with L^p-derivatives for $1 < p < \infty$, which generalize the H^σ-error estimates in Corollary 4.3.9 for Lipschitz-windows or, in other words, for $\mathcal{AC}$-windows with L^∞-derivatives.

Corollary 4.3.11 (H^σ-error estimate for $\mathcal{AC}$-windows with L^p-derivatives)**.** *For $\alpha > 0$, let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ and, for $k \in \mathbb{N}$, let the window W satisfy $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \leq j \leq k-1$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ for some $1 < p < \infty$ with*

$$W(0) = 1 \quad \textit{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Then, for $0 \leq \sigma \leq \alpha$, *the* H^σ*-norm of the FBP reconstruction error* $e_L = f - f_L$ *is bounded above by*

$$\|e_L\|_\sigma \leq \left(\frac{1}{(k-1)!} \left(\frac{1-1/p}{k-1/p} \right)^{1-1/p} \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1 \right) L^{\sigma-\alpha} \|f\|_\alpha$$

if $\alpha - \sigma \leq k - \frac{1}{p}$ *and, if* $\alpha - \sigma > k - \frac{1}{p}$, *by*

$$\|e_L\|_\sigma \leq \begin{cases} \left(\frac{1}{(k-1)!} \left(\frac{1-1/p}{k-1/p} \right)^{1-1/p} \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1 \right) L^{\sigma-\alpha} \|f\|_\alpha & \text{for } L < L^* \\ \left(\frac{c_{\alpha-\sigma,k,p}}{(k-1)!} \left(\frac{1-1/p}{k-1/p} \right)^{1-1/p} \|W^{(k)}\|_{\mathrm{L}^p([0,1])} L^{-(k-1/p)} + L^{\sigma-\alpha} \right) \|f\|_\alpha & \text{for } L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k-1/p}}{\sqrt{\alpha-\sigma-k+1/p}}$ *and the strictly monotonically decreasing constant*

$$c_{\alpha-\sigma,k,p} = \left(\frac{k-1/p}{\alpha-\sigma-k+1/p} \right)^{(k-1/p)/2} \left(\frac{\alpha-\sigma-k+1/p}{\alpha-\sigma} \right)^{(\alpha-\sigma)/2} \quad \text{for } \alpha - \sigma > k - \frac{1}{p}.$$

In particular,

$$\|e_L\|_\sigma = \mathcal{O}\Big(L^{-\min\{k-1/p,\alpha-\sigma\}} \Big) \quad \text{for} \quad L \longrightarrow \infty. \qquad \square$$

We again observe that the convergence rate of the error bound in Corollary 4.3.11 is determined by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error e_L is measured, as long as $\alpha - \sigma \leq k - \frac{1}{p}$. But for $\alpha - \sigma > k - \frac{1}{p}$ the order of convergence saturates at possibly fractional rate $\mathcal{O}(L^{-(k-1/p)})$. However, in this case the error bound still decreases at increasing α, since the involved constant $c_{\alpha-\sigma,k,p}$ is strictly monotonically decreasing in $\alpha - \sigma > k - \frac{1}{p}$. Thus, although the convergence rate saturates, a smoother target function still allows for a better approximation, as expected.

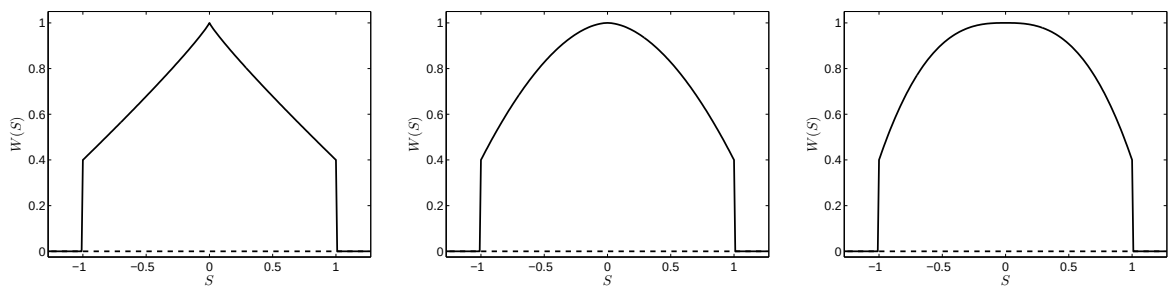
Finally, note that the bound on the inherent FBP reconstruction error in Corollary 4.3.11 is affine-linear with respect to $\|W^{(k)}\|_{\mathrm{L}^p([0,1])}$ so that this quantity can now be used to evaluate the approximation quality of the chosen window function W that satisfies the stated assumptions.

To close this section, we give an example of a window W that satisfies the requirements of our error theory in this section.

Example 4.3.12. *Consider the* generalized polynomial filter *of order* $\nu \in \mathbb{R}_{>0}$ *and with jump height* $\beta \in [0, 1)$, *whose window function is given by*

$$W(S) = \begin{cases} 1 - (1-\beta)\,|S|^\nu & \text{for } |S| \leq 1 \\ 0 & \text{for } |S| > 1. \end{cases}$$

Plots of the generalized polynomial window of different orders can be found in Figure 4.11.



(a) $\nu = 0.9$, $\beta = 0.4$ (b) $\nu = 1.8$, $\beta = 0.4$ (c) $\nu = 2.7$, $\beta = 0.4$

Figure 4.11: Window functions of the generalized polynomial filter with different parameters.

Note that $W \in \mathrm{L}^\infty(\mathbb{R})$ *is even and compactly supported with* $\operatorname{supp}(W) = [-1,1]$ *and that we have* $W \in \mathscr{C}([-1,1])$ *and* $W(0) = 1$.

Case 1: If $0 < \nu \leq 1$, *the pointwise derivative* W' *of* W *exists almost everywhere on* $[-1,1]$, *namely for all* $S \in (-1,1) \setminus \{0\}$ *with*

$$W'(S) = -(1-\beta)\,\nu\,\operatorname{sgn}(S)\,|S|^{\nu-1} \quad \forall\, S \in (-1,1) \setminus \{0\}.$$

Observe that $W' \in \mathrm{L}^1([-1,1])$ *with*

$$\|W'\|_{\mathrm{L}^1([-1,1])} = \nu(1-\beta)\int_{-1}^{1} |S|^{\nu-1}\,\mathrm{d}S = 2\nu(1-\beta)\int_0^1 S^{\nu-1}\,\mathrm{d}S = 2(1-\beta) < \infty.$$

Due to the fundamental theorem of calculus for $\mathcal{AC}$*-functions, Theorem A.4.4, the window* W *is in fact absolutely continuous on* $[-1,1]$, $W \in \mathcal{AC}([-1,1])$, *since, for* $-1 \leq S \leq 0$, *we have*

$$W(-1) + \int_{-1}^{S} W'(t)\,\mathrm{d}t = \beta + \nu(1-\beta)\int_{-1}^{S} (-t)^{\nu-1}\,\mathrm{d}t = 1 - (1-\beta)\,(-S)^\nu = 1 - (1-\beta)\,|S|^\nu = W(S)$$

and, for $0 < S \leq 1$,

$$W(-1) + \int_{-1}^{S} W'(t)\,\mathrm{d}t = \beta + \nu(1-\beta)\int_{-1}^{0} (-t)^{\nu-1}\,\mathrm{d}t - \nu(1-\beta)\int_0^S t^{\nu-1}\,\mathrm{d}t = 1 - (1-\beta)\,S^\nu = W(S).$$

In addition, for $\nu = 1$, *we obtain* $W' \in \mathrm{L}^\infty([-1,1])$ *with*

$$\|W'\|_{\mathrm{L}^\infty([-1,1])} = 1 - \beta = \|W'\|_{\mathrm{L}^\infty([0,1])}$$

and, for $0 < \nu < 1$, *we have* $W' \in \mathrm{L}^p([-1,1])$ *for all* $1 \leq p < \frac{1}{1-\nu}$ *with*

$$\|W'\|^p_{\mathrm{L}^p([-1,1])} = \nu^p(1-\beta)^p \int_{-1}^{1} |S|^{p(\nu-1)}\,\mathrm{d}S = 2\nu^p(1-\beta)^p \int_0^1 S^{p(\nu-1)}\,\mathrm{d}S = \frac{2\nu^p(1-\beta)^p}{p(\nu-1)+1} < \infty.$$

Consequently,

$$\|W'\|_{\mathrm{L}^p([0,1])} = \frac{\nu(1-\beta)}{\sqrt[p]{p(\nu-1)+1}} \quad \forall\, 1 \leq p < \frac{1}{1-\nu}.$$

Case 2: For $\nu > 1$, *we split* ν *into an integer and a fractional part, i.e.,*

$$\nu = k + \eta$$

with $k \in \mathbb{N}$ *and* $0 < \eta \leq 1$. *Then, the window* W *is* k*-times continuously differentiable in the interval* $[-1,1]$, $W \in \mathscr{C}^k([-1,1])$, *and the derivatives* $W^{(j)}$, *for* $1 \leq j \leq k$, *are given by*

$$W^{(j)}(S) = -j! \binom{\nu}{j} (1-\beta)\,(\operatorname{sgn}(S))^j\,|S|^{\nu-j} \quad \forall\, S \in (-1,1).$$

Thus, similar calculations as in Case 1 show that $W^{(j)} \in \mathcal{AC}([-1,1])$ *for all* $0 \leq j \leq k$ *with*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k.$$

In particular,

$$W^{(k)}(S) = -k! \binom{\nu}{k} (1-\beta)\,(\operatorname{sgn}(S))^k\,|S|^{\eta} \quad \forall\, S \in (-1,1)$$

so that, as in Case 1, for $\eta = 1$ *we have* $W^{(k+1)} \in \mathrm{L}^\infty([-1,1])$ *with*

$$\|W^{(k+1)}\|_{\mathrm{L}^\infty([-1,1])} = (k+1)!\,(1-\beta) = \|W^{(k+1)}\|_{\mathrm{L}^\infty([0,1])}.$$

On the other hand, for $0 < \eta < 1$ *we have* $W^{(k+1)} \in \mathrm{L}^p([-1,1])$ *for all* $1 \le p < \frac{1}{1-\eta}$ *with*

$$\begin{aligned}\|W^{(k+1)}\|^p_{\mathrm{L}^p([-1,1])} &= 2(1-\beta)^p((k+1)!)^p \binom{\nu}{k+1}^p \int_0^1 S^{p(\eta-1)}\,\mathrm{d}S \\ &= \frac{2(1-\beta)^p((k+1)!)^p}{p(\eta-1)+1} \binom{\nu}{k+1}^p < \infty\end{aligned}$$

so that

$$\|W^{(k+1)}\|_{\mathrm{L}^p([0,1])} = (1-\beta)\,\frac{\nu\cdot\ldots\cdot(\nu-k)}{\sqrt[p]{p(\nu-k-1)+1}} \quad \forall\, 1 \le p < \frac{1}{k+1-\nu}.$$

When using the generalized polynomial filter of order $\nu > 0$ from Example 4.3.12 in the FBP method, our error theory from Theorems 4.3.8 and 4.3.10 predicts that the error term $\Phi_{\gamma,W}(L)$ behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{\nu,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty$$

and the order of convergence saturates for $\gamma > \nu$ at (possibly) fractional rate $\mathcal{O}(L^{-2\nu})$.

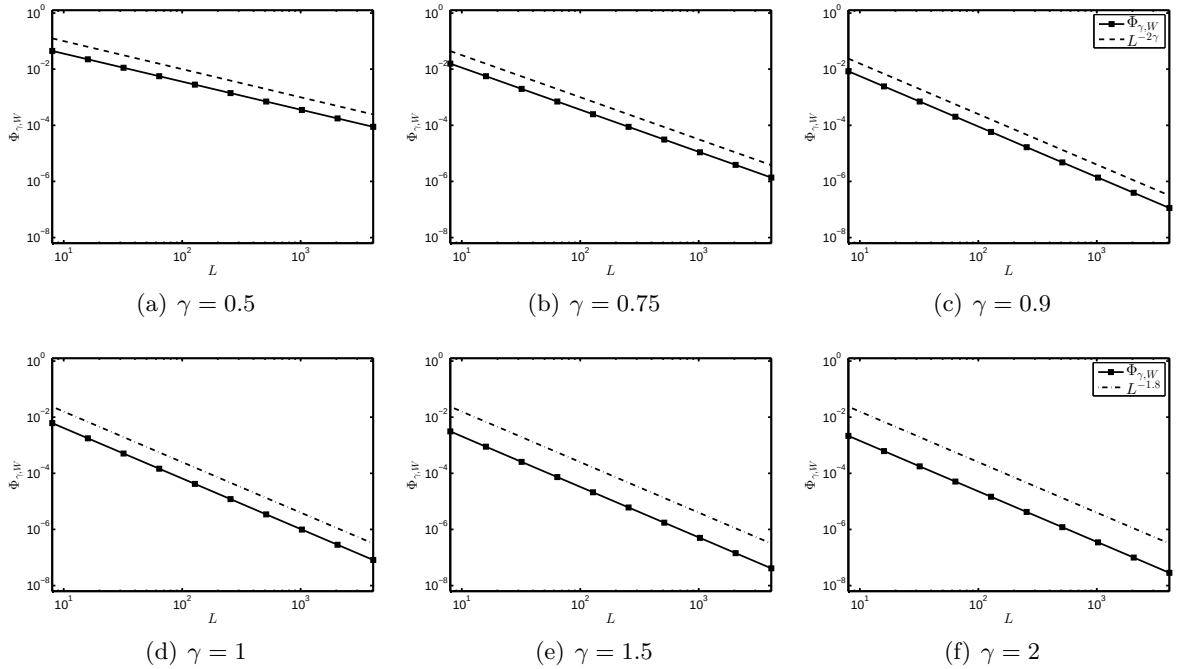
This behaviour was also observed in our numerical experiments, where we calculated $\Phi_{\gamma,W}(L)$ for different choices of $\nu > 0$ and $\beta \in [0,1)$ as well as different parameters $\gamma > 0$. As an example, Figure 4.12 shows $\Phi_{\gamma,W}(L)$ as a function of the bandwidth $L > 0$ in logarithmic scales for the generalized polynomial filter of order $\nu = 0.9$ with jump height $\beta = 0.4$ and for the parameters $\gamma \in \{0.5, 0.75, 0.9, 1, 1.5, 2\}$. For $\gamma \in \{0.5, 0.75, 0.9\}$ we observe that $\Phi_{\gamma,W}(L)$ exactly behaves as $L^{-2\gamma}$, whereas for $\gamma \in \{1, 1.5, 2\}$ the behaviour of $\Phi_{\gamma,W}(L)$ corresponds to $L^{-1.8}$, i.e., $L^{-2\nu}$. But in this case $\Phi_{\gamma,W}(L)$ continues to decrease at increasing $\gamma > 0.9$.

We can summarize our numerical results for any choice of $\nu > 0$ and $\beta \in [0,1)$ as follows. For $\gamma \le \nu$ we observe

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\gamma}) \quad \text{for} \quad L \longrightarrow \infty,$$

while for $\gamma > \nu$ the convergence order of $\Phi_{\gamma,W}(L)$ saturates at fractional rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2\nu}) \quad \text{for} \quad L \longrightarrow \infty.$$



(a) $\gamma = 0.5$ (b) $\gamma = 0.75$ (c) $\gamma = 0.9$

(d) $\gamma = 1$ (e) $\gamma = 1.5$ (f) $\gamma = 2$

Figure 4.12: Decay rate of $\Phi_{\gamma,W}(L)$ for the generalized polynomial filter with $\nu = 0.9$, $\beta = 0.4$.

In conclusion, our numerical observations for the generalized polynomial filter totally comply with our theoretical results in Theorems 4.3.8 and 4.3.10 for $\mathcal{AC}$-windows with L^p-derivatives for $1 < p \leq \infty$ whose first $k-1$ derivatives vanish at the origin. In particular, we have observed saturation of the convergence order of $\Phi_{\gamma,W}(L)$ for $\gamma > k - \frac{1}{p}$ at rate

$$\Phi_{\gamma,W}(L) = \mathcal{O}(L^{-2(k-1/p)}) \quad \text{for} \quad L \longrightarrow \infty$$

and that $\Phi_{\gamma,W}(L)$ continues to decrease at increasing values $\gamma > k - \frac{1}{p}$, which corresponds to the monotonic decrease of the constant $c_{\gamma,k,p}$ in the case $1 < p < \infty$ and $c_{\gamma,k}$ when $p = \infty$.

4.3.4 Convergence rates for Hölder-windows

In the previous Section 4.3.3 we generalized our results from Section 4.3.2 for Lipschitz-windows by considering $\mathcal{AC}$-windows with L^p-derivatives for $1 < p < \infty$, leading to fractional saturation rates for the order of convergence of the inherent FBP reconstruction error $e_L = f - f_L$ in (4.7).

In this section we follow a different approach by applying Hölder-windows. To be more precise, we again consider even window functions $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\mathrm{supp}(W) \subseteq [-1,1]$. But now we assume that, for $k \in \mathbb{N}_0$ and some $0 < \nu \leq 1$, the window W satisfies $W \in \mathscr{C}^{k,\nu}([-1,1])$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k.$$

In other words, W is k-times continuously differentiable on $[-1,1]$, $W \in \mathscr{C}^k([-1,1])$, and $W^{(k)}$ is Hölder continuous on $[-1,1]$, $W^{(k)} \in \mathscr{C}^{0,\nu}([-1,1])$, with Hölder exponent $0 < \nu \leq 1$. Then, there exists a Hölder constant $|W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])} > 0$ such that

$$|W^{(k)}(S) - W^{(k)}(t)| \leq |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])}\, |S-t|^\nu \quad \forall\, S,t \in [0,1].$$

Under these assumption, we will prove that the convergence rate of the error term

$$\Phi_{\gamma,W}(L) = \sup_{S\in[-1,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } L > 0$$

now behaves like

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k+\nu,\gamma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty$$

so that its order of convergence saturates for $\gamma > k+\nu$ at possibly fractional rate $\mathcal{O}(L^{-2(k+\nu)})$.

Theorem 4.3.13 (Convergence rate of $\Phi_{\gamma,W}$ for Hölder-windows)**.** *Let the window function W satisfy $W \in \mathscr{C}^{k,\nu}([-1,1])$ with $k \in \mathbb{N}_0$ and Hölder exponent $0 < \nu \leq 1$ such that*

$$W(0) = 1 \quad \textit{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k.$$

Then, for $\gamma \leq k+\nu$, we have

$$\Phi_{\gamma,W}(L) \leq \left(\frac{\Gamma(\nu+1)}{\Gamma(k+\nu+1)}\right)^2 |W^{(k)}|^2_{\mathscr{C}^{0,\nu}([0,1])}\, L^{-2\gamma}$$

and, for $\gamma > k+\nu$,

$$\Phi_{\gamma,W}(L) \leq \begin{cases} \left(\frac{\Gamma(\nu+1)}{\Gamma(k+\nu+1)}\right)^2 |W^{(k)}|^2_{\mathscr{C}^{0,\nu}([0,1])}\, L^{-2\gamma} & \textit{for } L < L^* \\ c^2_{\gamma,k,\nu} \left(\frac{\Gamma(\nu+1)}{\Gamma(k+\nu+1)}\right)^2 |W^{(k)}|^2_{\mathscr{C}^{0,\nu}([0,1])}\, L^{-2(k+\nu)} & \textit{for } L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k+\nu}}{\sqrt{\gamma-k-\nu}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k,\nu} = \Big(\frac{k+\nu}{\gamma-k-\nu}\Big)^{(k+\nu)/2} \Big(\frac{\gamma-k-\nu}{\gamma}\Big)^{\gamma/2} \quad \textit{for } \gamma > k+\nu.$$

In particular,

$$\Phi_{\gamma,W}(L) = \mathcal{O}\Big(L^{-2\min\{k+\nu,\gamma\}}\Big) \quad \textit{for} \quad L \longrightarrow \infty.$$

Proof. Since the window W is even and continuous on $[-1,1]$, the error term $\Phi_{\gamma,W}(L)$ is given by

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } L>0.$$

By assumption we have $W \in \mathscr{C}^{k,\nu}([-1,1])$ with $k \in \mathbb{N}_0$ and Hölder exponent $0<\nu\leq 1$ which implies that $W \in \mathscr{C}^k([-1,1])$ and

$$|W^{(k)}(S) - W^{(k)}(t)| \leq |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])}\, |S-t|^\nu \quad \forall\, S,t\in[0,1].$$

Thus, if $k=0$, the assumption $W(0)=1$ gives

$$|1-W(S)| \leq |W|_{\mathscr{C}^{0,\nu}([0,1])}\, S^\nu \quad \forall\, S\in[0,1].$$

On the other hand, if $k\geq 1$, by the fundamental theorem of calculus holds that

$$W^{(j)}(S) = W^{(j)}(0) + \int_0^S W^{(j+1)}(t)\,\mathrm{d}t \quad \forall\, 0\leq j\leq k-1.$$

In particular,

$$W(S) = W(0) + \int_0^S W'(t)\,\mathrm{d}t = 1 + \int_0^S W'(t)\,\mathrm{d}t \quad \forall\, S\in[0,1]$$

and, for $k>1$, iteratively applying integration by parts yields

$$W(S) = 1 + \frac{1}{(k-1)!}\int_0^S (S-t)^{k-1}\, W^{(k)}(t)\,\mathrm{d}t \quad \forall\, S\in[0,1],$$

where we used

$$W(0)=1 \quad \text{and} \quad W^{(j)}(0)=0 \quad \forall\, 1\leq j\leq k-1.$$

Since $W^{(k)}$ is Hölder continuous on $[-1,1]$ with Hölder exponent $0<\nu\leq 1$ and $W^{(k)}(0)=0$, we have

$$|W^{(k)}(t)| = |W^{(k)}(t) - W^{(k)}(0)| \leq |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])}\, t^\nu \quad \forall\, t\in[0,1].$$

Consequently, for all $S\in[0,1]$ follows that

$$\begin{aligned}|1-W(S)| &= \left|\frac{1}{(k-1)!}\int_0^S (S-t)^{k-1}\, W^{(k)}(t)\,\mathrm{d}t\right| \leq \frac{1}{(k-1)!}\int_0^S (S-t)^{k-1}\,|W^{(k)}(t)|\,\mathrm{d}t \\ &\leq \frac{1}{(k-1)!}\,|W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])}\int_0^S (S-t)^{k-1}\,t^\nu\,\mathrm{d}t,\end{aligned}$$

where, for $k=1$, we have

$$\int_0^S t^\nu\,\mathrm{d}t = \frac{1}{\nu+1}\,S^{\nu+1} = \frac{\Gamma(\nu+1)}{\Gamma(\nu+2)}\,S^{\nu+1}$$

and, for $k>1$,

$$\begin{aligned}\int_0^S (S-t)^{k-1}\,t^\nu\,\mathrm{d}t &= \underbrace{\left[(S-t)^{k-1}\,\frac{1}{\nu+1}\,t^{\nu+1}\right]_{t=0}^{t=S}}_{=0} + \frac{k-1}{\nu+1}\int_0^S (S-t)^{k-2}\,t^{\nu+1}\,\mathrm{d}t \\ &= \dots \\ &= \frac{(k-1)!}{(\nu+1)\cdot\ldots\cdot(\nu+k-1)}\int_0^S t^{\nu+k-1}\,\mathrm{d}t \\ &= \frac{(k-1)!}{(\nu+1)\cdot\ldots\cdot(\nu+k)}\,S^{\nu+k} = (k-1)!\,\frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)}\,S^{\nu+k}.\end{aligned}$$

Hence, for all $k \in \mathbb{N}_0$ holds that

$$|1 - W(S)| \leq \frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)} \, |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])} \, S^{k+\nu} \quad \forall \, S \in [0,1]$$

and, thus, the error term $\Phi_{\gamma,W}(L)$ is bounded above by

$$\Phi_{\gamma,W}(L) = \max_{S\in[0,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \leq \Big(\frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)}\Big)^2 \, |W^{(k)}|^2_{\mathscr{C}^{0,\nu}([0,1])} \max_{S\in[0,1]} \frac{S^{2(k+\nu)}}{(1+L^2S^2)^\gamma}.$$

Finally, applying Lemma 4.3.3 yields

$$\max_{S\in[0,1]} \frac{S^{2(k+\nu)}}{(1+L^2S^2)^\gamma} \leq \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq k+\nu \vee (\gamma > k+\nu \wedge L < L^*) \\ c^2_{\gamma,k,\nu} \, L^{-2(k+\nu)} & \text{for } \gamma > k+\nu \wedge L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k+\nu}}{\sqrt{\gamma-k-\nu}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k,\nu} = \Big(\frac{k+\nu}{\gamma-k-\nu}\Big)^{(k+\nu)/2} \Big(\frac{\gamma-k-\nu}{\gamma}\Big)^{\gamma/2} \qquad \text{for } \gamma > k+\nu.$$

Combining the estimates completes the proof. □

We remark that the error estimates in Theorem 4.3.13 for Hölder-windows are consistent with the error estimates in Theorem 4.3.8 for Lipschitz-windows, where we have $W \in \mathscr{C}^{k-1,1}([-1,1])$ with

$$|W^{(k-1)}|_{\mathscr{C}^{0,1}([0,1])} = \|W^{(k)}\|_{\mathrm{L}^\infty([0,1])}.$$

Indeed, choosing the differentiability order $k-1$ and Hölder exponent $\nu = 1$ in Theorem 4.3.13 gives

$$\frac{\Gamma(\nu+1)}{\Gamma(k-1+\nu+1)} = \frac{\Gamma(2)}{\Gamma(k+1)} = \frac{1}{k!}$$

and

$$c_{\gamma,k-1,\nu} = \Big(\frac{k-1+\nu}{\gamma-k+1-\nu}\Big)^{(k-1+\nu)/2} \Big(\frac{\gamma-k+1-\nu}{\gamma}\Big)^{\gamma/2} = \Big(\frac{k}{\gamma-k}\Big)^{k/2} \Big(\frac{\gamma-k}{\gamma}\Big)^{\gamma/2} = c_{\gamma,k}$$

so that for Lipschitz-windows the error bounds obtained from Theorem 4.3.13 agree with the estimates in Theorem 4.3.8. In particular, the predicted saturation rates for the convergence order of the error term $\Phi_{\gamma,W}(L)$ coincide.

In addition, we note that the results in Theorem 4.3.13 continue to hold true if we assume that $W^{(k)}$ satisfies a Hölder condition of order $0 < \nu \leq 1$ only at the origin in the sense that there exists a constant $C_{W,k,\nu} > 0$ such that

$$|W^{(k)}(0) - W^{(k)}(S)| \leq C_{W,k,\nu} \, S^\nu \quad \forall \, S \in [0,1].$$

Finally, combining the Theorems 4.2.3 and 4.3.13 gives the following generalized H^σ-error estimate for the approximate FBP reconstruction with Hölder-windows.

Corollary 4.3.14 (H^σ-error estimate for Hölder-windows)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha > 0$ and let the window function W satisfy $W \in \mathscr{C}^{k,\nu}([-1,1])$ with $k \in \mathbb{N}_0$ and Hölder exponent $0 < \nu \leq 1$ such that*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall \, 1 \leq j \leq k.$$

Then, for $0 \leq \sigma \leq \alpha$, the H^σ-norm of the FBP reconstruction error $e_L = f - f_L$ is bounded above by

$$\|e_L\|_\sigma \leq \Big(\frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)} \, |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])} + 1\Big) \, L^{\sigma-\alpha} \, \|f\|_\alpha$$

if $\alpha - \sigma \leq k + \nu$ *and, if* $\alpha - \sigma > k + \nu$, *by*

$$\|e_L\|_\sigma \leq \begin{cases} \Big(\frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)} \, |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])} + 1 \Big) \, L^{\sigma-\alpha} \, \|f\|_\alpha & \text{for } L < L^* \\ \Big(c_{\alpha-\sigma,k,\nu} \, \frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)} \, |W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])} \, L^{-(k+\nu)} + L^{\sigma-\alpha} \Big) \, \|f\|_\alpha & \text{for } L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k+\nu}}{\sqrt{\alpha-\sigma-k-\nu}}$ *and the strictly monotonically decreasing constant*

$$c_{\alpha-\sigma,k,\nu} = \Big(\frac{k+\nu}{\alpha-\sigma-k-\nu} \Big)^{(k+\nu)/2} \Big(\frac{\alpha-\sigma-k-\nu}{\alpha-\sigma} \Big)^{(\alpha-\sigma)/2} \qquad \text{for } \alpha - \sigma > k + \nu.$$

In particular,

$$\|e_L\|_\sigma = \mathcal{O}\Big(L^{-\min\{k+\nu, \alpha-\sigma\}} \Big) \quad \text{for} \quad L \longrightarrow \infty. \qquad \square$$

We observe that the convergence rate of the error bound in Corollary 4.3.14 is determined by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error e_L is measured, as long as $\alpha - \sigma \leq k + \nu$. But for $\alpha - \sigma > k + \nu$ the order of convergence saturates at fractional rate $\mathcal{O}(L^{-(k+\nu)})$. However, in this case the error bound still decreases at increasing α, since the involved constant $c_{\alpha-\sigma,k,\nu}$ is strictly monotonically decreasing in $\alpha - \sigma > k + \nu$. Thus, although the rate of convergence saturates, a smoother target function still allows for a better approximation, as expected.

We close this section by comparing the generalized error estimates in Theorem 4.3.13 for Hölder-windows with the generalized error estimates in Theorem 4.3.10 for $\mathcal{AC}$-windows with L^p-derivatives for $1 < p < \infty$.

Remark 4.3.15 (Special case)**.** *Let the even window* $W \in \mathrm{L}^\infty(\mathbb{R})$ *with* $\mathrm{supp}(W) \subseteq [-1,1]$ *satisfy* $W^{(j)} \in \mathcal{AC}([-1,1])$ *for all* $0 \leq j \leq k$ *with*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k.$$

Further, let $W^{(k+1)} \in \mathrm{L}^p([-1,1])$ *with* $1 < p < \infty$. *Then, we also have* $W \in \mathscr{C}^{k,\nu}([-1,1])$ *with Hölder exponent*

$$\nu = 1 - \frac{1}{p} \in (0,1)$$

and Hölder constant

$$|W^{(k)}|_{\mathscr{C}^{0,\nu}([0,1])} \leq \|W^{(k+1)}\|_{\mathrm{L}^p([0,1])}.$$

Indeed, for $0 \leq t \leq S \leq 1$, *the fundamental theorem of calculus for* $\mathcal{AC}$*-functions, Theorem A.4.4, and Hölder's inequality, Theorem A.4.1, yield*

$$\begin{aligned} |W^{(k)}(S) - W^{(k)}(t)| = \Big| \int_t^S W^{(k+1)}(\tau) \, \mathrm{d}\tau \Big| &\leq \Big(\int_t^S |W^{(k+1)}(\tau)|^p \, \mathrm{d}\tau \Big)^{1/p} \Big(\int_t^S 1 \, \mathrm{d}\tau \Big)^{1-1/p} \\ &\leq \|W^{(k+1)}\|_{\mathrm{L}^p([0,1])} \, |S-t|^{1-1/p}. \end{aligned}$$

Note that the converse is not true in general. As a counterexample we refer to the classical Weierstraß function, which is Hölder continuous but not absolutely continuous on $[-1,1]$.

Therefore, we can apply the error estimates in Theorem 4.3.13 for Hölder-windows to bound the error term $\Phi_{\gamma,W}(L)$ *from above by*

$$\Phi_{\gamma,W}(L) \leq b_{k,\nu}^2 \, \|W^{(k+1)}\|_{\mathrm{L}^p([0,1])}^2 \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq k + \nu \\ c_{\gamma,k,\nu}^2 \, L^{-2(k+\nu)} & \text{for } \gamma > k + \nu \end{cases}$$

with

$$b_{k,\nu} = \frac{\Gamma(\nu+1)}{\Gamma(\nu+k+1)} \qquad \text{for } k \in \mathbb{N}_0 \text{ and } 0 < \nu < 1.$$

On the other hand, using $\nu = 1 - \frac{1}{p}$ *in the error analysis from Theorem 4.3.10 for* $\mathcal{AC}$*-windows with* L^p*-derivatives yields*

$$\Phi_{\gamma,W}(L) \le a_{k,\nu}^2 \, \|W^{(k+1)}\|_{\mathrm{L}^p([0,1])}^2 \begin{cases} L^{-2\gamma} & \text{for } \gamma \le k+\nu \\ c_{\gamma,k,\nu}^2 \, L^{-2(k+\nu)} & \text{for } \gamma > k+\nu \end{cases}$$

with

$$a_{k,\nu} = \frac{1}{k!}\Big(\frac{\nu}{\nu+k}\Big)^{\nu} \quad \text{for } k \in \mathbb{N}_0 \text{ and } 0 < \nu < 1,$$

since

$$c_{\gamma,k+1,p} = \Big(\frac{k+1-1/p}{\gamma-k-1+1/p}\Big)^{(k+1-1/p)/2} \Big(\frac{\gamma-k-1+1/p}{\gamma}\Big)^{\gamma/2} = c_{\gamma,k,\nu}.$$

Consequently, the bounds in the above estimates obtained by Theorems 4.3.10 and 4.3.13 only differ in the constants $a_{k,\nu}$ *and* $b_{k,\nu}$*, which we now compare for all* $k \in \mathbb{N}_0$ *and* $\nu \in (0,1)$*. For* $k = 0$ *we have*

$$a_{0,\nu} = 1 = b_{0,\nu} \quad \forall\, 0 < \nu < 1,$$

whereas for $k \ge 1$ *we will show that*

$$a_{k,\nu} < b_{k,\nu} \quad \forall\, 0 < \nu < 1.$$

To this end, recall that a function $g : I \longrightarrow \mathbb{R}_{\ge 0}$ *is called* completely monotonic *on an interval* $I \subseteq \mathbb{R}$ *if* g *has derivatives of all orders on* I *and, for all* $j \in \mathbb{N}_0$*,*

$$(-1)^j \, g^{(j)}(x) \ge 0 \quad \forall\, x \in I.$$

The set of completely monotonic functions on I *is denoted by* $\mathcal{CM}(I)$*. For further details and properties we refer to [125]. In [28] it is pointed out that if a non-constant function is completely monotonic on* $(0,\infty)$*, then the above inequality is strictly satisfied. In particular, any non-constant function* $g \in \mathcal{CM}((0,\infty))$ *is positive and strictly monotonically decreasing on* $(0,\infty)$*.*

Recall further that a positive function $g : I \longrightarrow \mathbb{R}_{>0}$ *is called* logarithmically completely monotonic *on an interval* $I \subseteq \mathbb{R}$ *if* g *has derivatives of all orders on* I *and, for all* $j \in \mathbb{N}$*, its logarithm* $\log(g)$ *satisfies*

$$(-1)^j \, (\log(g))^{(j)}(x) \ge 0 \quad \forall\, x \in I.$$

The set of logarithmically completely monotonic functions on I *is denoted by* $\mathcal{LCM}(I)$ *and we have*

$$\mathcal{LCM}(I) \subset \mathcal{CM}(I).$$

We refer to [91] for details and an overview of (logarithmically) completely monotonic functions.

For $0 < \nu < 1$ *we now consider the function*

$$h_\nu(x) = \frac{(x+\nu)^{1-\nu}\,\Gamma(x+\nu)}{x\,\Gamma(x)} \quad \text{for } x > 0.$$

In [91, Theorem 1.2] it is proven that h_ν *is logarithmically completely monotonic on* $(0,\infty)$ *for all* $\nu \in (0,1)$*, i.e.,*

$$h_\nu \in \mathcal{LCM}((0,\infty)) \subset \mathcal{CM}((0,\infty)),$$

with

$$\lim_{x \searrow 0} h_\nu(x) = \frac{\Gamma(\nu+1)}{\nu^\nu} \quad \text{and} \quad \lim_{x \to \infty} h_\nu(x) = 1.$$

Thus, our above discussion about non-constant completely monotonic functions on $(0,\infty)$ *gives*

$$h_\nu(x) > h_\nu(y) \quad \forall\, 0 < x < y,$$

which implies that

$$\lim_{y\to\infty} h_\nu(y) < h_\nu(x) < \lim_{y\searrow 0} h_\nu(y) \quad \forall\, x > 0.$$

Consequently, we have

$$1 < \frac{(x+\nu)^{1-\nu}\,\Gamma(x+\nu)}{x\,\Gamma(x)} < \frac{\Gamma(\nu+1)}{\nu^\nu} \quad \forall\, x > 0$$

such that

$$\begin{aligned}\frac{b_{k,\nu}}{a_{k,\nu}} &= \frac{\Gamma(k+1)\,(k+\nu)^\nu}{\Gamma(k+\nu+1)}\,\frac{\Gamma(\nu+1)}{\nu^\nu} > \frac{\Gamma(k+1)\,(k+\nu)^\nu}{\Gamma(k+\nu+1)}\,\frac{(k+\nu)^{1-\nu}\,\Gamma(k+\nu)}{k\,\Gamma(k)} \\ &= \frac{\Gamma(k+1)}{\Gamma(k+\nu+1)}\,\frac{(k+\nu)\,\Gamma(k+\nu)}{k\,\Gamma(k)} = \frac{\Gamma(k+1)}{\Gamma(k+\nu+1)}\,\frac{\Gamma(k+\nu+1)}{\Gamma(k+1)} = 1.\end{aligned}$$

Therefore, for $k \geq 1$ we finally obtain

$$a_{k,\nu} < b_{k,\nu} \quad \forall\, 0 < \nu < 1.$$

In conclusion, we observe that the weaker assumptions in Theorem 4.3.13 for Hölder-windows with $k \in \mathbb{N}_0$ and $\nu \in (0,1)$ lead to the same convergence rates for the error term $\Phi_{\gamma,W}(L)$ we obtain in Theorem 4.3.10 for $\mathcal{AC}$-windows with L^p-derivatives. However, for $k \geq 1$ the involved constant in the error bounds is smaller under the stronger assumptions in Theorem 4.3.10.

4.4 Asymptotic error estimates

In the previous Section 4.3 we derived H^σ-error estimates for the FBP method and analysed the convergence rate of the approximate FBP reconstruction for some special cases, where we assumed a certain kind of *regularity* of the low-pass filter's window W on the whole interval $[-1,1]$. The presented results are obtained by utilizing the H^σ-error estimate (4.10) from Theorem 4.2.3 and by estimating the error term $\Phi_{\gamma,W}$ in (4.9) for $\gamma > 0$ given by

$$\Phi_{\gamma,W}(L) = \operatorname*{ess\,sup}_{S\in[-1,1]} \frac{(1-W(S))^2}{(1+L^2S^2)^\gamma} \quad \text{for } L > 0.$$

In this section we take a different approach to prove *asymptotic* H^σ-error estimates for the FBP method, where we assume regularity of the filter's window function only at the origin. For this purpose, we consider even window functions $W \in \mathrm{L}^\infty(\mathbb{R})$ with compact support in $[-1,1]$ that are k-times differentiable at 0 for some $k \in \mathbb{N}$ with

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

As before, we consider target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with some $\alpha > 0$ and analyse the H^σ-norm of the inherent FBP reconstruction error $e_L = f - f_L$ in (4.7) for all $0 \leq \sigma \leq \alpha$. For the sake of brevity, we again define

$$r(x,y) = \sqrt{x^2+y^2} \quad \text{for } (x,y) \in \mathbb{R}^2$$

and, for $\gamma \geq 0$,

$$r_\gamma(x,y) = (1+r(x,y)^2)^\gamma = (1+x^2+y^2)^\gamma \quad \text{for } (x,y) \in \mathbb{R}^2.$$

As in the derivation of Theorem 4.2.3, we start with splitting the H^σ-norm of e_L into the sum of two integrals

$$\|e_L\|_\sigma^2 = \frac{1}{4\pi^2}\int_{\mathbb{R}}\int_{\mathbb{R}} r_\sigma(x,y)\,|\mathcal{F}(f-f_L)(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y = I_1 + I_2,$$

where

$$I_1 = \frac{1}{4\pi^2} \int_{r(x,y)\le L} r_\sigma(x,y)\,|1 - W_L(x,y)|^2\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y)$$

and

$$I_2 = \frac{1}{4\pi^2} \int_{r(x,y)>L} r_\sigma(x,y)\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y).$$

Recall that, for $0 \le \sigma \le \alpha$, the integral I_2 can be bounded above by

$$I_2 \le L^{2(\sigma-\alpha)}\,\frac{1}{4\pi^2} \int_{r(x,y)>L} r_\alpha(x,y)\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y) \le L^{2(\sigma-\alpha)}\,\|f\|_\alpha^2,$$

whereas the integral I_1 can be expressed as

$$I_1 = \frac{1}{4\pi^2} \int_{r(x,y)\le L} r_\sigma(x,y)\left|1 - W\Big(\frac{r(x,y)}{L}\Big)\right|^2 |\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y).$$

Because $W \in \mathrm{L}^\infty(\mathbb{R})$ is k-times differentiable at zero, we can apply Taylor's Theorem A.4.2 and, thus, there exists a function $h_k : \mathbb{R} \longrightarrow \mathbb{R}$ satisfying

$$W(S) = \sum_{j=0}^{k} \frac{W^{(j)}(0)}{j!}\,S^j + h_k(S)\,S^k \quad \forall\, S \in \mathbb{R}$$

and

$$\lim_{S\to 0} h_k(S) = 0.$$

By assumption, the window W satisfies

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1.$$

Hence, for $(x,y) \in \mathbb{R}^2$ and $L > 0$ follows that

$$W\Big(\frac{r(x,y)}{L}\Big) = 1 + \left(\frac{W^{(k)}(0)}{k!} + h_k\Big(\frac{r(x,y)}{L}\Big)\right)\Big(\frac{r(x,y)}{L}\Big)^k$$

so that we obtain the representation

$$I_1 = \frac{1}{4\pi^2} \int_{r(x,y)\le L} \left(\frac{W^{(k)}(0)}{k!} + h_k\Big(\frac{r(x,y)}{L}\Big)\right)^2 \Big(\frac{r(x,y)}{L}\Big)^{2k} r_\sigma(x,y)\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y).$$

For convenience we now define, for $\gamma \ge 0$,

$$\phi^*_{\gamma,L,k} = \max_{r(x,y)\le L} \frac{\Big(\frac{r(x,y)}{L}\Big)^{2k}}{r_\gamma(x,y)} = \max_{S\in[0,1]} \frac{S^{2k}}{(1+L^2\,S^2)^\gamma}.$$

Then, I_1 can be bounded above by

$$\begin{aligned} I_1 &\le \phi^*_{\alpha-\sigma,L,k}\,\frac{1}{4\pi^2} \int_{r(x,y)\le L} \left(\frac{W^{(k)}(0)}{k!} + h_k\Big(\frac{r(x,y)}{L}\Big)\right)^2 r_\alpha(x,y)\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y) \\ &= \phi^*_{\alpha-\sigma,L,k}\left(I_3 + 2\,\frac{W^{(k)}(0)}{k!}\,I_4 + I_5\right), \end{aligned}$$

where we let

$$I_3 = \frac{1}{4\pi^2} \int_{r(x,y)\le L} \Big(\frac{W^{(k)}(0)}{k!}\Big)^2 r_\alpha(x,y)\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y)$$

and

$$I_j = \frac{1}{4\pi^2} \int_{r(x,y)\le L} \left(h_k\Big(\frac{r(x,y)}{L}\Big)\right)^{j-3} r_\alpha(x,y)\,|\mathcal{F}f(x,y)|^2\,\mathrm{d}(x,y) \quad \text{for } j \in \{4,5\}.$$

To continue our analysis, we observe that, for $S \neq 0$, the function h_k can be written as

$$h_k(S) = (W(S) - 1)\, S^{-k} - \frac{W^{(k)}(0)}{k!}.$$

Since the window W has compact support in $[-1, 1]$, we further obtain

$$h_k(S) = -S^{-k} - \frac{W^{(k)}(0)}{k!} \quad \forall\, |S| > 1,$$

which implies that

$$h_k(S) \longrightarrow -\frac{W^{(k)}(0)}{k!} \quad \text{for} \quad S \longrightarrow \pm\infty.$$

From $W \in \mathrm{L}^\infty(\mathbb{R})$ and

$$h_k(S) \longrightarrow 0 \quad \text{for} \quad S \longrightarrow 0$$

it follows that h_k is essentially bounded on $\mathbb{R}$, so that there exists a constant $M > 0$ satisfying

$$|h_k(S)| \leq M \quad \text{for almost all } S \in \mathbb{R}.$$

Hence, for $j \in \{4, 5\}$ and all $L > 0$ the integrand in I_j,

$$h_{k,L}^j(x, y) = \Big(h_k\Big(\frac{r(x, y)}{L}\Big)\Big)^{j-3} r_\alpha(x, y)\, |\mathcal{F}f(x, y)|^2,$$

is bounded almost everywhere on $\mathbb{R}^2$ by the function

$$\Phi_j(x, y) = M^{j-3}\, r_\alpha(x, y) |\mathcal{F}f(x, y)|^2 \quad \text{for } (x, y) \in \mathbb{R}^2,$$

which is integrable over $\mathbb{R}^2$ due to the assumption $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$. Moreover, we have

$$h_k\Big(\frac{r(x, y)}{L}\Big) \longrightarrow 0 \quad \text{for} \quad \frac{r(x, y)}{L} \longrightarrow 0,$$

implying that, for any $(x, y) \in \mathbb{R}^2$, the integrand $h_{k,L}^j(x, y)$ tends to zero as L goes to ∞. Consequently, we can apply Lebesgue's theorem on dominated convergence and for $j \in \{4, 5\}$ follows that

$$I_j = \frac{1}{4\pi^2} \int_{r(x,y) \leq L} h_{k,L}^j(x, y)\, \mathrm{d}(x, y) = o(1) \quad \text{for} \quad L \longrightarrow \infty.$$

On the other hand, we have

$$I_3 \leq \Big(\frac{W^{(k)}(0)}{k!}\Big)^2 \|f\|_\alpha^2$$

leading to the asymptotic estimate

$$I_1 \leq \phi_{\alpha-\sigma,L,k}^* \Big(\frac{W^{(k)}(0)}{k!}\Big)^2 \|f\|_\alpha^2 + \phi_{\alpha-\sigma,L,k}^*\, o(1).$$

In Lemma 4.3.3 we have shown that, for $\gamma \geq 0$, the maximum $\phi_{\gamma,L,k}^*$ is bounded above by

$$\phi_{\gamma,L,k}^* \leq \begin{cases} L^{-2\gamma} & \text{for } \gamma \leq k \vee (\gamma > k \wedge L < L^*) \\ c_{\gamma,k}^2\, L^{-2k} & \text{for } \gamma > k \wedge L \geq L^* \end{cases}$$

with the critical bandwidth $L^* = \frac{\sqrt{k}}{\sqrt{\gamma - k}}$ and the strictly monotonically decreasing constant

$$c_{\gamma,k} = \Big(\frac{k}{\gamma - k}\Big)^{k/2} \Big(\frac{\gamma - k}{\gamma}\Big)^{\gamma/2} \quad \text{for } \gamma > k.$$

Thus, for $\alpha - \sigma \leq k$ follows that

$$I_1 \leq \frac{1}{(k!)^2} \, |W^{(k)}(0)|^2 \, L^{2(\sigma-\alpha)} \, \|f\|_\alpha^2 + o\Big(L^{2(\sigma-\alpha)}\Big)$$

and, for $\alpha - \sigma > k$,

$$I_1 \leq \frac{c_{\alpha-\sigma,k}^2}{(k!)^2} \, |W^{(k)}(0)|^2 \, L^{-2k} \, \|f\|_\alpha^2 + o\Big(L^{-2k}\Big).$$

By combining our derived bounds for I_1 and I_2, we finally get the asymptotic H^σ-error estimate

$$\|e_L\|_\sigma \leq \Big(C_{\alpha-\sigma,k} \, |W^{(k)}(0)| \, L^{-\min\{k,\alpha-\sigma\}} + L^{\sigma-\alpha}\Big) \, \|f\|_\alpha + o\Big(L^{-\min\{k,\alpha-\sigma\}}\Big).$$

In summary, we have proven the following error theorem for the approximate FBP method.

Theorem 4.4.1 (Asymptotic H^σ-error estimate)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha > 0$ and let the window $W \in \mathrm{L}^\infty(\mathbb{R})$ be even and compactly supported with $\operatorname{supp}(W) \subseteq [-1,1]$. Moreover, let W be k-times differentiable at the origin for $k \in \mathbb{N}$ with*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Then, for $0 \leq \sigma \leq \alpha$, the H^σ-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by

$$\|e_L\|_\sigma \leq \left(\frac{1}{k!} \, |W^{(k)}(0)| + 1\right) L^{\sigma-\alpha} \, \|f\|_\alpha + o(L^{\sigma-\alpha}) \tag{4.16}$$

if $\alpha - \sigma \leq k$ and, if $\alpha - \sigma > k$, by

$$\|e_L\|_\sigma \leq \frac{c_{\alpha-\sigma,k}}{k!} \, |W^{(k)}(0)| \, L^{-k} \, \|f\|_\alpha + o(L^{-k}) \tag{4.17}$$

with the strictly monotonically decreasing constant

$$c_{\alpha-\sigma,k} = \Big(\frac{k}{\alpha-\sigma-k}\Big)^{k/2} \Big(\frac{\alpha-\sigma-k}{\alpha-\sigma}\Big)^{(\alpha-\sigma)/2} \qquad \text{for } \alpha - \sigma > k.$$

In particular,

$$\|e_L\|_\sigma = \mathcal{O}\Big(L^{-\min\{k,\alpha-\sigma\}}\Big) \quad \text{for} \quad L \longrightarrow \infty. \qquad \square$$

We wish to draw the following conclusions from Theorem 4.4.1.

Firstly, the *flatness* of the filter's window function W determines the convergence rate of the error bounds (4.16) and (4.17) for the inherent FBP reconstruction error. Indeed, if W is k-times differentiable at the origin such that the first $k-1$ derivatives of W vanish at zero, then the convergence rate in (4.16) is given by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error e_L is measured, as long as $\alpha - \sigma \leq k$. For $\alpha - \sigma > k$, however, the order of convergence in (4.17) saturates at rate $\mathcal{O}(L^{-k})$.

Note that this observation is consistent with our results in Corollary 4.3.5, where we more restrictively assumed that the window W is k-times continuously differentiable on the whole interval $[-1,1]$ where it is supported. Consequently, the weaker assumptions in Theorem 4.4.1 lead to the same convergence rates for the H^σ-norm of the inherent FBP reconstruction error as our error theory for $\mathscr{C}^k$-windows. However, the estimates (4.16) and (4.17) are only asymptotic in contrast to the estimates (4.14) and (4.15) in Corollary 4.3.5.

Secondly, the quantity $|W^{(k)}(0)|$, i.e., the k-th derivative of W at the origin, dominates the error bound in both (4.16) and (4.17). Therefore, the value $|W^{(k)}(0)|$ can be used as an indicator to predict the approximation quality of the proposed FBP method in the H^σ-norm.

As a corollary, choosing $\sigma = 0$ in Theorem 4.4.1 yields the following asymptotic L^2-error estimate for the approximate FBP reconstruction.

Corollary 4.4.2 (Asymptotic L^2-error estimate)**.** *For $\alpha > 0$ let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$. Moreover, let $W \in \mathrm{L}^\infty(\mathbb{R})$ be even with $\operatorname{supp}(W) \subseteq [-1,1]$ and k-times differentiable at the origin with*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1.$$

Then, for $\alpha \le k$, the L^2-norm of the inherent FBP reconstruction error $e_L = f - f_L$ is bounded above by

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \left(\frac{1}{k!}\,|W^{(k)}(0)| + 1\right) L^{-\alpha}\,\|f\|_\alpha + o(L^{-\alpha}).$$

On the other hand, for $\alpha > k$ and sufficiently large $L > 0$, the L^2-norm of e_L is bounded above by

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \frac{c_{\alpha,k}}{k!}\,|W^{(k)}(0)|\,L^{-k}\,\|f\|_\alpha + o(L^{-k})$$

with the strictly monotonically decreasing constant

$$c_{\alpha,k} = \Big(\frac{k}{\alpha-k}\Big)^{k/2}\Big(\frac{\alpha-k}{\alpha}\Big)^{\alpha/2} \quad \text{for } \alpha > k.$$

In particular,

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(L^{-\min\{k,\alpha\}}\Big) \quad \text{for} \quad L \longrightarrow \infty. \qquad \square$$

To conclude our discussion, we finally consider the following special case.

Remark 4.4.3 (Special case)**.** *Let the filter's window function W fulfil the assumptions of Corollary 4.4.2 with $k \in \mathbb{N}$ and let the smoothness α of the target function $f \in \mathrm{H}^\alpha(\mathbb{R}^2)$ satisfy*

$$\alpha > k.$$

Then, the asymptotic L^2-error estimate of the FBP method reduces to

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le C_{\alpha,k}\,|W^{(k)}(0)|\,L^{-k}\,\|f\|_\alpha + o(L^{-k}).$$

Consequently, the L^2-norm of the intrinsic FBP reconstruction error is proportional to $|W^{(k)}(0)|$, if we neglect the higher order terms.

We remark that the commonly used low-pass filters $A_L(S) = |S|\,W(S/L)$ we considered after Example 4.3.2 satisfy the assumptions of our theory with $k = 2$ so that we obtain

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le C\,|W''(0)|\,L^{-2}\,\|f\|_\alpha + o(L^{-2}).$$

Therefore, in this case the intrinsic FBP reconstruction error is proportional to $|W''(0)|$ and its decay rate is of order $\mathcal{O}(L^{-2})$. This observation complies with the results of Munshi [76] and Munshi et al. [78], [80], see also the asymptotic pointwise error formula (4.6). Finally, Table 4.1 lists the quality indicator $|W''(0)|$ for the commonly used low-pass filters mentioned above.

Name	$W(S)$ for $\lvert S\rvert \le 1$	Parameter	$\lvert W''(0)\rvert$
Shepp-Logan	$\operatorname{sinc}(\pi S/2)$	-	$\pi^2/12$
Cosine	$\cos(\pi S/2)$	-	$\pi^2/4$
Hamming	$\beta + (1-\beta)\cos(\pi S)$	$\beta \in [1/2, 1]$	$(1-\beta)\,\pi^2$
Gaussian	$\exp(-(\pi S/\beta)^2)$	$\beta > 1$	$2\pi^2/\beta^2$

Table 4.1: Quality indicator $|W''(0)|$ for commonly used low-pass filters.

4.5 Error estimates for noisy data

To close this chapter on error estimates for the method of filtered back projection, we reconsider the important case of noisy data. Recall that, in many relevant applications, the Radon data $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ is usually not known exactly, but only up to a noise level $\delta > 0$ so that we have to reconstruct the target function $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$, for $\alpha > 0$, from given noisy measurements $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ satisfying

$$\|\mathcal{R}f - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \leq \delta.$$

By applying the FBP method (3.4) to the noisy data g^δ, we obtain the reconstruction

$$f_L^\delta = \frac{1}{2}\,\mathcal{B}(q_L * g^\delta) \tag{4.18}$$

and, as mentioned at the beginning of Section 4.2, the overall FBP reconstruction error

$$e_L^\delta = f - f_L^\delta \tag{4.19}$$

can be split into an approximation error term and a data error term,

$$e_L^\delta = \underbrace{f - f_L}_{\substack{\text{approximation}\\ \text{error}}} + \underbrace{f_L - f_L^\delta}_{\substack{\text{data}\\ \text{error}}}.$$

In the following, we analyse the L^2-norm of the overall FBP reconstruction error e_L^δ in (4.19) with respect to the noise level δ as well as the filter's window function W and bandwidth L. To this end, we first show that for even window $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\mathrm{supp}(W) \subseteq [-1,1]$ the noisy FBP reconstruction f_L^δ in (4.18) also satisfies $f_L^\delta \in \mathrm{L}^2(\mathbb{R}^2)$. By the triangle inequality, we then have

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} + \|f_L - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)}$$

and, consequently, we can treat the approximation error and the data error separately. Note that the approximation error is just the inherent FBP reconstruction error e_L in (4.7) we considered earlier in this chapter. Hence, for the analysis of the approximation error we can in particular rely on our results from Section 4.3.

4.5.1 Analysis of the data error

We start with analysing the data error $f_L - f_L^\delta$ in the L^2-norm. To this end, recall that the Radon transform $\mathcal{R}$ is continuous as a mapping

$$\mathcal{R} : \mathcal{S}(\mathbb{R}^2) \longrightarrow \mathcal{S}(\mathbb{R} \times [0,\pi)).$$

In particular, the Radon transform of a Schwartz function $f \in \mathcal{S}(\mathbb{R}^2)$ belongs to the Schwartz space $\mathcal{S}(\mathbb{R} \times [0,\pi)) \subset \mathrm{L}^2(\mathbb{R} \times [0,\pi))$, see [43, Theorem 4.1]. Since, up to the constant $\frac{1}{\pi}$, the back projection operator $\mathcal{B}$ is the formal dual operator of the Radon transform $\mathcal{R}$ satisfying

$$(\mathcal{R}f, g)_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} = \pi\,(f, \mathcal{B}g)_{\mathrm{L}^2(\mathbb{R}^2)} \quad \forall\, f \in \mathcal{S}(\mathbb{R}^2),\, g \in \mathcal{S}(\mathbb{R} \times [0,\pi)),$$

we can conclude that $\mathcal{B}g$ is a tempered distribution on $\mathbb{R}^2$, $\mathcal{B}g \in \mathcal{S}'(\mathbb{R}^2)$, for all $g \in \mathcal{S}'(\mathbb{R}\times[0,\pi))$. Because further $\mathrm{L}^2(\mathbb{R} \times [0,\pi)) \subset \mathcal{S}'(\mathbb{R} \times [0,\pi))$, we especially obtain

$$\mathcal{B}g \in \mathcal{S}'(\mathbb{R}^2) \quad \forall\, g \in \mathrm{L}^2(\mathbb{R} \times [0,\pi)). \tag{4.20}$$

Based on this, we now show that

$$R_L g = \frac{1}{2}\,\mathcal{B}(q_L * g) \tag{4.21}$$

defines a continuous linear operator

$$R_L : \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi)) \longrightarrow \mathrm{L}^2(\mathbb{R}^2),$$

which we call *FBP regularization operator*.

Theorem 4.5.1 (FBP regularization operator)**.** *Let* $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ *and let* $W \in \mathrm{L}^\infty(\mathbb{R})$ *be even with* $\operatorname{supp}(W) \subseteq [-1,1]$*. Then, we have* $R_L g \in \mathrm{L}^2(\mathbb{R}^2)$ *with*

$$\|R_L g\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \frac{1}{\sqrt{2\pi}} \left(\operatorname*{ess\,sup}_{S \in [-1,1]} |S|\,|W(S)|^2 \right)^{1/2} L^{1/2}\, \|g\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))}.$$

Proof. Recall that the band-limited function q_L defined in (3.5) is well-defined on $\mathbb{R} \times [0,\pi)$ and we have $q_L \in \mathrm{L}^2(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^\infty(\mathbb{R} \times [0,\pi))$. In addition, for all $\theta \in [0,\pi)$ the Fourier inversion formula

$$A_L(S) = \mathcal{F}(\mathcal{F}^{-1} A_L)(S) = \mathcal{F} q_L(S,\theta)$$

holds in L^2-sense and, in particular, for almost all $S \in \mathbb{R}$. Because further $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$, we obtain the representation

$$A_L(S)\, \mathcal{F} g(S,\theta) = \mathcal{F}(q_L * g)(S,\theta) \quad \text{for almost all } S \in \mathbb{R}$$

by the Fourier convolution Theorem 3.1.1. Moreover, Young's inequality, Theorem A.1.16, yields $(q_L * g)(\cdot,\theta) \in \mathrm{L}^2(\mathbb{R})$ for any $\theta \in [0,\pi)$ and by the Fourier inversion formula (in L^2-sense) follows that

$$(q_L * g)(S,\theta) = \mathcal{F}^{-1}[A_L(S)\, \mathcal{F} g(S,\theta)] \quad \text{for almost all } S \in \mathbb{R}.$$

In particular, for $g \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ we have $(q_L * g) \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ and, therefore,

$$R_L g = \frac{1}{2}\, \mathcal{B}(q_L * g)$$

is defined almost everywhere on $\mathbb{R}^2$ and satisfies $R_L g \in \mathrm{L}^2_{\mathrm{loc}}(\mathbb{R}^2)$ due to Proposition 3.1.4.

On the other hand, we also have $R_L g \in \mathcal{S}'(\mathbb{R}^2)$ according to (4.20). This allows us to determine the (distributional) Fourier transform of $R_L g$, as being defined via the duality relation

$$\langle \mathcal{F}(R_L g), w \rangle = \langle R_L g, \mathcal{F} w \rangle = \frac{1}{2}\, (\mathcal{B}(q_L * g), \mathcal{F} w)_{\mathrm{L}^2(\mathbb{R}^2)} \quad \forall\, w \in \mathcal{S}(\mathbb{R}^2).$$

Now, for an arbitrary Schwartz function $w \in \mathcal{S}(\mathbb{R}^2)$ the definition of the back projection $\mathcal{B}$ gives

$$\langle R_L g, \mathcal{F} w \rangle = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^\pi (q_L * g)(x\cos(\theta) + y\sin(\theta), \theta)\, \mathrm{d}\theta\, \mathcal{F} w(x,y)\, \mathrm{d}x\, \mathrm{d}y$$

and by applying the parameter transformation

$$x = t\cos(\theta) - s\sin(\theta) \quad \text{and} \quad y = t\sin(\theta) + s\cos(\theta),$$

we obtain

$$\begin{aligned} \langle R_L g, \mathcal{F} w \rangle &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^\pi (q_L * g)(t,\theta)\, \mathcal{F} w(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\, \mathrm{d}\theta\, \mathrm{d}t\, \mathrm{d}s \\ &= \frac{1}{2\pi} \int_0^\pi \int_{\mathbb{R}} (q_L * g)(t,\theta)\, \mathcal{R}(\mathcal{F} w)(t,\theta)\, \mathrm{d}t\, \mathrm{d}\theta \end{aligned}$$

by Fubini's theorem and the definition of the Radon transform $\mathcal{R}$. Recall that the operators

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^2) \longrightarrow \mathcal{S}(\mathbb{R}^2) \quad \text{and} \quad \mathcal{R} : \mathrm{L}^1(\mathbb{R}^2) \longrightarrow \mathrm{L}^1(\mathbb{R} \times [0,\pi))$$

are continuous and $\mathcal{S}(\mathbb{R}^2) \subset \mathrm{L}^1(\mathbb{R}^2)$. Consequently, for any $\theta \in [0,\pi)$, we obtain

$$\mathcal{R}(\mathcal{F} w)(\cdot,\theta) \in \mathrm{L}^1(\mathbb{R}) \quad \forall\, w \in \mathcal{S}(\mathbb{R}^2)$$

and the application of Parseval's identity

$$\int_{\mathbb{R}} \mathcal{F}^{-1} f(x)\, h(x)\, \mathrm{d}x = \int_{\mathbb{R}} f(x)\, \mathcal{F}^{-1} h(x)\, \mathrm{d}x \quad \forall\, f, h \in \mathrm{L}^1(\mathbb{R})$$

in combination with the representation

$$(q_L * g)(t,\theta) = \mathcal{F}^{-1}[A_L(t)\,\mathcal{F}g(t,\theta)] \quad \text{for almost all } t \in \mathbb{R}$$

yields

$$\langle R_L g, \mathcal{F}w\rangle = \frac{1}{2\pi}\int_0^\pi \int_{\mathbb{R}} A_L(t)\,\mathcal{F}g(t,\theta)\,\mathcal{F}^{-1}(\mathcal{R}(\mathcal{F}w))(t,\theta)\,\mathrm{d}t\,\mathrm{d}\theta.$$

To continue our analysis, we note that the Fourier transform $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$ are related via

$$\mathcal{F}^{-1}f = (2\pi)^{-n}\,(\mathcal{F}f)^* \quad \forall\, f \in \mathrm{L}^1(\mathbb{R}^n),$$

where $^* : \mathrm{L}^p(\mathbb{R}^n) \longrightarrow \mathrm{L}^p(\mathbb{R}^n)$ denotes the parity operator, defined as

$$f^*(x) = f(-x) \quad \text{for } x \in \mathbb{R}^n.$$

Since $\mathcal{F}w \in \mathrm{L}^1(\mathbb{R}^2)$, the Fourier slice Theorem 2.2.11 now gives, for all $(t,\theta) \in \mathbb{R}\times[0,\pi)$,

$$\begin{aligned}\mathcal{F}^{-1}\big(\mathcal{R}(\mathcal{F}w)\big)(t,\theta) &= (2\pi)^{-1}\,\mathcal{F}(\mathcal{R}(\mathcal{F}w))(-t,\theta) = (2\pi)^{-1}\,\mathcal{F}(\mathcal{F}w)(-t\cos(\theta),-t\sin(\theta))\\ &= 2\pi\,\mathcal{F}^{-1}(\mathcal{F}w)(t\cos(\theta),t\sin(\theta)) = 2\pi\,w(t\cos(\theta),t\sin(\theta))\end{aligned}$$

by applying the Fourier inversion formula on $\mathcal{S}(\mathbb{R}^2)$. Using this, we subsequently obtain

$$\begin{aligned}\langle R_L g, \mathcal{F}w\rangle &= \frac{1}{2\pi}\int_0^\pi\int_{\mathbb{R}} A_L(t)\,\mathcal{F}g(t,\theta)\,2\pi\,w(t\cos(\theta),t\sin(\theta))\,\mathrm{d}t\,\mathrm{d}\theta\\ &= \int_0^\pi\int_{\mathbb{R}} W_L(t)\,\mathcal{F}g(t,\theta)\,w(t\cos(\theta),t\sin(\theta))\,|t|\,\mathrm{d}t\,\mathrm{d}\theta.\end{aligned}$$

Transforming back to Cartesian coordinates, i.e., $(x,y) = (t\cos(\theta),t\sin(\theta))$, finally shows that the (distributional) Fourier transform of $R_L g$ is given by

$$\mathcal{F}(R_L g)(S\cos(\theta),S\sin(\theta)) = W_L(S)\,\mathcal{F}g(S,\theta) \quad \text{for almost all } (S,\theta) \in \mathbb{R}\times[0,\pi).$$

Since $W \in \mathrm{L}^\infty(\mathbb{R})$ is compactly supported with $\mathrm{supp}(W) \subseteq [-1,1]$ and $g \in \mathrm{L}^2(\mathbb{R}\times[0,\pi))$, we can conclude that $\mathcal{F}(R_L g) \in \mathrm{L}^2(\mathbb{R}^2)$. Indeed, by transforming to polar coordinates we obtain

$$\begin{aligned}\|\mathcal{F}(R_L g)\|^2_{\mathrm{L}^2(\mathbb{R}^2)} &= \int_0^\pi\int_{\mathbb{R}} |\mathcal{F}(R_L g)(S\cos(\theta),S\sin(\theta))|^2\,|S|\,\mathrm{d}S\,\mathrm{d}\theta\\ &= \int_0^\pi\int_{\mathbb{R}} |W_L(S)|^2\,|S|\,|\mathcal{F}g(S,\theta)|^2\,\mathrm{d}S\,\mathrm{d}\theta.\end{aligned}$$

Because the scaled window function W_L has compact support in $[-L,L]$, we finally obtain

$$\begin{aligned}\|\mathcal{F}(R_L g)\|^2_{\mathrm{L}^2(\mathbb{R}^2)} &\le \left(\operatorname*{ess\,sup}_{S\in[-L,L]} |S|\,|W_L(S)|^2\right)\int_0^\pi\int_{\mathbb{R}} |\mathcal{F}g(S,\theta)|^2\,\mathrm{d}S\,\mathrm{d}\theta\\ &= 2\pi\,L\left(\operatorname*{ess\,sup}_{S\in[-1,1]} |S|\,|W(S)|^2\right)\|g\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} < \infty.\end{aligned}$$

By the Rayleigh-Plancherel Theorem A.1.12, we then also have $R_L g \in \mathrm{L}^2(\mathbb{R}^2)$ with

$$\|R_L g\|^2_{\mathrm{L}^2(\mathbb{R}^2)} = \frac{1}{4\pi^2}\,\|\mathcal{F}(R_L g)\|^2_{\mathrm{L}^2(\mathbb{R}^2)} \le \frac{L}{2\pi}\left(\operatorname*{ess\,sup}_{S\in[-1,1]} |S|\,|W(S)|^2\right)\|g\|^2_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))},$$

i.e.,

$$\|R_L g\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \frac{1}{\sqrt{2\pi}}\left(\operatorname*{ess\,sup}_{S\in[-1,1]} |S|\,|W(S)|^2\right)^{1/2} L^{1/2}\,\|g\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))},$$

and the proof is complete. □

We are now prepared to analyse the data error $f_L - f_L^\delta$ in the L^2-norm for target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with some $\alpha > 0$ satisfying $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$, where

$$f_L = \frac{1}{2}\,\mathcal{B}(q_L * \mathcal{R}f) = R_L(\mathcal{R}f) \quad \text{and} \quad f_L^\delta = \frac{1}{2}\,\mathcal{B}(q_L * g^\delta) = R_L(g^\delta)$$

with noisy measurements $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$.

Theorem 4.5.2 (Data error)**.** *For $\alpha > 0$ let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ satisfy $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$. Further, let $W \in \mathrm{L}^\infty([-1,1])$ be even and, for $\delta > 0$, let $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ be given with*

$$\|\mathcal{R}f - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \le \delta.$$

Then, the L^2-norm of the data error $f_L - f_L^\delta$ is bounded above by

$$\|f_L - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le c_W\, L^{1/2}\, \delta,$$

where

$$c_W^2 = \frac{1}{2\pi} \operatorname*{ess\,sup}_{S\in[-1,1]} |S|\,|W(S)|^2.$$

Proof. Since $f \in \mathrm{L}^1(\mathbb{R}^2)$ implies $\mathcal{R}f \in \mathrm{L}^1(\mathbb{R} \times [0,\pi))$ due to Proposition 2.2.7 and, additionally, $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ as well as $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ by assumption, we can use the linear FBP regularization operator $R_L : \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi)) \longrightarrow \mathrm{L}^2(\mathbb{R}^2)$ defined in (4.21) to obtain the representation

$$f_L - f_L^\delta = R_L(\mathcal{R}f) - R_L(g^\delta) = R_L(\mathcal{R}f - g^\delta) \in \mathrm{L}^2(\mathbb{R}^2).$$

Consequently, applying Theorem 4.5.1 gives the estimate

$$\|f_L - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \frac{1}{\sqrt{2\pi}} \left(\operatorname*{ess\,sup}_{S\in[-1,1]} |S|\,|W(S)|^2 \right)^{1/2} L^{1/2}\, \|\mathcal{R}f - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \le c_W\, L^{1/2}\, \delta,$$

as stated. □

4.5.2 Analysis of the overall FBP reconstruction error

We now combine the above results for the data error with previous of our findings for the approximation error to estimate the L^2-norm of the overall FBP reconstruction error e_L^δ in (4.19) via

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} + \|f_L - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)}.$$

For that purpose, we first recall two representative estimates on the approximation error $f - f_L$ from Section 4.3 in the L^2-norm depending on properties of the filter's window function W.

The first estimate on the approximation error $f - f_L$ relies on the basic assumption that the smallest maximizer $S^*_{\alpha,W,L} \in [0,1]$ of the even function

$$\Phi_{\alpha,W,L}(S) = \frac{(1 - W(S))^2}{(1 + L^2S^2)^\alpha} \quad \text{for } S \in [-1,1]$$

is uniformly bounded away from 0, i.e., there exists a constant $c_{\alpha,W} > 0$ such that

$$S^*_{\alpha,W,L} \ge c_{\alpha,W} \quad \forall\, L > 0. \tag{A}$$

Theorem 4.5.3 (see Theorem 4.3.1)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$, for $\alpha > 0$, and $W \in \mathrm{L}^\infty([-1,1])$. Then, the L^2-norm of the approximation error $f - f_L$ is under Assumption* (A) *bounded above by*

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \left(c_{\alpha,W}^{-\alpha}\, \|1 - W\|_{\mathrm{L}^\infty([0,1])} + 1 \right) L^{-\alpha}\, \|f\|_\alpha.$$ □

The second L^2-error estimate on the approximation error $f - f_L$ from Section 4.3 works with conditions only on the window function W and predicts that the decay rate of the error saturates depending on the flatness of W at the origin.

Theorem 4.5.4 (see Corollaries 4.3.9 and 4.3.11)**.** *For $\alpha > 0$ let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$. Further, for $k \in \mathbb{N}$ let $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \le j \le k-1$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ with $1 < p \le \infty$ such that*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1.$$

Then, the L^2-norm of the approximation error $f - f_L$ is bounded above by

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,k,p}\, \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1\Big)\, L^{-\min\{k-1/p,\alpha\}}\, \|f\|_\alpha$$

with some constant $c_{\alpha,k,p} > 0$ independent of W and f. □

We are now in the position to bound the overall FBP reconstruction error e_L^δ in the L^2-norm. On the one hand, combining Theorem 4.5.2 with Theorem 4.5.3 gives the estimate

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,W}^{-\alpha}\, \|1 - W\|_{\mathrm{L}^\infty([0,1])} + 1\Big)\, L^{-\alpha}\, \|f\|_\alpha + c_W\, L^{1/2}\, \delta.$$

By coupling the bandwidth L with the noise level δ via

$$L = \delta^{-\frac{2}{2\alpha+1}}\, \|f\|_\alpha^{\frac{2}{2\alpha+1}}$$

we obtain

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,W}^{-\alpha}\, \|1 - W\|_{\mathrm{L}^\infty([0,1])} + c_W + 1\Big)\, \|f\|_\alpha^{\frac{1}{2\alpha+1}}\, \delta^{\frac{2\alpha}{2\alpha+1}}.$$

Thus, we have just established the following result.

Corollary 4.5.5 (Convergence rates for noisy data I)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$ satisfy $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ and let $W \in \mathrm{L}^\infty([-1,1])$ such that Assumption* (A) *is fulfilled. Further, let $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ be given with*

$$\|\mathcal{R}f - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \le \delta.$$

Then, the L^2-norm of the overall FBP reconstruction error $e_L^\delta = f - f_L^\delta$ is bounded above by

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,W}^{-\alpha}\, \|1 - W\|_{\mathrm{L}^\infty([0,1])} + c_W + 1\Big)\, \|f\|_\alpha^{\frac{1}{2\alpha+1}}\, \delta^{\frac{2\alpha}{2\alpha+1}},$$

where the bandwidth L is chosen as

$$L = \delta^{-\frac{2}{2\alpha+1}}\, \|f\|_\alpha^{\frac{2}{2\alpha+1}}.$$

In particular,

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(\delta^{\frac{2\alpha}{2\alpha+1}}\Big) \quad \text{for} \quad \delta \searrow 0.$$ □

On the other hand, combining Theorem 4.5.2 with Theorem 4.5.4 yields the estimate

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,k,p}\, \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1\Big)\, L^{-\min\{k-1/p,\alpha\}}\, \|f\|_\alpha + c_W\, L^{1/2}\, \delta.$$

By choosing

$$L = \delta^{-\frac{2}{2\min\{k-1/p,\alpha\}+1}}\, \|f\|_\alpha^{\frac{2}{2\min\{k-1/p,\alpha\}+1}}$$

we now obtain

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,k,p}\, \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + c_W + 1\Big)\, \|f\|_\alpha^{\frac{1}{2\min\{k-1/p,\alpha\}+1}}\, \delta^{\frac{2\min\{k-1/p,\alpha\}}{2\min\{k-1/p,\alpha\}+1}}.$$

We summarize our discussion as follows.

Corollary 4.5.6 (Convergence rates for noisy data II)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$ satisfy $\mathcal{R}f \in \mathrm{L}^2(\mathbb{R} \times [0,\pi))$. Further, for $k \in \mathbb{N}$ let $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \le j \le k-1$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ with $1 < p \le \infty$ such that*

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1.$$

Finally, let $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ be given with

$$\|\mathcal{R}f - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \le \delta.$$

Then, the L^2-norm of the overall FBP reconstruction error $e_L^\delta = f - f_L^\delta$ is bounded above by

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le \Big(c_{\alpha,k,p}\,\|W^{(k)}\|_{\mathrm{L}^p([0,1])} + c_W + 1\Big)\,\|f\|_\alpha^{\frac{1}{2\min\{k-1/p,\alpha\}+1}}\,\delta^{\frac{2\min\{k-1/p,\alpha\}}{2\min\{k-1/p,\alpha\}+1}},$$

where the bandwidth L is chosen as

$$L = \delta^{-\frac{2}{2\min\{k-1/p,\alpha\}+1}}\,\|f\|_\alpha^{\frac{2}{2\min\{k-1/p,\alpha\}+1}}.$$

In particular,

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(\delta^{\frac{2\min\{k-1/p,\alpha\}}{2\min\{k-1/p,\alpha\}+1}}\Big) \quad \text{for} \quad \delta \searrow 0. \qquad \square$$

Note that the decay rate of the L^2-error bound in Corollary 4.5.5 is given by

$$\|f - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(\delta^{\frac{2\alpha}{2\alpha+1}}\Big) \quad \text{for} \quad \delta \searrow 0,$$

where the filter's bandwidth $L > 0$ has to go to ∞ as the noise level $\delta > 0$ goes to 0 with rate

$$L = \mathcal{O}\Big(\delta^{-\frac{2}{2\alpha+1}}\Big) \quad \text{for} \quad \delta \searrow 0.$$

In particular, we observe that the L^2-error bound approaches the optimal convergence order

$$\|f - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}(\delta) \quad \text{for} \quad \delta \searrow 0$$

as the smoothness α of the target function f goes to ∞. In contrast to that, the order of convergence of the L^2-error bound in Corollary 4.5.6 saturates at rate

$$\|f - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(\delta^{\frac{2(k-1/p)}{2(k-1/p)+1}}\Big) \quad \text{for} \quad \delta \searrow 0.$$

To achieve convergence, the bandwidth L again has to go to ∞ as the noise level δ tends to 0, but now with rate

$$L = \mathcal{O}\Big(\delta^{-\frac{2}{2(k-1/p)+1}}\Big) \quad \text{for} \quad \delta \searrow 0.$$

Optimal choice of the bandwidth L

In this last paragraph we now determine the optimal bandwidth L minimizing the error bound in

$$\|f - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \le c_1\,L^{-\gamma}\,\|f\|_\alpha + c_2\,L^{1/2}\,\delta,$$

where

$$c_1 = c_{\alpha,W}^{-\alpha}\,\|1 - W\|_{\mathrm{L}^\infty([0,1])} + 1, \quad c_2 = c_W, \quad \gamma = \alpha$$

or

$$c_1 = c_{\alpha,k,p}\,\|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1, \quad c_2 = c_W, \quad \gamma = \min\{k - 1/p, \alpha\},$$

respectively, see Corollaries 4.5.5 and 4.5.6.

To this end, we define the auxiliary function

$$\varphi(L) = c_1\, L^{-\gamma}\, \|f\|_\alpha + c_2\, L^{1/2}\, \delta \quad \text{for } L > 0.$$

Differentiating yields

$$\varphi'(L) = -\gamma\, c_1\, L^{-\gamma-1}\, \|f\|_\alpha + \frac{1}{2}\, c_2\, L^{-1/2}\, \delta \quad \forall\, L > 0$$

such that

$$\varphi'(L) = 0 \quad \Longleftrightarrow \quad L = \Big(2\gamma\, \frac{c_1}{c_2}\Big)^{\frac{2}{2\gamma+1}}\, \delta^{-\frac{2}{2\gamma+1}}\, \|f\|_\alpha^{\frac{2}{2\gamma+1}}.$$

This choice of L yields the global minimum of the continuously differentiable function φ on $\mathbb{R}_{>0}$, since it is the only stationary point of $\varphi : \mathbb{R}_{>0} \longrightarrow \mathbb{R}_{>0}$ and

$$\varphi(L) \longrightarrow \infty \quad \text{for} \quad L \searrow 0$$

as well as

$$\varphi(L) \longrightarrow \infty \quad \text{for} \quad L \longrightarrow \infty.$$

Consequently, the minimal error bound on $\|f - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)}$ is given by

$$\begin{aligned}\|f - f_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} &\leq c_1 \Big(2\gamma\, \frac{c_1}{c_2}\Big)^{-\frac{2\gamma}{2\gamma+1}} \delta^{\frac{2\gamma}{2\gamma+1}}\, \|f\|_\alpha^{-\frac{2\gamma}{2\gamma+1}}\, \|f\|_\alpha + c_2 \Big(2\gamma\, \frac{c_1}{c_2}\Big)^{\frac{1}{2\gamma+1}} \delta^{-\frac{1}{2\gamma+1}}\, \|f\|_\alpha^{\frac{1}{2\gamma+1}}\, \delta \\ &= c_1^{\frac{1}{2\gamma+1}}\, c_2^{\frac{2\gamma}{2\gamma+1}}\, (2\gamma+1)\, (2\gamma)^{-\frac{2\gamma}{2\gamma+1}}\, \delta^{\frac{2\gamma}{2\gamma+1}}\, \|f\|_\alpha^{\frac{1}{2\gamma+1}}.\end{aligned}$$

In summary, we obtain the following result concerning the convergence rates for noisy data.

Corollary 4.5.7 (Convergence rates for noisy data III)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha > 0$ and let the window function W satisfy one of the following assumptions:*

(i) $W \in \mathrm{L}^\infty([-1,1])$ such that Assumption (A) *is fulfilled,*

(ii) $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \leq j \leq k-1$ with $k \in \mathbb{N}$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ with $1 < p \leq \infty$ such that

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

Further, let $g^\delta \in \mathrm{L}^1(\mathbb{R} \times [0,\pi)) \cap \mathrm{L}^2(\mathbb{R} \times [0,\pi))$ be given with

$$\|\mathcal{R}f - g^\delta\|_{\mathrm{L}^2(\mathbb{R}\times[0,\pi))} \leq \delta.$$

Then, the L^2-norm of the overall FBP reconstruction error $e_L^\delta = f - f_L^\delta$ is bounded above by

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq c_\gamma\, \delta^{\frac{2\gamma}{2\gamma+1}}\, \|f\|_\alpha^{\frac{1}{2\gamma+1}}$$

with

$$c_\gamma = c_1^{\frac{1}{2\gamma+1}}\, c_2^{\frac{2\gamma}{2\gamma+1}}\, (2\gamma+1)\, (2\gamma)^{-\frac{2\gamma}{2\gamma+1}}$$

and

$$L = \Big(2\gamma\, \frac{c_1}{c_2}\Big)^{\frac{2}{2\gamma+1}}\, \delta^{-\frac{2}{2\gamma+1}}\, \|f\|_\alpha^{\frac{2}{2\gamma+1}},$$

where

(i) $$c_1 = c_{\alpha,W}^{-\alpha}\, \|1 - W\|_{\mathrm{L}^\infty([0,1])} + 1, \quad c_2 = c_W, \quad \gamma = \alpha$$

or

(ii) $$c_1 = c_{\alpha,k,p}\, \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1, \quad c_2 = c_W, \quad \gamma = \min\{k - 1/p, \alpha\}.$$

In particular,

$$\|e_L^\delta\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(\delta^{\frac{2\gamma}{2\gamma+1}}\Big) \quad \text{for} \quad \delta \searrow 0. \qquad \square$$

In conclusion, we observe that the choice of the bandwidth L in Corollaries 4.5.5 and 4.5.6 leads to optimal convergence rates of the error bounds as the noise level δ goes to 0.

Chapter 5

Numerical experiments

In Chapter 2 we have seen that we can reconstruct a bivariate function $f \in \mathrm{L}^1(\mathbb{R}^2)$ from its given Radon data

$$\{\mathcal{R}f(t,\theta) \mid t \in \mathbb{R},\, \theta \in [0,\pi)\}$$

under suitable assumptions by applying the classical filtered back projection (FBP) formula (2.9), i.e.,

$$f(x,y) = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[|S|\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

However, the FBP formula is highly sensitive with respect to noise and, thus, cannot be used in practice. For that reason, in Chapter 3 we have explained how the FBP formula can be stabilized by incorporating a low-pass filter $A_L : \mathbb{R} \longrightarrow \mathbb{R}$ of the form

$$A_L(S) = |S|\, W_L(S) = |S|\, W(S/L) \qquad \text{for } S \in \mathbb{R}$$

with finite bandwidth $L > 0$ and an even window function $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\operatorname{supp}(W) \subseteq [-1,1]$. This reduces the noise sensitivity of the FBP formula, but only leads to an approximate FBP reconstruction f_L given by

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f).$$

In Chapter 4 we have then analysed the inherent FBP reconstruction error $e_L = f - f_L$ which is incurred by the application of the low-pass filter A_L. To this end, we have proven error estimates on e_L for target functions f from Sobolev spaces $\mathrm{H}^\alpha(\mathbb{R}^2)$ of fractional order $\alpha > 0$ depending on the bandwidth L and window W. To be more precise, we have estimated the H^σ-norm of e_L for all $0 \le \sigma \le \alpha$. In particular, we have proven convergence rates for the approximate FBP reconstruction f_L as the bandwidth L goes to infinity, where we have observed saturation of the order of convergence at fractional rates depending on smoothness properties of the window W.

In this chapter we finally provide numerical experiments to evaluate the inherent FBP reconstruction error e_L numerically and to validate our error theory from Chapter 4. In the following, however, we restrict ourselves to the classical L^2-case and let $\sigma = 0$. Further, we remark that the approximate FBP reconstruction formula assumes the Radon data $\mathcal{R}f(t,\theta)$ to be available for *all* parameters $(t,\theta) \in \mathbb{R} \times [0,\pi)$. In practice, however, only finitely many Radon samples are given and we have to recover the function f from a finite set of Radon data

$$\{\mathcal{R}f(t_j,\theta_j) \mid j = 1,\dots,J\} \qquad \text{for some } J \in \mathbb{N}.$$

This requires a suitable discretization of the approximate FBP method, leading to inevitable discretization errors that are not covered by our error theory in Chapter 4. In the first section of this chapter we describe a standard discretization of the FBP method. Following this, we explain the notion of *mathematical phantoms* and introduce analytical test cases for our numerical experiments. In the last part, we finally investigate the reconstruction error of the FBP method numerically, where our findings will be consistent with our theoretical results and, in particular, the predicted saturation of the order of convergence at fractional rates will be observed.

5.1 Discretization of the FBP method

In this section we describe how to implement the method of filtered back projection for the reconstruction of a bivariate function $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ from *finitely* many Radon samples

$$\{(\mathcal{R}f)_{j,k} \mid j = -M, \dots, M,\, k = 0, \dots, N-1\}\,.$$

The starting point for our elaborations is the representation (3.4) of the approximate FBP reconstruction f_L, i.e.,

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f), \tag{5.1}$$

where A_L denotes a fixed low-pass filter of the form

$$A_L(S) = |S|\,W_L(S) = |S|\,W(S/L) \quad \text{for } S \in \mathbb{R}$$

with finite bandwidth $L > 0$ and an even window function $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\operatorname{supp}(W) \subseteq [-1, 1]$.

The implementation of the FBP method requires a suitable discretization of the approximate reconstruction formula. To be more precise, we have to discretize the convolution product $*$ and the back projection operator $\mathcal{B}$. This also includes the specification of a sampling scheme for the Radon transform $\mathcal{R}f$ and the inverse Fourier transform $\mathcal{F}^{-1}A_L$ of the chosen low-pass filter. To this end, we follow a standard discretization approach and proceed as in [82, Section V.1].

5.1.1 Parallel beam geometry

A commonly used sampling scheme for the data acquisition is given by the so called *parallel beam geometry*, where the Radon lines $\ell_{t,\theta}$ are equally spaced in both the radial parameter $t \in \mathbb{R}$ and the angular parameter $\theta \in [0, \pi)$. More precisely, for N uniformly distributed angles we collect Radon samples along $2M + 1$ parallel lines per angle with a fixed spacing $d > 0$. Hence, the Radon data are of the form

$$(\mathcal{R}f)_{j,k} = \mathcal{R}f(t_j, \theta_k)$$

with

$$\theta_k = k \cdot \frac{\pi}{N} \quad \text{for } k = 0, \dots, N-1$$

and

$$t_j = j \cdot d \quad \text{for } j = -M, \dots, M$$

so that in total $N \cdot (2M + 1)$ Radon samples are taken. For illustration, Figure 5.1 shows the arrangement of 108 Radon lines in $[-1, 1]^2$ with $N = 12$, $M = 4$ and sampling spacing $d = 0.25$.

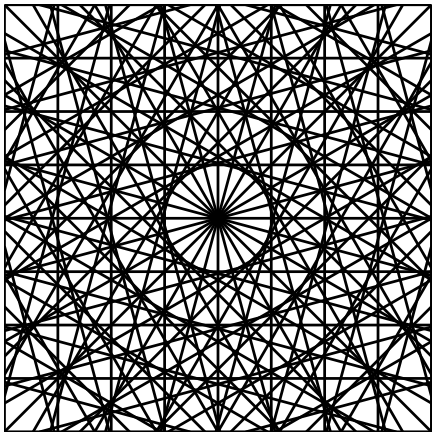


Figure 5.1: Parallel beam geometry with $N = 12$, $M = 4$ and $d = 0.25$.

Shannon sampling

Before we come to the discretization of the FBP method (5.1), we first discuss the sampling process in more detail, which will help to select a sample spacing d. In this paragraph we will see that we can uniquely recover a band-limited function $h \in \mathrm{L}^2(\mathbb{R})$ from uniformly spaced discrete function values $h(t_j)$, for $j \in \mathbb{Z}$, if the sampling distance $d = t_{j+1} - t_j$ is chosen reasonably. This is the statement of the classical *Shannon sampling theorem* and the sample spacing d corresponds to the smallest detail in h that is still recognizable after sampling the function.

Theorem 5.1.1 (Shannon sampling theorem, see [16, Theorem 4.35] or [19, Theorem 3.10.1])**.** *Let $h \in \mathrm{L}^2(\mathbb{R})$ be a band-limited function with $\operatorname{supp}(\mathcal{F}f) \subseteq [-L, L]$ for some bandwidth $L > 0$. Then, h is uniquely determined by the discrete values $h\big(\frac{\pi k}{L}\big)$, for $k \in \mathbb{Z}$, and we have*

$$h(t) = \sum_{k\in\mathbb{Z}} h\Big(\frac{\pi k}{L}\Big)\, \operatorname{sinc}(Lt - k\pi) \quad \forall\, t \in \mathbb{R}. \tag{5.2}$$

Proof. Let $h \in \mathrm{L}^2(\mathbb{R})$ be band-limited with $\operatorname{supp}(\mathcal{F}h) \subseteq [-L, L]$ for some bandwidth $L > 0$. By the Rayleigh-Plancherel Theorem A.1.12 we then have $\mathcal{F}h \in \mathrm{L}^2([-L, L]) \subset \mathrm{L}^1([-L, L])$, which implies that h has a continuous representative due to the Riemann-Lebesgue Lemma A.1.2. Thus, the Fourier inversion formula (A.3) holds pointwise and in L^2-sense yielding

$$h(t) = \mathcal{F}^{-1}(\mathcal{F}h)(t) = \frac{1}{2\pi}\int_{-L}^{L} \mathcal{F}h(\omega)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}\omega = \frac{1}{2L}\int_{-\pi}^{\pi} \mathcal{F}h\Big(\frac{L\omega}{\pi}\Big)\, \mathrm{e}^{\mathrm{i}\omega Lt/\pi}\, \mathrm{d}\omega \quad \forall\, t \in \mathbb{R}.$$

For fixed $t \in \mathbb{R}$ we now consider the Fourier expansion of $\mathrm{e}^{\mathrm{i}\omega Lt/\pi}$ as a function in $\mathrm{L}^2([-\pi, \pi])$, which is given by

$$\mathrm{e}^{\mathrm{i}\omega Lt/\pi} = \sum_{k\in\mathbb{Z}} c_k\, \mathrm{e}^{\mathrm{i}\omega k} \quad \forall\, \omega \in [-\pi, \pi]$$

with the Fourier coefficients

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\omega Lt/\pi}\, \mathrm{e}^{-\mathrm{i}\omega k}\, \mathrm{d}\omega = \frac{1}{2}\int_{-1}^{1} \mathrm{e}^{\mathrm{i}\,(Lt - k\pi)\,\omega}\, \mathrm{d}\omega = \operatorname{sinc}(Lt - k\pi) \quad \forall\, k \in \mathbb{Z}.$$

Recall that the partial sums of the above Fourier series converge in the L^2-norm, i.e.,

$$\int_{-\pi}^{\pi} \Big| \mathrm{e}^{\mathrm{i}\omega Lt/\pi} - \sum_{k=-n}^{n} c_k\, \mathrm{e}^{\mathrm{i}\omega k} \Big|^2\, \mathrm{d}\omega \longrightarrow 0 \quad \text{for} \quad n \longrightarrow \infty.$$

Since $\mathcal{F}h \in \mathrm{L}^2([-L, L])$, this in combination with the Cauchy-Schwarz inequality implies that

$$\left| \int_{-\pi}^{\pi} \mathcal{F}h\Big(\frac{L\omega}{\pi}\Big)\, \mathrm{e}^{\mathrm{i}\omega Lt/\pi}\, \mathrm{d}\omega - \sum_{k=-n}^{n} c_k \int_{-\pi}^{\pi} \mathcal{F}h\Big(\frac{L\omega}{\pi}\Big)\, \mathrm{e}^{\mathrm{i}\omega k}\, \mathrm{d}\omega \right| \longrightarrow 0 \quad \text{for} \quad n \longrightarrow \infty.$$

Consequently, we can interchange the order of summation and integration so that, for all $t \in \mathbb{R}$,

$$\begin{aligned} h(t) &= \frac{1}{2L}\int_{-\pi}^{\pi} \mathcal{F}h\Big(\frac{L\omega}{\pi}\Big) \sum_{k\in\mathbb{Z}} \operatorname{sinc}(Lt - k\pi)\, \mathrm{e}^{\mathrm{i}\omega k}\, \mathrm{d}\omega = \sum_{k\in\mathbb{Z}} \operatorname{sinc}(Lt - k\pi)\, \frac{1}{2L}\int_{-\pi}^{\pi} \mathcal{F}h\Big(\frac{L\omega}{\pi}\Big)\, \mathrm{e}^{\mathrm{i}\omega k}\, \mathrm{d}\omega \\ &= \sum_{k\in\mathbb{Z}} \operatorname{sinc}(Lt - k\pi)\, \frac{1}{2\pi}\int_{-L}^{L} \mathcal{F}h(\omega)\, \mathrm{e}^{\mathrm{i}\omega\pi k/L}\, \mathrm{d}\omega = \sum_{k\in\mathbb{Z}} h\Big(\frac{\pi k}{L}\Big)\, \operatorname{sinc}(Lt - k\pi), \end{aligned}$$

where we again use the Fourier inversion formula (A.3). □

We remark that the formula (5.2) is also called *Shannon-Whittaker interpolation formula* and the sample spacing $d = \frac{\pi}{L}$ is known as the *Nyquist rate*. Since a band-limited function h of bandwidth $L > 0$ contains no details smaller than $\frac{2\pi}{L}$, the Nyquist rate requires that h is sampled with a sampling distance d at most half of the smallest detail contained in h.

Remark 5.1.2. *If $h \in \mathcal{S}(\mathbb{R})$ is only* essentially L-band-limited *in the sense that, for $0 < \varepsilon \ll 1$,*

$$\int_{\mathbb{R}\setminus[-L,L]} |\mathcal{F}h(\omega)| \, \mathrm{d}\omega \leq \varepsilon,$$

the reconstruction of h from discrete samples $h(k \cdot d)$, for $k \in \mathbb{Z}$ and sampling distance $d > 0$, by using

$$S_d h(t) = \sum_{k\in\mathbb{Z}} h(k \cdot d) \, \operatorname{sinc}\left(\frac{\pi}{d}\,(t - k \cdot d)\right) \quad \text{for } t \in \mathbb{R}$$

is no longer exact. If $d \leq \frac{\pi}{L}$, however, [82, Theorem III.1.3] shows that the reconstruction error can be bounded by

$$\|S_d h - h\|_{\mathrm{L}^\infty(\mathbb{R})} \leq \frac{\varepsilon}{\pi}.$$

5.1.2 Discrete FBP reconstruction formula

We now address the discretization of the FBP method (5.1) for the approximate reconstruction of a target function f from discrete Radon data $\{(\mathcal{R}f)_{j,k}\}$ given in parallel beam geometry, i.e.,

$$(\mathcal{R}f)_{j,k} = \mathcal{R}f(t_j, \theta_k)$$

with

$$t_j = j \cdot d \quad \text{for } -M \leq j \leq M \qquad \text{and} \qquad \theta_k = k \cdot \frac{\pi}{N} \quad \text{for } 0 \leq k \leq N-1.$$

To this end, from now on we assume that f is compactly supported with

$$\operatorname{supp}(f) \subseteq B_r(0) \quad \text{for some } r \in \mathbb{N}.$$

Note that this is typically satisfied in relevant applications, cf. Remark 4.2.9, and implies that the Radon transform $\mathcal{R}f$ of f has compact support as well due to Proposition 2.2.9, since for all angles $\theta \in [0, \pi)$ follows that

$$\mathcal{R}f(t, \theta) = 0 \quad \forall\, |t| > r.$$

We start with discretizing the convolution product $*$ in (5.1) between the Radon data $\mathcal{R}f$ and the inverse Fourier transform $\mathcal{F}^{-1}A_L$ of the low-pass filter A_L. Here, for fixed angle $\theta \in [0, \pi)$, we have to approximate the convolution integral

$$(\mathcal{F}^{-1}A_L * \mathcal{R}f)(S, \theta) = \int_{\mathbb{R}} \mathcal{F}^{-1}A_L(S - t)\, \mathcal{R}(t, \theta) \, \mathrm{d}t \quad \text{for } S \in \mathbb{R}$$

by only using the discrete data

$$\mathcal{R}f(t_j, \theta) = \mathcal{R}f(j \cdot d, \theta) \quad \text{for } j \in \mathbb{Z}$$

taken at equally spaced sampling points $t_j = j \cdot d$, for $j \in \mathbb{Z}$, with fixed sampling distance $d > 0$. To achieve this, we apply the composite trapezoidal rule and replace the above convolution integral by the (infinite) sum

$$(\mathcal{F}^{-1}A_L * \mathcal{R}f)(S, \theta) \approx d \sum_{j\in\mathbb{Z}} \mathcal{F}^{-1}A_L(S - t_j)\, \mathcal{R}f(t_j, \theta) \quad \text{for } (S, \theta) \in \mathbb{R} \times [0, \pi).$$

Since f is assumed to have compact support, the above sum is in fact finite and, consequently, we obtain

$$(\mathcal{F}^{-1}A_L * \mathcal{R}f)(S, \theta) \approx d \sum_{j=-M}^{M} \mathcal{F}^{-1}A_L(S - t_j)\, \mathcal{R}f(t_j, \theta) \quad \text{for } (S, \theta) \in \mathbb{R} \times [0, \pi),$$

where $M \in \mathbb{N}$ is chosen sufficiently large such that, for any angle $\theta \in [0, \pi)$,

$$\mathcal{R}f(t, \theta) = 0 \quad \forall\, |t| > M \cdot d.$$

Let us continue with the discretization of the back projection operator $\mathcal{B}$. By Definition 2.2.15 the back projection $\mathcal{B}h$ of a function $h \equiv h(S,\theta)$ in polar coordinates is given by

$$\mathcal{B}h(x,y) = \frac{1}{\pi}\int_0^\pi h(x\cos(\theta)+y\sin(\theta),\theta)\,\mathrm{d}\theta \quad \text{for } (x,y)\in\mathbb{R}^2.$$

In (5.1), this has to be applied to the function

$$h(S,\theta) = (\mathcal{F}^{-1}A_L * \mathcal{R}f)(S,\theta) \quad \text{for } (S,\theta)\in\mathbb{R}\times[0,\pi),$$

where the Radon data $\mathcal{R}f(t,\theta)$ is only known for a finite set of N angles

$$\theta_k = k\cdot\frac{\pi}{N} \quad \text{for } k = 0,\dots,N-1.$$

Thus, for the discretization of $\mathcal{B}$ we again use the composite trapezoidal rule and replace the above integral by the sum

$$\mathcal{B}h(x,y) \approx \frac{1}{N}\sum_{k=0}^{N-1} h(x\cos(\theta_k)+y\sin(\theta_k),\theta_k) \quad \text{for } (x,y)\in\mathbb{R}^2.$$

Combining the discretization steps leads us to a discrete version of the FBP method (5.1) given by

$$f_D(x,y) = \frac{d}{2N}\sum_{k=0}^{N-1}\sum_{j=-M}^{M} \mathcal{F}^{-1}A_L(x\cos(\theta_k)+y\sin(\theta_k)-t_j)\,\mathcal{R}f(t_j,\theta_k) \quad \text{for } (x,y)\in\mathbb{R}^2.$$

By defining the *discrete convolution* $*_D$ of $\mathcal{F}^{-1}A_L$ and $\mathcal{R}f$ as

$$(\mathcal{F}^{-1}A_L *_D \mathcal{R}f)(S,\theta) = d\sum_{j=-M}^{M}\mathcal{F}^{-1}A_L(S-t_j)\,\mathcal{R}f(t_j,\theta) \quad \text{for } (S,\theta)\in\mathbb{R}\times[0,\pi)$$

and the *discrete back projection* $\mathcal{B}_D$ of a function $h\equiv h(S,\theta)$ in polar coordinates via

$$\mathcal{B}_D h(x,y) = \frac{1}{N}\sum_{k=0}^{N-1} h(x\cos(\theta_k)+y\sin(\theta_k),\theta_k) \quad \text{for } (x,y)\in\mathbb{R}^2,$$

we can rewrite f_D in compact form as

$$f_D = \frac{1}{2}\,\mathcal{B}_D(\mathcal{F}^{-1}A_L *_D \mathcal{R}f).$$

The evaluation of the discrete reconstruction f_D requires the computation of the values

$$(\mathcal{F}^{-1}A_L *_D \mathcal{R}f)(x\cos(\theta_k)+y\sin(\theta_k),\theta_k) \quad \forall\, 0\le k\le N-1$$

for each reconstruction point $(x,y)\in\mathbb{R}^2$. To reduce the computational costs, we evaluate, for each $0\le k\le N-1$, the function

$$h(t,\theta_k) = (\mathcal{F}^{-1}A_L *_D \mathcal{R}f)(t,\theta_k) = d\sum_{j=-M}^{M}\mathcal{F}^{-1}A_L(t-t_j)\,\mathcal{R}f(t_j,\theta_k) \quad \text{for } t\in\mathbb{R}$$

only at the sampling points $t_i = i\cdot d$ for $i\in I$ with a sufficiently large index set $I\subset\mathbb{Z}$. For each reconstruction point $(x,y)\in\mathbb{R}^2$ we then interpolate the value $h(t,\theta_k)$ at $t = x\cos(\theta_k)+y\sin(\theta_k)$ by using a suitable interpolation method $\mathcal{I}$. This leads us to the *discrete FBP reconstruction formula*

$$f_{\mathrm{FBP}} = \frac{1}{2}\,\mathcal{B}_D\Big(\mathcal{I}[\mathcal{F}^{-1}A_L *_D \mathcal{R}f]\Big). \tag{5.3}$$

According to [82, Section III.3] the optimal sampling conditions for the reconstruction of an essentially L-band-limited target function f supported in the unit disc $B_1(0)$, i.e., $r = 1$, are given by

$$d \leq \frac{\pi}{L}, \quad M \geq \frac{1}{d}, \quad N \geq L$$

leading to the well-known optimal sampling relation

$$N = \pi \cdot M.$$

Here, the restriction $d \leq \frac{\pi}{L}$ ensures that the convolution $*$ in (5.1) is properly discretized, while $N \geq L$ guarantees a satisfactory discretization of the back projection $\mathcal{B}$ via the trapezoidal rule. Since for fixed angle $\theta \in [0, \pi)$ the function

$$h(S) = (\mathcal{F}^{-1} A_L * \mathcal{R}f)(S, \theta) \quad \text{for } S \in \mathbb{R}$$

is band-limited with bandwidth L, the condition on d corresponds to the Nyquist rate for h according to the Shannon sampling Theorem 5.1.1. Finally, the relation $M \geq \frac{1}{d}$ ensures that the whole support of the target function f is covered during the acquisition of the Radon data.

Since we assume that f is supported in $B_r(0)$ for some $r \in \mathbb{N}$ and N, M have to be integers, we couple the discretization parameters $d > 0$ and $M, N \in \mathbb{N}$ with the bandwidth L via

$$d = \frac{\pi}{L}, \quad M = r \cdot \frac{L}{\pi}, \quad N = 3 \cdot M$$

and choose L to be a multiple of π, i.e., $L = \pi \cdot K$ for some $K \in \mathbb{N}$.

We summarize the discrete FBP method in the following image reconstruction algorithm, where we assume that the reconstruction f_{FBP} in (5.3) is evaluated in Cartesian grid points

$$\Big\{(x_m, y_n) \in \mathbb{R}^2 \mid (m, n) \in I_x \times I_y\Big\}$$

with finite index sets $I_x \times I_y \subset \mathbb{N} \times \mathbb{N}$.

Algorithm 1 Discrete FBP method

Input: Radon data $(\mathcal{R}f)_{j,k} = \mathcal{R}f(t_j, \theta_k)$ for $j = -M, \ldots, M$, $k = 0, \ldots, N-1$

choose low-pass filter A_L with bandwidth $L > 0$

for $k = 0, \ldots, N-1$ **do** ▷ Computation of the discrete convolution
  for $i \in I$ **do**
    $$h(t_i, \theta_k) = \frac{\pi}{L} \sum_{j=-M}^{M} \mathcal{F}^{-1} A_L(t_i - t_j)\, \mathcal{R}f(t_j, \theta_k)$$
  end for
end for

choose interpolation method $\mathcal{I}$

for $m \in I_x$ **do** ▷ Computation of the discrete back projection
  for $n \in I_y$ **do**
    $$f_{\mathrm{FBP}}(x_m, y_n) = \frac{1}{2N} \sum_{k=0}^{N-1} \mathcal{I}h(x_m \cos(\theta_k) + y_n \sin(\theta_k), \theta_k)$$
  end for
end for

Output: Approximate reconstruction f_{FBP} on Cartesian grid $\{(x_m, y_n) \mid (m, n) \in I_x \times I_y\}$

There are many possible choices for the interpolation method $\mathcal{I}$. In the following, we give two examples that are commonly used.

Example 5.1.3 (Interpolation schemes, see [33, Section 8.8])**.** *Let the samples $g_j = g(t_j)$ of a univariate function $g \equiv g(t)$ be given at equidistant sampling points $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and a fixed sampling distance $d > 0$. To approximate the function value $g(t)$ at a given point $t \in \mathbb{R}$ we use one of the following interpolation schemes.*

- **Nearest neighbour interpolation:**

  Let $t \in [t_m, t_{m+1})$ for some $m \in \mathbb{Z}$. Then, the function value $g(t)$ is approximated by

  $$\mathcal{I}_0 g(t) = \begin{cases} g(t_m) & \text{for } t - t_m \leq t_{m+1} - t \\ g(t_{m+1}) & \text{for } t - t_m > t_{m+1} - t. \end{cases}$$

  This defines a piecewise constant interpolant $\mathcal{I}_0 g$ of g, which is discontinuous in general. Note that by considering the characteristic function

  $$\chi_{[-\frac{1}{2},\frac{1}{2})}(t) = \begin{cases} 1 & \text{for } t \geq -\frac{1}{2} \wedge t < \frac{1}{2} \\ 0 & \text{for } t < -\frac{1}{2} \vee t \geq \frac{1}{2}, \end{cases}$$

  we can express $\mathcal{I}_0 g$ as

  $$\mathcal{I}_0 g(t) = \sum_{m \in \mathbb{Z}} g_m \cdot \chi_{[-\frac{1}{2},\frac{1}{2})}\left(\frac{t}{d} - m\right) \quad \forall\, t \in \mathbb{R}.$$

- **Linear spline interpolation:**

  Let $t \in [t_m, t_{m+1})$ for some $m \in \mathbb{Z}$. Then, the function value $g(t)$ is approximated by

  $$\mathcal{I}_1 g(t) = \frac{1}{d} \left[(t - t_m)\, g(t_{m+1}) + (t_{m+1} - t)\, g(t_m) \right].$$

  This defines a piecewise linear interpolant $\mathcal{I}_1 g$ of g, which is globally continuous. Note that by considering the hat function

  $$\bigwedge(t) = \begin{cases} 1 - |t| & \text{for } |t| \leq 1 \\ 0 & \text{for } |t| > 1, \end{cases}$$

  we can express $\mathcal{I}_1 g$ as

  $$\mathcal{I}_1 g(t) = \sum_{m \in \mathbb{Z}} g_m \cdot \bigwedge\left(\frac{t}{d} - m\right) \quad \forall\, t \in \mathbb{R}.$$

We remark that for target functions f of low regularity it is sufficient to use linear spline interpolation in the discrete FBP method (5.3), whereas nearest neighbour interpolation is usually not satisfactory. This has also been found in practical tests in [106]. To exploit a higher regularity of f, we additionally use cubic spline interpolation with not-a-knot end conditions in our numerical simulations. Here, we refrain from giving the details on standard spline theory and instead refer the reader to the textbooks [3], [15], [56].

Finally, we remark that the different discretization steps introduce additional discretization errors that are not included in our error theory in Chapter 4. However, the discretization of the FBP method (5.1) is necessary for providing numerical examples and analysing the resulting discretization errors is beyond the aims and scopes of this thesis. For work in this direction, we refer to the standard reference [82, Section V.1] and to the research articles [32], [57].

5.1.3 Inverse Fourier transform of low-pass filters

The computation of the discretized approximate FBP reconstruction f_{FBP} in (5.3), i.e.,

$$f_{\mathrm{FBP}} = \frac{1}{2}\,\mathcal{B}_D\Big(\mathcal{I}\big[\mathcal{F}^{-1}A_L *_D \mathcal{R}f\big]\Big),$$

requires the evaluation of the inverse Fourier transform $\mathcal{F}^{-1}A_L$ of the utilized low-pass filter A_L at the sampling points

$$t_j = j \cdot \frac{\pi}{L} \quad \text{for } j \in \mathbb{Z}.$$

Thus, in the following we derive analytical expressions for the samples $\mathcal{F}^{-1}A_L\big(\frac{j\pi}{L}\big)$, for $j \in \mathbb{Z}$, for typical choices of the low-pass filter A_L, including the Ram-Lak, Shepp-Logan, Cosine, Hamming and Gaussian filters we introduced in Examples 3.2.8 – 3.2.12.

We begin with the Ram-Lak filter from Example 3.2.8, see also Example 3.3.8.

Proposition 5.1.4. *The inverse Fourier transform of the Ram-Lak filter*

$$A_L(S) = |S| \cdot \sqcap_L(S) \quad \textit{for } S \in \mathbb{R}$$

is given by

$$\mathcal{F}^{-1}A_L(t) = \frac{L^2}{2\pi}\Big(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2\big(Lt/2\big)\Big) \quad \forall\, t \in \mathbb{R}.$$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ at $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and sample spacing $d = \frac{\pi}{L}$ yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \begin{cases} \frac{L^2}{2\pi} & \textit{for } j = 0 \\ 0 & \textit{for } j \neq 0 \textit{ even} \\ -\frac{2L^2}{\pi^3 j^2} & \textit{for } j \neq 0 \textit{ odd.} \end{cases}$$

Proof. In Example 3.3.8 we have already seen that the inverse Fourier transform of A_L is given by

$$\mathcal{F}^{-1}A_L(t) = \frac{L^2}{2\pi}\Big(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2\big(Lt/2\big)\Big) \quad \forall\, t \in \mathbb{R}.$$

Thus, the evaluation at $t_j = \frac{j\pi}{L}$, for $j \in \mathbb{Z}$, yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{L^2}{2\pi}\Big(2\,\mathrm{sinc}(j\pi) - \mathrm{sinc}^2\big(j\pi/2\big)\Big) = \begin{cases} \frac{L^2}{2\pi} & \text{for } j = 0 \\ 0 & \text{for } j \neq 0 \text{ even} \\ -\frac{2L^2}{\pi^3 j^2} & \text{for } j \neq 0 \text{ odd.} \end{cases}$$

and the proof is complete. □

Let us continue with the Shepp-Logan filter from Example 3.2.9.

Proposition 5.1.5. *The inverse Fourier transform of the Shepp-Logan filter*

$$A_L(S) = |S| \cdot \mathrm{sinc}\Big(\frac{\pi S}{2L}\Big) \cdot \sqcap_L(S) \quad \textit{for } S \in \mathbb{R}$$

is given by

$$\mathcal{F}^{-1}A_L(t) = \frac{4L^2}{\pi^2}\,\frac{\pi - 2Lt\,\sin(Lt)}{\pi^2 - 4L^2t^2} \quad \forall\, t \in \mathbb{R}.$$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ at $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and sample spacing $d = \frac{\pi}{L}$ yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{4L^2}{\pi^3(1-4j^2)}.$$

Proof. Since A_L is even and compactly supported with $\operatorname{supp}(A_L) \subseteq [-L, L]$, its inverse Fourier transform is given by the inverse cosine transform and for all $t \in \mathbb{R}$ follows that

$$\begin{aligned}\mathcal{F}^{-1}A_L(t) &= \frac{1}{\pi}\int_0^L S\,\operatorname{sinc}\Big(\frac{\pi S}{2L}\Big)\cos(tS)\,\mathrm{d}S = \frac{L}{\pi^2}\int_0^L \sin\Big(\Big(\frac{\pi}{2L}-t\Big)S\Big) + \sin\Big(\Big(\frac{\pi}{2L}+t\Big)S\Big)\,\mathrm{d}S \\ &= \frac{L}{\pi^2}\left(\frac{\cos(Lt-\pi/2)-1}{t-\pi/(2L)} - \frac{\cos(Lt+\pi/2)-1}{t+\pi/(2L)}\right) = \frac{4L^2}{\pi^2}\,\frac{\pi-2Lt\,\sin(Lt)}{\pi^2-4L^2t^2}\end{aligned}$$

by using the trigonometric identity

$$\sin(x)\cos(y) = \frac{\sin(x-y)+\sin(x+y)}{2} \quad \forall\, x, y \in \mathbb{R}.$$

Thus, by evaluating at $t_j = \frac{j\pi}{L}$, for $j \in \mathbb{Z}$, we obtain

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{4L^2\pi}{\pi^2(\pi^2-4j^2\pi^2)} = \frac{4L^2}{\pi^3(1-4j^2)},$$

as stated. □

We now consider the Cosine filter from Example 3.2.10.

Proposition 5.1.6. *The inverse Fourier transform of the Cosine filter*

$$A_L(S) = |S| \cdot \cos\Big(\frac{\pi S}{2L}\Big) \cdot \sqcap_L(S) \quad \text{for } S \in \mathbb{R}$$

is given by

$$\mathcal{F}^{-1}A_L(t) = \frac{2L^2\cos(Lt)}{\pi^2-4L^2t^2} + \frac{16L^3t\sin(Lt)}{(\pi^2-4L^2t^2)^2} - \frac{4L^2(\pi^2+4L^2t^2)}{\pi(\pi^2-4L^2t^2)^2} \quad \forall\, t \in \mathbb{R}.$$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ at $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and sample spacing $d = \frac{\pi}{L}$ yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{2L^2}{\pi^2}\left(\frac{(-1)^j}{1-4j^2} - \frac{2(1+4j^2)}{\pi(1-4j^2)^2}\right).$$

Proof. Since A_L is even, we again apply the inverse cosine transform and obtain, for all $t \in \mathbb{R}$,

$$\begin{aligned}\mathcal{F}^{-1}A_L(t) &= \frac{1}{\pi}\int_0^L S\,\cos\Big(\frac{\pi S}{2L}\Big)\cos(tS)\,\mathrm{d}S = \frac{1}{2\pi}\int_0^L S\left(\cos\left(tS-\frac{\pi S}{2L}\right)+\cos\left(tS+\frac{\pi S}{2L}\right)\right)\mathrm{d}S \\ &= \frac{L^2}{\pi}\left(\frac{\sin(Lt-\pi/2)}{2Lt-\pi} + \frac{\sin(Lt+\pi/2)}{2Lt+\pi}\right) - \frac{L}{\pi}\int_0^L \frac{\sin\big((t-\frac{\pi}{2L})S\big)}{2Lt-\pi} + \frac{\sin\big((t+\frac{\pi}{2L})S\big)}{2Lt+\pi}\,\mathrm{d}S\end{aligned}$$

by using the trigonometric identity

$$\cos(x)\cos(y) = \frac{\cos(x-y)+\cos(x+y)}{2} \quad \forall\, x, y \in \mathbb{R}.$$

Consequently, we have

$$\begin{aligned}\mathcal{F}^{-1}A_L(t) &= \frac{2L^2\cos(Lt)}{\pi^2-4L^2t^2} + \frac{2L^2}{\pi}\left(\frac{\cos(Lt-\pi/2)-1}{(2Lt-\pi)^2} + \frac{\cos(Lt+\pi/2)-1}{(2Lt+\pi)^2}\right) \\ &= \frac{2L^2\cos(Lt)}{\pi^2-4L^2t^2} + \frac{16L^3t\,\sin(Lt)}{(\pi^2-4L^2t^2)^2} - \frac{4L^2(\pi^2+4L^2t^2)}{\pi(\pi^2-4L^2t^2)^2}.\end{aligned}$$

Hence, for $t = \frac{j\pi}{L}$ with $j \in \mathbb{Z}$ follows that

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{2L^2\cos(j\pi)}{\pi^2-4j^2\pi^2} - \frac{4L^2(\pi^2+4j^2\pi^2)}{\pi(\pi^2-4j^2\pi^2)^2} = \frac{2L^2}{\pi^2}\left(\frac{(-1)^j}{1-4j^2} - \frac{2(1+4j^2)}{\pi(1-4j^2)^2}\right),$$

which completes the proof. □

The next proposition deals with the Hamming filter from Example 3.2.11.

Proposition 5.1.7. *The inverse Fourier transform of the Hamming filter*

$$A_L(S) = |S| \cdot \left(\beta + (1-\beta)\cos\left(\frac{\pi S}{L}\right)\right) \cdot \sqcap_L(S) \quad \text{for } S \in \mathbb{R}$$

with parameter $\beta \in \left[\frac{1}{2}, 1\right]$ *is given by*

$$\mathcal{F}^{-1}A_L(t) = \beta\, R_L(t) + (1-\beta)\, C_L(t) \quad \forall\, t \in \mathbb{R}$$

with

$$R_L(t) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2(Lt/2)\right) \quad \text{for } t \in \mathbb{R}$$

and

$$C_L(t) = \frac{L^2}{\pi}\left(\frac{Lt\sin(Lt)}{\pi^2 - L^2t^2} - \frac{(\cos(Lt)+1)\,(\pi^2+L^2t^2)}{(\pi^2-L^2t^2)^2}\right) \quad \text{for } t \in \mathbb{R}.$$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ *at* $t_j = j \cdot d$ *for* $j \in \mathbb{Z}$ *and sample spacing* $d = \frac{\pi}{L}$ *yields*

$$\mathcal{F}^{-1}A_L\left(\frac{j\pi}{L}\right) = \begin{cases} \beta\,\frac{L^2}{2\pi} - (1-\beta)\,\frac{2L^2}{\pi^3} & \text{for } j = 0 \\ (1-\beta)\,\frac{L^2}{4\pi} - \beta\,\frac{2L^2}{\pi^3} & \text{for } j \in \{-1,1\} \\ -(1-\beta)\,\frac{2L^2(1+j^2)}{\pi^3(1-j^2)^2} & \text{for } j \notin \{-1,0,1\} \text{ even} \\ -\beta\,\frac{2L^2}{\pi^3 j^2} & \text{for } j \notin \{-1,0,1\} \text{ odd.} \end{cases}$$

Proof. The Hamming filter is a combination of the Ram-Lak filter and a modified Cosine filter. Thus, for $t \in \mathbb{R}$ we have

$$\mathcal{F}^{-1}A_L(t) = \beta\, R_L(t) + (1-\beta)\, C_L(t),$$

where

$$R_L(t) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2(Lt/2)\right)$$

due to Proposition 5.1.4 and, analogous to the calculations in Proposition 5.1.6,

$$\begin{aligned} C_L(t) &= \frac{1}{\pi}\int_0^L S\,\cos\left(\frac{\pi S}{L}\right)\cos(tS)\,\mathrm{d}S = \frac{1}{2\pi}\int_0^L S\left(\cos\left(tS - \frac{\pi S}{L}\right) + \cos\left(tS + \frac{\pi S}{L}\right)\right)\,\mathrm{d}S \\ &= \frac{L^2}{2\pi}\left(\frac{\sin(Lt-\pi)}{Lt-\pi} + \frac{\sin(Lt+\pi)}{Lt+\pi}\right) - \frac{L}{2\pi}\int_0^L \frac{\sin\left((t-\frac{\pi}{L})S\right)}{Lt-\pi} + \frac{\sin\left((t+\frac{\pi}{L})S\right)}{Lt+\pi}\,\mathrm{d}S \\ &= \frac{L^2}{2\pi}\left(\frac{2Lt\sin(Lt)}{\pi^2-L^2t^2} + \frac{\cos(Lt-\pi)-1}{(Lt-\pi)^2} + \frac{\cos(Lt+\pi)-1}{(Lt+\pi)^2}\right) \\ &= \frac{L^2}{\pi}\left(\frac{Lt\sin(Lt)}{\pi^2-L^2t^2} - \frac{(\cos(Lt)+1)\,(\pi^2+L^2t^2)}{(\pi^2-L^2t^2)^2}\right). \end{aligned}$$

Thus, evaluating at $t_j = \frac{j\pi}{L}$, for $j \in \mathbb{Z}$, yields

$$\mathcal{F}^{-1}A_L\left(\frac{j\pi}{L}\right) = \beta\, R_L\left(\frac{j\pi}{L}\right) + (1-\beta)\, C_L\left(\frac{j\pi}{L}\right)$$

with

$$R_L\left(\frac{j\pi}{L}\right) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(j\pi) - \mathrm{sinc}^2(j\pi/2)\right) = \begin{cases} \frac{L^2}{2\pi} & \text{for } j = 0 \\ 0 & \text{for } j \neq 0 \text{ even} \\ -\frac{2L^2}{\pi^3 j^2} & \text{for } j \neq 0 \text{ odd} \end{cases}$$

and

$$C_L\left(\frac{j\pi}{L}\right) = \frac{L^2}{\pi}\left(\frac{j\sin(j\pi)}{\pi(1-j^2)} - \frac{(\cos(j\pi)+1)\,(1+j^2)}{\pi^2(1-j^2)^2}\right) = \begin{cases} \frac{L^2}{4\pi} & \text{for } j \in \{-1,1\} \\ 0 & \text{for } j \notin \{-1,1\} \text{ odd} \\ -\frac{2L^2(1+j^2)}{\pi^3(1-j^2)^2} & \text{for } j \notin \{-1,1\} \text{ even,} \end{cases}$$

where we use l'Hospital's rule in the last step. Combining the results completes the proof. □

Before we come to the inverse Fourier transform of the Gaussian filter from Example 3.2.12, we first recall the definitions of the *error function*

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \mathrm{e}^{-w^2}\,\mathrm{d}w \quad \text{for } z \in \mathbb{C}$$

and the *Dawson function*

$$D(z) = \mathrm{e}^{-z^2} \int_0^z \mathrm{e}^{w^2}\,\mathrm{d}w \quad \text{for } z \in \mathbb{C},$$

which are related via

$$\mathrm{i}\,\mathrm{e}^{z^2} D(z) = \frac{\sqrt{\pi}}{2} \operatorname{erf}(\mathrm{i}z) \quad \forall\, z \in \mathbb{C}. \tag{5.4}$$

Proposition 5.1.8. *The inverse Fourier transform of the Gaussian filter*

$$A_L(S) = |S| \cdot \exp\Big(-\Big(\frac{\pi S}{\beta L}\Big)^2 \Big) \cdot \sqcap_L(S) \quad \text{for } S \in \mathbb{R}$$

with parameter $\beta > 1$ is given by

$$\mathcal{F}^{-1}A_L(t) = -\frac{\beta^2 L^2}{2\pi^3} \left(\cos(Lt)\,\mathrm{e}^{-(\frac{\pi}{\beta})^2} - 1 + \frac{\beta L t}{\pi} D_L(t) \right) \quad \forall\, t \in \mathbb{R}$$

with

$$D_L(t) = \operatorname{Re}\left(D\Big(\frac{\beta L t}{2\pi}\Big) - \mathrm{e}^{\mathrm{i}Lt - (\frac{\pi}{\beta})^2} D\Big(\frac{\beta L t}{2\pi} + \frac{\mathrm{i}\pi}{\beta}\Big) \right) \quad \text{for } t \in \mathbb{R}.$$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ at $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and sample spacing $d = \frac{\pi}{L}$ yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{\beta^2 L^2}{2\pi^3} \left(1 - (-1)^j\,\mathrm{e}^{-(\frac{\pi}{\beta})^2} - \beta j \operatorname{Re}\left(D\Big(\frac{\beta j}{2}\Big) - (-1)^j\,\mathrm{e}^{-(\frac{\pi}{\beta})^2} D\Big(\frac{\beta j}{2} + \frac{\mathrm{i}\pi}{\beta}\Big) \right) \right).$$

Proof. Since A_L is even, its inverse Fourier transform is given by the inverse cosine transform and for $t \in \mathbb{R}$ follows that

$$\begin{aligned}
\mathcal{F}^{-1}A_L(t) &= \frac{1}{\pi} \int_0^L S \exp\Big(-\Big(\frac{\pi S}{\beta L}\Big)^2 \Big) \cos(tS)\,\mathrm{d}S = -\frac{\beta^2 L^2}{2\pi^3} \int_0^L \frac{\mathrm{d}}{\mathrm{d}S} \Big(\mathrm{e}^{-(\frac{\pi S}{\beta L})^2} \Big) \cos(tS)\,\mathrm{d}S \\
&= -\frac{\beta^2 L^2}{2\pi^3} \left(\cos(Lt)\,\mathrm{e}^{-(\frac{\pi}{\beta})^2} - 1 + t \int_0^L \sin(tS)\,\mathrm{e}^{-(\frac{\pi S}{\beta L})^2}\,\mathrm{d}S \right).
\end{aligned}$$

By using the identity

$$\sin(x) = \frac{1}{2\mathrm{i}} \left(\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x} \right) \quad \forall\, x \in \mathbb{R},$$

we obtain

$$\begin{aligned}
\int_0^L \sin(tS)\,\mathrm{e}^{-(\frac{\pi S}{\beta L})^2}\,\mathrm{d}S &= \frac{1}{2\mathrm{i}} \left(\int_0^L \mathrm{e}^{\mathrm{i}tS - (\frac{\pi S}{\beta L})^2}\,\mathrm{d}S - \int_0^L \mathrm{e}^{-\mathrm{i}tS - (\frac{\pi S}{\beta L})^2}\,\mathrm{d}S \right) \\
&= \frac{1}{2\mathrm{i}}\,\mathrm{e}^{-(\frac{\beta L t}{2\pi})^2} \left(\int_0^L \mathrm{e}^{-(\frac{\pi S}{\beta L} - \frac{\mathrm{i}\beta L t}{2\pi})^2}\,\mathrm{d}S - \int_0^L \mathrm{e}^{-(\frac{\pi S}{\beta L} + \frac{\mathrm{i}\beta L t}{2\pi})^2}\,\mathrm{d}S \right),
\end{aligned}$$

where

$$\begin{aligned}
\int_0^L \mathrm{e}^{-(\frac{\pi S}{\beta L} \pm \frac{\mathrm{i}\beta L t}{2\pi})^2}\,\mathrm{d}S &= \frac{\beta L}{\pi} \int_{\pm\mathrm{i}\frac{\beta L t}{2\pi}}^{\frac{\pi}{\beta} \pm \mathrm{i}\frac{\beta L t}{2\pi}} \mathrm{e}^{-w^2}\,\mathrm{d}w = \frac{\beta L}{\pi} \left(\int_0^{\frac{\pi}{\beta} \pm \mathrm{i}\frac{\beta L t}{2\pi}} \mathrm{e}^{-w^2}\,\mathrm{d}w - \int_0^{\pm\mathrm{i}\frac{\beta L t}{2\pi}} \mathrm{e}^{-w^2}\,\mathrm{d}w \right) \\
&= \frac{\beta L}{2\sqrt{\pi}} \left(\operatorname{erf}\Big(\frac{\pi}{\beta} \pm \mathrm{i}\frac{\beta L t}{2\pi}\Big) - \operatorname{erf}\Big(\pm\,\mathrm{i}\frac{\beta L t}{2\pi}\Big) \right).
\end{aligned}$$

This, in combination with the property

$$\operatorname{erf}(\overline{z}) = \overline{\operatorname{erf}(z)} \quad \forall\, z \in \mathbb{C},$$

implies that

$$\int_0^L \sin(tS)\, \mathrm{e}^{-(\frac{\pi S}{\beta L})^2}\, \mathrm{d}S = \frac{\beta L}{2\sqrt{\pi}}\, \mathrm{e}^{-(\frac{\beta L t}{2\pi})^2}\, \operatorname{Im}\Big(\operatorname{erf}\Big(\frac{\pi}{\beta} - \frac{\mathrm{i}\beta L t}{2\pi}\Big) - \operatorname{erf}\Big(- \frac{\mathrm{i}\beta L t}{2\pi}\Big)\Big).$$

Further, the relation (5.4) between the Dawson function D and the error function erf yields

$$\frac{\sqrt{\pi}}{2}\, \mathrm{e}^{-(\frac{\beta L t}{2\pi})^2}\, \operatorname{erf}\Big(- \frac{\mathrm{i}\beta L t}{2\pi}\Big) = \mathrm{i}\, D\Big(- \frac{\beta L t}{2\pi}\Big)$$

and

$$\frac{\sqrt{\pi}}{2}\, \mathrm{e}^{-(\frac{\beta L t}{2\pi})^2}\, \operatorname{erf}\Big(\frac{\pi}{\beta} - \frac{\mathrm{i}\beta L t}{2\pi}\Big) = \mathrm{i}\, \mathrm{e}^{\mathrm{i}Lt - (\frac{\pi}{\beta})^2}\, D\Big(- \frac{\beta L t}{2\pi} - \frac{\mathrm{i}\pi}{\beta}\Big)$$

so that

$$\int_0^L \sin(tS)\, \mathrm{e}^{-(\frac{\pi S}{\beta L})^2}\, \mathrm{d}S = \frac{\beta L}{\pi}\, \operatorname{Re}\Big(e^{\mathrm{i}Lt - (\frac{\pi}{\beta})^2}\, D\Big(- \frac{\beta L t}{2\pi} - \frac{\mathrm{i}\pi}{\beta}\Big) - D\Big(- \frac{\beta L t}{2\pi}\Big)\Big).$$

Since D is an odd function, combining the results now shows that

$$\mathcal{F}^{-1} A_L(t) = -\frac{\beta^2 L^2}{2\pi^3}\, \Big(\cos(Lt)\, \mathrm{e}^{-(\frac{\pi}{\beta})^2} - 1 + \frac{\beta L t}{\pi}\, D_L(t)\Big)$$

with

$$D_L(t) = \operatorname{Re}\Big(D\Big(\frac{\beta L t}{2\pi}\Big) - \mathrm{e}^{\mathrm{i}Lt - (\frac{\pi}{\beta})^2}\, D\Big(\frac{\beta L t}{2\pi} + \frac{\mathrm{i}\pi}{\beta}\Big)\Big).$$

Finally, the evaluation at $t = \frac{j\pi}{L}$, for $j \in \mathbb{Z}$, yields

$$\mathcal{F}^{-1} A_L\Big(\frac{j\pi}{L}\Big) = \frac{\beta^2 L^2}{2\pi^3}\, \Big(1 - (-1)^j\, \mathrm{e}^{-(\frac{\pi}{\beta})^2} - \beta j\, \operatorname{Re}\Big(D\Big(\frac{\beta j}{2}\Big) - (-1)^j\, \mathrm{e}^{-(\frac{\pi}{\beta})^2}\, D\Big(\frac{\beta j}{2} + \frac{\mathrm{i}\pi}{\beta}\Big)\Big)\Big),$$

as stated. □

In Example 4.3.2 we defined the generalized Ramp filter. The next proposition deals with its inverse Fourier transform.

Proposition 5.1.9. *The inverse Fourier transform of the generalized Ramp filter*

$$A_L(S) = \begin{cases} |S| & \text{for } |S| \le \beta L \\ \frac{1}{1-\beta}\Big(1 - \beta\lambda - \frac{1-\lambda}{L}\, |S|\Big)|S| & \text{for } \beta L < |S| \le L \\ 0 & \text{for } |S| > L \end{cases}$$

of width $\beta \in [0,1)$ and jump height $\lambda \in [0,1]$ is given by

$$\mathcal{F}^{-1} A_L(t) = \frac{L^2}{2\pi(1-\beta)}\, F_L(t) - \frac{2(1-\lambda)}{\pi(1-\beta)}\, G_L(t) \quad \forall\, t \in \mathbb{R}$$

with

$$F_L(t) = 2\lambda(1-\beta)\operatorname{sinc}(Lt) - (1-\beta\lambda)\operatorname{sinc}^2\Big(\frac{Lt}{2}\Big) + (1-\lambda)\beta^3 \operatorname{sinc}^2\Big(\frac{\beta L t}{2}\Big) \quad \text{for } t \in \mathbb{R}$$

and

$$G_L(t) = \begin{cases} \frac{L^2}{3}(\beta^3 - 1) & \text{for } t = 0 \\ \frac{\cos(Lt) - \operatorname{sinc}(Lt)}{t^2} - \beta\, \frac{\cos(\beta L t) - \operatorname{sinc}(\beta L t)}{t^2} & \text{for } t \ne 0. \end{cases}$$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ at $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and sample spacing $d = \frac{\pi}{L}$ yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \begin{cases} \frac{L^2}{6\pi}\left(\beta(1+\beta)(1-\lambda)+2\lambda+1\right) & \text{for } j=0 \\ \frac{L^2}{2\pi(1-\beta)}\,F_j - \frac{2(1-\lambda)L^2}{j^2\pi^3(1-\beta)}\,G_j & \text{for } j\neq 0 \end{cases}$$

with

$$F_j = (1-\lambda)\beta^3 \operatorname{sinc}^2\Big(\frac{\beta j\pi}{2}\Big) - (1-\beta\lambda)\operatorname{sinc}^2\Big(\frac{j\pi}{2}\Big)$$

and

$$G_j = (-1)^j - \beta\left(\cos(\beta j\pi) - \operatorname{sinc}(\beta j\pi)\right).$$

Proof. Due to the evenness of A_L we again apply the inverse cosine transform and, for $t \in \mathbb{R}$, by the definition of A_L follows that

$$\mathcal{F}^{-1}A_L(t) = \frac{1}{\pi}\int_0^{\beta L} S\,\cos(tS)\,\mathrm{d}S + \frac{1-\beta\lambda}{\pi(1-\beta)}\int_{\beta L}^{L} S\,\cos(tS)\,\mathrm{d}S - \frac{1-\lambda}{\pi L(1-\beta)}\int_{\beta L}^{L} S^2\,\cos(tS)\,\mathrm{d}S.$$

If $t = 0$, we obtain

$$\begin{aligned}\mathcal{F}^{-1}A_L(0) &= \frac{L^2\beta^2}{2\pi} + \frac{(1-\beta\lambda)(1-\beta^2)L^2}{2\pi(1-\beta)} - \frac{(1-\lambda)L^2(1-\beta^3)}{3\pi(1-\beta)} \\ &= \frac{L^2}{6\pi}\left(\beta(1+\beta)(1-\lambda)+2\lambda+1\right).\end{aligned}$$

On the other hand, for $t \neq 0$ we have

$$\int_0^{\beta L} S\,\cos(tS)\,\mathrm{d}S = \frac{\beta Lt\sin(\beta Lt)+\cos(\beta Lt)-1}{t^2}$$

and

$$\int_{\beta L}^{L} S\,\cos(tS)\,\mathrm{d}S = \frac{Lt\sin(Lt)+\cos(Lt)}{t^2} - \frac{\beta Lt\sin(\beta Lt)+\cos(\beta Lt)}{t^2}$$

as well as

$$\int_{\beta L}^{L} S^2\,\cos(tS)\,\mathrm{d}S = \frac{(L^2t^2-2)\sin(Lt)+2Lt\cos(Lt)}{t^3} - \frac{(\beta^2L^2t^2-2)\sin(\beta Lt)+2\beta Lt\cos(\beta Lt)}{t^3}.$$

Because $\cos(\varphi) = 1-2\sin^2(\varphi/2)$ for all $\varphi \in \mathbb{R}$, this implies that

$$\mathcal{F}^{-1}A_L(t) = \frac{L^2}{2\pi(1-\beta)}\,F_L(t) - \frac{2(1-\lambda)}{\pi(1-\beta)}\,G_L(t)$$

with

$$F_L(t) = 2\lambda(1-\beta)\operatorname{sinc}(Lt) - (1-\beta\lambda)\operatorname{sinc}^2(Lt/2) + (1-\lambda)\beta^3\operatorname{sinc}^2(\beta Lt/2) \quad \text{for } t\in\mathbb{R}$$

and

$$G_L(t) = \frac{\cos(Lt)-\operatorname{sinc}(Lt)}{t^2} - \beta\,\frac{\cos(\beta Lt)-\operatorname{sinc}(\beta Lt)}{t^2} \quad \text{for } t\in\mathbb{R},$$

where

$$\lim_{t\to 0} G_L(t) = \frac{L^2}{3}(\beta^3-1).$$

Finally, evaluating at $t_j = \frac{j\pi}{L}$, for $j \in \mathbb{Z}\setminus\{0\}$, yields

$$\mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{L^2}{2\pi(1-\beta)}\,F_L\Big(\frac{j\pi}{L}\Big) - \frac{2(1-\lambda)}{\pi(1-\beta)}\,G_L\Big(\frac{j\pi}{L}\Big)$$

with

$$F_L\Big(\frac{j\pi}{L}\Big) = (1-\lambda)\beta^3\operatorname{sinc}^2(\beta j\pi/2) - (1-\beta\lambda)\operatorname{sinc}^2(j\pi/2)$$

and

$$G_L\Big(\frac{j\pi}{L}\Big) = \frac{L^2}{j^2\pi^2}\Big((-1)^j - \beta\left(\cos(\beta j\pi) - \operatorname{sinc}(\beta j\pi)\right)\Big).$$

Combining the results completes the proof. □

We close this section with considering the generalized polynomial filter from Example 4.3.12. This involves a special integral representation of the generalized hypergeometric function ${}_1F_2$ as

$$ {}_1F_2\Big(a;\frac{1}{2},b;-z\Big) = \frac{\Gamma(b)}{\Gamma(a)\,\Gamma(b-a)} \int_0^1 (1-\sigma)^{b-a-1}\,\sigma^{a-1}\,\cos(2\sqrt{\sigma z})\,\mathrm{d}\sigma \quad \text{for } z \geq 0 \tag{5.5} $$

with parameters $b > a > 0$. For details, we refer the reader to the textbook [115].

Proposition 5.1.10. *The inverse Fourier transform of the generalized polynomial filter*

$$ A_L(S) = \begin{cases} |S|\left(1 - \frac{1-\beta}{L^\nu}\,|S|^\nu\right) & \text{for } |S| \leq L \\ 0 & \text{for } |S| > L \end{cases} $$

of order $\nu \in \mathbb{R}_{>0}$ and with jump height $\beta \in [0,1)$ is given by

$$ \mathcal{F}^{-1}A_L(t) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2\Big(\frac{Lt}{2}\Big)\right) - \frac{L^2(1-\beta)}{\pi(\nu+2)}\,{}_1F_2\Big(\frac{\nu}{2}+1;\frac{1}{2},\frac{\nu}{2}+2;-\frac{L^2t^2}{4}\Big) \quad \forall\, t \in \mathbb{R}. $$

Further, the evaluation of $\mathcal{F}^{-1}A_L$ at $t_j = j \cdot d$ for $j \in \mathbb{Z}$ and sample spacing $d = \frac{\pi}{L}$ yields

$$ \mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(j\pi) - \mathrm{sinc}^2\Big(\frac{j\pi}{2}\Big) - \frac{2(1-\beta)}{\nu+2}\,{}_1F_2\Big(\frac{\nu}{2}+1;\frac{1}{2},\frac{\nu}{2}+2;-\frac{j^2\pi^2}{4}\Big)\right). $$

Proof. Because A_L is even and supported in $[-L,L]$, we apply the inverse cosine transform and for $t \in \mathbb{R}$ follows that

$$ \mathcal{F}^{-1}A_L(t) = \frac{1}{\pi}\int_0^L A_L(S)\,\cos(tS)\,\mathrm{d}S = \frac{1}{\pi}\int_0^L S\cos(tS)\,\mathrm{d}S - \frac{1-\beta}{\pi L^\nu}\int_0^L S^{\nu+1}\cos(tS)\,\mathrm{d}S. $$

We already know that

$$ \frac{1}{\pi}\int_0^L S\cos(tS)\,\mathrm{d}S = \frac{L^2}{2\pi}\Big(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2(Lt/2)\Big). $$

Further, applying the substitution $\sigma = S^2$ and the integral representation (5.5) of ${}_1F_2$ yields

$$ \begin{aligned} \frac{1}{\pi L^\nu}\int_0^L S^{\nu+1}\cos(tS)\,\mathrm{d}S &= \frac{L^2}{\pi}\int_0^1 S^{\nu+1}\cos(LtS)\,\mathrm{d}S = \frac{L^2}{2\pi}\int_0^1 \sigma^{\nu/2}\cos\Big(2\sqrt{\sigma L^2t^2/4}\Big)\,\mathrm{d}\sigma \\ &= \frac{L^2}{2\pi}\,\frac{\Gamma(\frac{\nu}{2}+1)\Gamma(1)}{\Gamma(\frac{\nu}{2}+2)}\,{}_1F_2\Big(\frac{\nu}{2}+1;\frac{1}{2},\frac{\nu}{2}+2;-\frac{L^2t^2}{4}\Big) \\ &= \frac{L^2}{\pi(\nu+2)}\,{}_1F_2\Big(\frac{\nu}{2}+1;\frac{1}{2},\frac{\nu}{2}+2;-\frac{L^2t^2}{4}\Big). \end{aligned} $$

Thus, in total we have

$$ \mathcal{F}^{-1}A_L(t) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(Lt) - \mathrm{sinc}^2\Big(\frac{Lt}{2}\Big)\right) - \frac{L^2(1-\beta)}{\pi(\nu+2)}\,{}_1F_2\Big(\frac{\nu}{2}+1;\frac{1}{2},\frac{\nu}{2}+2;-\frac{L^2t^2}{4}\Big). $$

Finally, evaluating at $t_j = \frac{j\pi}{L}$, for $j \in \mathbb{Z}$, yields

$$ \mathcal{F}^{-1}A_L\Big(\frac{j\pi}{L}\Big) = \frac{L^2}{2\pi}\left(2\,\mathrm{sinc}(j\pi) - \mathrm{sinc}^2\Big(\frac{j\pi}{2}\Big)\right) - \frac{L^2(1-\beta)}{\pi(\nu+2)}\,{}_1F_2\Big(\frac{\nu}{2}+1;\frac{1}{2},\frac{\nu}{2}+2;-\frac{j^2\pi^2}{4}\Big) $$

and the proof is complete. □

5.2 Mathematical phantoms

Testing the accuracy of a reconstruction algorithm requires the exact knowledge of the function we wish to recover. Otherwise, we cannot determine differences between the true function and its reconstruction. Consequently, we have to apply the algorithm to a test object whose internal structure is known. Nevertheless, some obstacles may still occur.

Our discrete reconstruction algorithm, Algorithm 1, is based on sampled Radon data of the object's attenuation function. Although the internal structure of the test object is known, errors may be incurred during the sampling of the Radon data. In turn, these errors cause inaccuracies in the image reconstruction process. Unfortunately, we are not able to distinguish between the errors caused by the algorithm itself and the errors incurred by the data acquisition.

To resolve this problem, Shepp and Logan introduced the concept of a so called *mathematical phantom* in their innovative paper [113]. This denotes a simulated test object whose structure is entirely determined by mathematical formulas. Hence, we can analytically compute its Radon transform and, thus, no errors occur during the acquisition of the Radon data so that all observed errors in the image are cause by the discretized reconstruction algorithm we applied.

In this section we describe two different phantoms we use for the validation of our error theory from Chapter 4 and give analytic expressions for their Radon transforms. We start with the well-known *Shepp-Logan phantom*, which consists of ten ellipses of various sizes and simulates a cross-section of the human brain. Since the Shepp-Logan phantom is of low regularity, we additionally consider a smooth phantom of higher regularity to observe higher orders of convergence and, in particular, saturation at the rates given in Chapter 4.

5.2.1 Additional properties of the Radon transform

To calculate the Radon transform of the phantoms mentioned above, we need some additional basic properties of the Radon transform, namely its shift, scaling and rotation property.

We start with studying the effect of shifting the argument in the target function f.

Proposition 5.2.1 (Shift property of the Radon transform)**.** *Let $f \equiv f(x,y)$ be a bivariate function with Radon transform $\mathcal{R}f \equiv \mathcal{R}f(t,\theta)$. For a given vector $c = (c_x, c_y) \in \mathbb{R}^2$ we define the shifted function f_c via*

$$f_c(x,y) = f(x - c_x, y - c_y) \quad \text{for } (x,y) \in \mathbb{R}^2.$$

Then, the Radon transform $\mathcal{R}f_c$ of f_c is given by

$$\mathcal{R}f_c(t,\theta) = \mathcal{R}f(t - c_x\cos(\theta) - c_y\sin(\theta), \theta) \quad \forall\,(t,\theta) \in \mathbb{R} \times [0,\pi).$$

Proof. For fixed $(t,\theta) \in \mathbb{R} \times [0,\pi)$, the definition of the Radon transform $\mathcal{R}$ yields

$$\begin{aligned}\mathcal{R}f_c(t,\theta) &= \int_{\mathbb{R}} f_c(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}s \\ &= \int_{\mathbb{R}} f(t\cos(\theta) - s\sin(\theta) - c_x, t\sin(\theta) + s\cos(\theta) - c_y)\,\mathrm{d}s,\end{aligned}$$

where

$$\begin{aligned}t\cos(\theta) - s\sin(\theta) - c_x &= (t - c_x\cos(\theta))\cos(\theta) - (s + c_x\sin(\theta))\sin(\theta) \\ &= (t - c_x\cos(\theta) - c_y\sin(\theta))\cos(\theta) - (s + c_x\sin(\theta) - c_y\cos(\theta))\sin(\theta)\end{aligned}$$

and

$$\begin{aligned}t\sin(\theta) + s\cos(\theta) - c_y &= (t - c_y\sin(\theta))\sin(\theta) + (s - c_y\cos(\theta))\cos(\theta) \\ &= (t - c_x\cos(\theta) - c_y\sin(\theta))\sin(\theta) + (s + c_x\sin(\theta) - c_y\cos(\theta))\cos(\theta).\end{aligned}$$

Consequently, by substituting

$$\tau = t - c_x\cos(\theta) - c_y\sin(\theta) \quad \text{and} \quad \sigma = s + c_x\sin(\theta) - c_y\cos(\theta),$$

we obtain $\mathrm{d}\sigma = \mathrm{d}s$ and can conclude that

$$\begin{aligned}\mathcal{R}f_c(t,\theta) &= \int_{\mathbb{R}} f(\tau\cos(\theta) - \sigma\sin(\theta), \tau\sin(\theta) + \sigma\cos(\theta))\,\mathrm{d}\sigma = \mathcal{R}f(\tau,\theta)\\ &= \mathcal{R}f(t - c_x\cos(\theta) - c_y\sin(\theta), \theta),\end{aligned}$$

which completes the proof. □

We continue with the effect of scaling the argument in the target function f.

Proposition 5.2.2 (Scaling property of the Radon transform)**.** *Let $f \equiv f(x,y)$ be a bivariate function with Radon transform $\mathcal{R}f \equiv \mathcal{R}f(t,\theta)$. For given positive constants $a, b > 0$ we define the scaled function $f_{a,b}$ via*

$$f_{a,b}(x,y) = f\Big(\frac{x}{a}, \frac{y}{b}\Big) \quad \textit{for } (x,y) \in \mathbb{R}^2.$$

Then, the Radon transform $\mathcal{R}f_{a,b}$ of $f_{a,b}$ is given by

$$\mathcal{R}f_{a,b}(t,\theta) = \frac{ab}{c_{a,b}(\theta)}\,\mathcal{R}f\Big(\frac{t}{c_{a,b}(\theta)}, \operatorname{atan}\Big(\frac{b}{a}\tan(\theta)\Big)\Big) \quad \forall\,(t,\theta) \in \mathbb{R}\times[0,\pi),$$

where

$$c_{a,b}(\theta) = \sqrt{a^2\cos^2(\theta) + b^2\sin^2(\theta)} > 0$$

and

$$\operatorname{atan}\Big(\frac{b}{a}\tan(\theta)\Big) = \begin{cases} \arctan\left(\frac{b}{a}\tan(\theta)\right) & \textit{for } \sin(\theta)\cos(\theta) > 0\\ 0 & \textit{for } \sin(\theta) = 0\\ \frac{\pi}{2} & \textit{for } \cos(\theta) = 0\\ \arctan\left(\frac{b}{a}\tan(\theta)\right) + \pi & \textit{for } \sin(\theta)\cos(\theta) < 0 \end{cases} \in [0,\pi).$$

Proof. For fixed $(t,\theta) \in \mathbb{R}\times[0,\pi)$, the definition of the Radon transform yields

$$\begin{aligned}\mathcal{R}f_{a,b}(t,\theta) &= \int_{\mathbb{R}} f_{a,b}(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}s\\ &= \int_{\mathbb{R}} f\Big(\frac{t}{a}\cos(\theta) - \frac{s}{a}\sin(\theta), \frac{t}{b}\sin(\theta) + \frac{s}{b}\cos(\theta)\Big)\,\mathrm{d}s.\end{aligned}$$

By considering the modified angle

$$\vartheta = \operatorname{atan}\Big(\frac{b}{a}\tan(\theta)\Big) \in [0,\pi),$$

we have

$$\cos(\vartheta) = \frac{a\cos(\theta)}{c_{a,b}(\theta)} \quad \text{and} \quad \sin(\vartheta) = \frac{b\sin(\theta)}{c_{a,b}(\theta)}.$$

Consequently, using the relation

$$c_{a,b}^2(\theta) = a^2\cos^2(\theta) + b^2\sin^2(\theta)$$

we obtain

$$\begin{aligned}\frac{t}{a}\cos(\theta) - \frac{s}{a}\sin(\theta) &= \frac{t}{c_{a,b}(\theta)}\,\frac{a\cos(\theta)}{c_{a,b}(\theta)} - \frac{t}{a}\,\frac{a^2 - c_{a,b}^2(\theta)}{c_{a,b}^2(\theta)}\,\cos(\theta) - \frac{s}{a}\,\sin(\theta)\\ &= \frac{t}{c_{a,b}(\theta)}\,\cos(\vartheta) - \frac{c_{a,b}(\theta)}{ab}\left(s + t\,\frac{a^2 - b^2}{c_{a,b}^2(\theta)}\,\sin(\theta)\cos(\theta)\right)\sin(\vartheta)\end{aligned}$$

and

$$\begin{aligned}\frac{t}{b}\sin(\theta)+\frac{s}{b}\cos(\theta) &= \frac{t}{c_{a,b}(\theta)}\,\frac{b\sin(\theta)}{c_{a,b}(\theta)} - \frac{t}{b}\,\frac{b^2-c_{a,b}^2(\theta)}{c_{a,b}^2(\theta)}\,\sin(\theta)+\frac{s}{b}\,\cos(\theta)\\ &= \frac{t}{c_{a,b}(\theta)}\,\sin(\vartheta)+\frac{c_{a,b}(\theta)}{ab}\left(s+t\,\frac{a^2-b^2}{c_{a,b}^2(\theta)}\,\sin(\theta)\cos(\theta)\right)\cos(\vartheta).\end{aligned}$$

Therefore, by substituting

$$\tau=\frac{t}{c_{a,b}(\theta)}\quad\text{and}\quad\sigma=\frac{c_{a,b}(\theta)}{ab}\left(s+t\,\frac{a^2-b^2}{c_{a,b}^2(\theta)}\,\sin(\theta)\cos(\theta)\right),$$

we have

$$\mathrm{d}\sigma=\frac{c_{a,b}(\theta)}{ab}\,\mathrm{d}s$$

and can conclude that

$$\begin{aligned}\mathcal{R}f_c(t,\theta) &= \frac{ab}{c_{a,b}(\theta)}\int_{\mathbb{R}} f(\tau\cos(\vartheta)-\sigma\sin(\vartheta),\tau\sin(\vartheta)+\sigma\cos(\vartheta))\,\mathrm{d}\sigma=\frac{ab}{c_{a,b}(\theta)}\,\mathcal{R}f(\tau,\vartheta)\\ &= \frac{ab}{c_{a,b}(\theta)}\,\mathcal{R}f\Big(\frac{t}{c_{a,b}(\theta)},\operatorname{atan}\Big(\frac{b}{a}\tan(\theta)\Big)\Big),\end{aligned}$$

as stated. □

Finally, we come to the effect of rotating the argument in the target function f.

Proposition 5.2.3 (Rotation property of the Radon transform)**.** *Let $f\equiv f(x,y)$ be a bivariate function with Radon transform $\mathcal{R}f\equiv\mathcal{R}f(t,\theta)$. For a given rotation angle $\varphi\in[-\pi,\pi)$ we define the rotated function f_φ via*

$$f_\varphi(x,y)=f(x\cos(\varphi)+y\sin(\varphi),-x\sin(\varphi)+y\cos(\varphi))\quad\text{for }(x,y)\in\mathbb{R}^2.$$

Then, the Radon transform $\mathcal{R}f_\varphi$ of f_φ is given by

$$\mathcal{R}f_\varphi(t,\theta)=\mathcal{R}f(t,\theta-\varphi)\quad\forall\,(t,\theta)\in\mathbb{R}\times[0,\pi).$$

Proof. For fixed $(t,\theta)\in\mathbb{R}\times[0,\pi)$, the definition of the Radon transform $\mathcal{R}$ yields

$$\mathcal{R}f_\varphi(t,\theta)=\int_{\mathbb{R}} f_\varphi(t\cos(\theta)-s\sin(\theta),t\sin(\theta)+s\cos(\theta))\,\mathrm{d}s=\int_{\mathbb{R}} f(x(s),y(s))\,\mathrm{d}s$$

with

$$\begin{aligned}x(s) &= (t\cos(\theta)-s\sin(\theta))\cos(\varphi)+(t\sin(\theta)+s\cos(\theta))\sin(\varphi)\\ &= t(\cos(\theta)\cos(\varphi)+\sin(\theta)\sin(\varphi))-s(\sin(\theta)\cos(\varphi)-\cos(\theta)\sin(\varphi))\\ &= t\cos(\theta-\varphi)-s\sin(\theta-\varphi)\end{aligned}$$

and

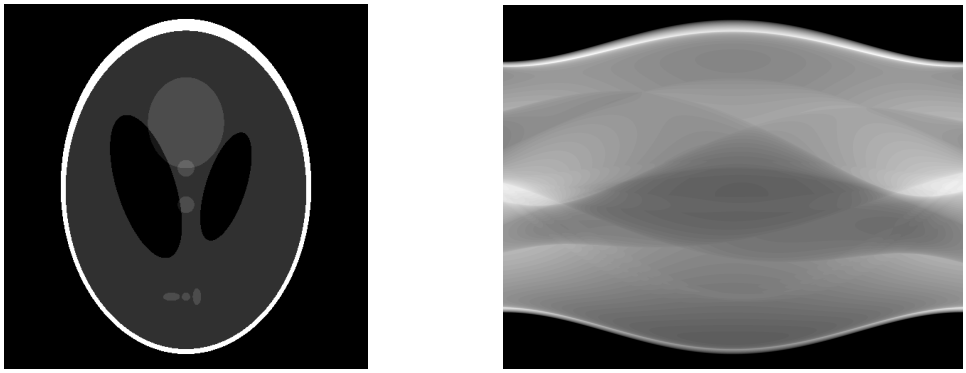
$$\begin{aligned}y(s) &= -(t\cos(\theta)-s\sin(\theta))\sin(\varphi)+(t\sin(\theta)+s\cos(\theta))\cos(\varphi)\\ &= t(\sin(\theta)\cos(\varphi)-\cos(\theta)\sin(\varphi))+s(\sin(\theta)\sin(\varphi)+\cos(\theta)\cos(\varphi))\\ &= t\sin(\theta-\varphi)+s\cos(\theta-\varphi).\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}\mathcal{R}f_\varphi(t,\theta) &= \int_{\mathbb{R}} f(t\cos(\theta-\varphi)-s\sin(\theta-\varphi),t\sin(\theta-\varphi)+s\cos(\theta-\varphi))\,\mathrm{d}s\\ &= \mathcal{R}f(t,\theta-\varphi)\end{aligned}$$

and the proof is complete. □

We are now prepared to describe the phantoms we use in our numerical experiments and, moreover, to determine analytic expressions for their Radon transforms.



(a) Phantom (b) Sinogram

Figure 5.2: The Shepp-Logan phantom and its sinogram.

5.2.2 The Shepp-Logan phantom

We start with the well-known Shepp-Logan phantom, which was introduced by Shepp and Logan in [113] and provides a standard test case for tomographic reconstruction methods. It schematically depicts a highly simplified cross-section of the human head and consists of ten ellipses of constant densities, but different sizes, eccentricities and locations, see Figure 5.2(a). In our version, however, we modified the densities of the different ellipses compared to the original phantom in order to get a higher contrast in the image for a better visual perception.

The corresponding attenuation function f_{SL} of the Shepp-Logan phantom is given by

$$f_{\mathrm{SL}} = \sum_{j=1}^{10} c_j \, f_j,$$

where each function f_j, for $1 \leq j \leq 10$, is of the form of the characteristic function f_e of an ellipse, defined as

$$f_e(x,y) = \chi_{B_1(0)}\Big(\frac{x_r}{a}, \frac{y_r}{b}\Big) \quad \text{for } (x,y) \in \mathbb{R}^2$$

with

$$x_r = (x-h)\cos(\varphi) + (y-k)\sin(\varphi) \quad \text{and} \quad y_r = -(x-h)\sin(\varphi) + (y-k)\cos(\varphi),$$

where the parameters a, b, h, k and φ denote the following attributes of the ellipse:

a: major axis, h: x-coordinate of the center, φ: rotation angle,
b: minor axis, k: y-coordinate of the center.

The parameters of the ellipses the Shepp-Logan phantom consists of are summarized in Table 5.1.

Before we compute the Radon transform of the Shepp-Logan phantom, we first determine the Radon transform $\mathcal{R}f_e$ of the characteristic function f_e of an ellipse with parameters a, b, h, k, φ.

j	1	2	3	4	5	6	7	8	9	10
a_j	0.69	0.6624	0.11	0.16	0.21	0.046	0.046	0.046	0.023	0.023
b_j	0.92	0.874	0.31	0.41	0.25	0.046	0.046	0.023	0.023	0.046
h_j	0	0	0.22	-0.22	0	0	0	-0.08	0	0.06
k_j	0	0.0184	0	0	$-0,35$	$-0,1$	0.1	0.605	0.605	0.605
φ_j	0	0	$\frac{\pi}{10}$	$-\frac{\pi}{10}$	0	0	0	0	0	0
c_j	1	-0.8	-0.2	-0.2	0.1	0.1	0.1	0.1	0.1	0.1

Table 5.1: Parameters of the ellipses in the Shepp-Logan phantom.

Proposition 5.2.4 (Radon transform of an ellipse)**.** *We consider the characteristic function f_e of an ellipse with parameters $a, b > 0$, $h, k \in \mathbb{R}$ and $\varphi \in [-\pi, \pi)$, i.e.,*

$$f_e(x,y) = \chi_{B_1(0)}\Big(\frac{x_r}{a}, \frac{y_r}{b}\Big) \quad \text{for } (x,y) \in \mathbb{R}^2$$

with

$$x_r = (x-h)\cos(\varphi) + (y-k)\sin(\varphi) \quad \text{and} \quad y_r = -(x-h)\sin(\varphi) + (y-k)\cos(\varphi).$$

Then, the Radon transform $\mathcal{R}f_e$ of f_e is given by

$$\mathcal{R}f_e(t,\theta) = \frac{2ab}{c^2_{a,b,\varphi}(\theta)} \sqrt{c^2_{a,b,\varphi}(\theta) - t^2_{h,k}(t,\theta)}\, \prod_{c_{a,b,\varphi}(\theta)}(t_{h,k}(t,\theta)) \quad \forall\, (t,\theta) \in \mathbb{R} \times [0,\pi)$$

with

$$c_{a,b,\varphi}(\theta) = \sqrt{a^2\cos^2(\theta-\varphi) + b^2\sin^2(\theta-\varphi)} \quad \text{and} \quad t_{h,k}(t,\theta) = t - h\cos(\theta) - k\sin(\theta).$$

Proof. For the sake of brevity, we define the functions

$$g(x,y) = \chi_{B_1(0)}\Big(\frac{x}{a}, \frac{y}{b}\Big) \quad \text{for } (x,y) \in \mathbb{R}^2$$

and

$$h(x,y) = g(x\cos(\varphi) + y\sin(\varphi), -x\sin(\varphi) + y\cos(\varphi)) \quad \text{for } (x,y) \in \mathbb{R}^2$$

so that the function f_e can be written as

$$f_e(x,y) = h(x-h, y-k) \quad \forall\, (x,y) \in \mathbb{R}^2.$$

Therefore, we can determine $\mathcal{R}f_e$ by applying the shift, scaling and rotation properties of the Radon transform. In Example 2.2.14 we have seen that the Radon transform of the characteristic function $\chi_{B_R(0)}$ of a ball with radius $R > 0$ around 0 is given by

$$\mathcal{R}\chi_{B_R(0)}(t,\theta) = \begin{cases} 2\sqrt{R^2 - t^2} & \text{for } |t| \le R \\ 0 & \text{for } |t| > R \end{cases} = 2\sqrt{R^2 - t^2}\, \prod_R(t) \quad \forall\, (t,\theta) \in \mathbb{R} \times [0,\pi).$$

Consequently, for fixed $(t,\theta) \in \mathbb{R} \times [0,\pi)$, applying the scaling property of the Radon transform, Proposition 5.2.2, yields

$$\begin{aligned} \mathcal{R}g(t,\theta) &= \frac{ab}{\sqrt{a^2\cos^2(\theta) + b^2\sin^2(\theta)}}\, \mathcal{R}\chi_{B_1(0)}\Bigg(\frac{t}{\sqrt{a^2\cos^2(\theta) + b^2\sin^2(\theta)}}, \operatorname{atan}\Big(\frac{b}{a}\tan(\theta)\Big)\Bigg) \\ &= \frac{2ab}{a^2\cos^2(\theta) + b^2\sin^2(\theta)} \sqrt{a^2\cos^2(\theta) + b^2\sin^2(\theta) - t^2}\, \prod_{\sqrt{a^2\cos^2(\theta)+b^2\sin^2(\theta)}}(t). \end{aligned}$$

Further, by defining

$$c_{a,b,\varphi}(\theta) = \sqrt{a^2\cos^2(\theta-\varphi) + b^2\sin^2(\theta-\varphi)}$$

the rotation property of the Radon transform, Proposition 5.2.3, shows that

$$\mathcal{R}h(t,\theta) = \mathcal{R}g(t,\theta-\varphi) = \frac{2ab}{c^2_{a,b,\varphi}(\theta)} \sqrt{c^2_{a,b,\varphi}(\theta) - t^2}\, \prod_{c_{a,b,\varphi}(\theta)}(t).$$

Finally, by applying the shift property of the Radon transform, Proposition 5.2.1, and setting

$$t_{h,k}(t,\theta) = t - h\cos(\theta) - k\sin(\theta),$$

for the Radon transform $\mathcal{R}f_e$ of f_e follows that

$$\mathcal{R}f_e(t,\theta) = \mathcal{R}h(t - h\cos(\theta) - k\sin(\theta), \theta) = \frac{2ab}{c^2_{a,b,\varphi}(\theta)} \sqrt{c^2_{a,b,\varphi}(\theta) - t^2_{h,k}(t,\theta)}\, \prod_{c_{a,b,\varphi}(\theta)}(t_{h,k}(t,\theta)),$$

as stated. □

With Proposition 5.2.4 we can now determine the Radon transform $\mathcal{R}f_{\mathrm{SL}}$ of the Shepp-Logan phantom. Indeed, due to the linearity of $\mathcal{R}$ we obtain

$$\mathcal{R}f_{\mathrm{SL}}(t,\theta) = \sum_{j=1}^{10} c_j \, \mathcal{R}f_j(t,\theta) \quad \forall\, (t,\theta) \in \mathbb{R} \times [0,\pi),$$

where

$$\mathcal{R}f_j(t,\theta) = \frac{2a_jb_j}{c^2_{a_j,b_j,\varphi_j}(\theta)} \sqrt{c^2_{a_j,b_j,\varphi_j}(\theta) - t^2_{h_j,k_j}(t,\theta)} \, \chi_{[-c_{a_j,b_j,\varphi_j}(\theta),\, c_{a_j,b_j,\varphi_j}(\theta)]}(t_{h_j,k_j}(t,\theta))$$

with

$$c_{a_j,b_j,\varphi_j}(\theta) = \sqrt{a_j^2 \cos^2(\theta-\varphi_j) + b_j^2 \sin^2(\theta-\varphi_j)} \quad \text{for } \theta \in [0,\pi)$$

and

$$t_{h_j,k_j}(t,\theta) = t - h_j\cos(\theta) - k_j \sin(\theta) \quad \text{for } (t,\theta) \in \mathbb{R} \times [0,\pi).$$

The sinogram of the Shepp-Logan phantom f_{SL}, i.e., the plot of its Radon transform $\mathcal{R}f_{\mathrm{SL}}$ in the (t,θ)-plane, is shown in Figure 5.2(b).

5.2.3 The smooth phantom

As explained at the beginning of Section 4.2, the attenuation function f_{SL} of the Shepp-Logan phantom belongs to the Sobolev space $\mathrm{H}^\alpha_0(\mathbb{R}^2)$ with $\alpha < \frac{1}{2}$, which in turn is an upper bound for the convergence rate of the inherent FBP reconstruction error $e_L = f_{\mathrm{SL}} - f_L$ in the L^2-norm. To observe higher orders of convergence and saturation at the rates given in Chapter 4 we need a test case of higher regularity and with an analytically computable Radon transform.

To this end, we consider the radially symmetric *bump-shaped function* $p_\nu : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$p_\nu(x,y) = \begin{cases} \big(1-x^2-y^2\big)^\nu & \text{for } x^2+y^2 \le 1 \\ 0 & \text{for } x^2+y^2 > 1 \end{cases}$$

with smoothness parameter $\nu \in \mathbb{R}_{>0}$, which is in $\mathrm{H}^\alpha_0(\mathbb{R}^2)$ for any $\alpha < \nu + \frac{1}{2}$. Adapting the approach in [100], we now define the *smooth phantom of order* ν via

$$f_{\mathrm{smooth}} = f_1 - \frac{3}{2} f_2 + \frac{3}{2} f_3 \in \mathrm{H}^\alpha_0(\mathbb{R}^2) \quad \forall\, \alpha < \nu + \frac{1}{2},$$

where each function f_j, for $j = 1,2,3$, is given by

$$f_j(x,y) = p_\nu\Big(\frac{x_r}{a_j}, \frac{y_r}{b_j}\Big) \quad \text{for } (x,y) \in \mathbb{R}^2$$

with

$$x_r = (x-h_j)\cos(\varphi_j) + (y-k_j)\sin(\varphi_j) \quad \text{and} \quad y_r = -(x-h_j)\sin(\varphi_j) + (y-k_j)\cos(\varphi_j)$$

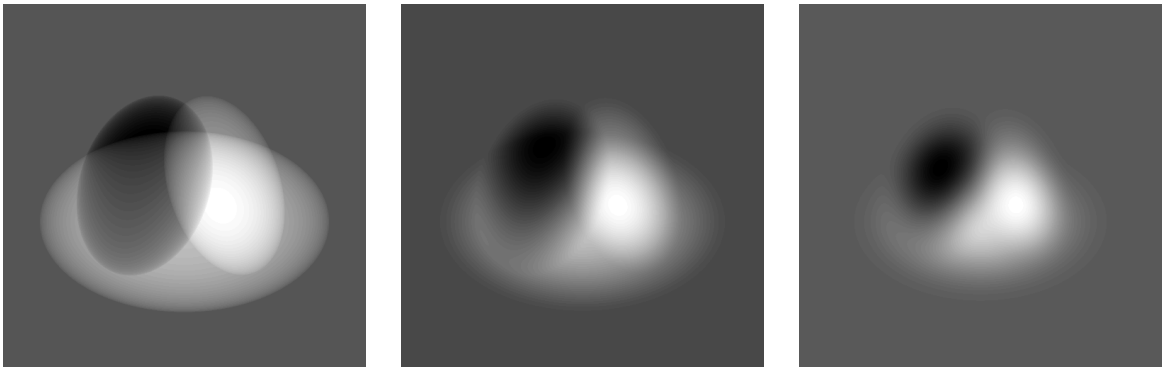
and the parameters

$$\begin{aligned}
&a_1 = 0.51, && b_1 = 0.31, && h_1 = 0.22, && k_1 = 0, && \varphi_1 = \frac{2}{5}\pi, \\
&a_2 = 0.51, && b_2 = 0.36, && h_2 = -0.22, && k_2 = 0, && \varphi_2 = \frac{3}{5}\pi, \\
&a_3 = 0.5, && b_3 = 0.8, && h_3 = 0, && k_3 = 0.2, && \varphi_3 = \frac{1}{2}\pi.
\end{aligned}$$

For illustration, Figure 5.3 shows the smooth phantom f_{smooth} of order $\nu \in \{0.5, 1.5, 3\}$, which satisfies

$$f_{\mathrm{smooth}} \in \mathrm{H}^\alpha_0(\mathbb{R}^2) \quad \forall\, \alpha < \alpha_{\max}$$

with maximal smoothness $\alpha_{\max} \in \{1, 2, 3.5\}$, respectively.



(a) $\nu = 0.5$ (b) $\nu = 1.5$ (c) $\nu = 3$

Figure 5.3: The smooth phantom of order $\nu \in \{0.5, 1.5, 3\}$.

Before we compute the Radon transform of the smooth phantom, we first have to determine the Radon transform $\mathcal{R}p_\nu$ of the bump-shaped function p_ν with parameter $\nu > 0$.

Proposition 5.2.5 (Radon transform of p_ν)**.** *We consider the radially symmetric bump-shaped function $p_\nu : \mathbb{R}^2 \longrightarrow \mathbb{R}$ with smoothness parameter $\nu > 0$, i.e.,*

$$p_\nu(x,y) = \begin{cases} (1-x^2-y^2)^\nu & \text{for } x^2+y^2 \le 1 \\ 0 & \text{for } x^2+y^2 > 1. \end{cases}$$

Then, the Radon transform $\mathcal{R}p_\nu$ of p_ν is given by

$$\mathcal{R}p_\nu(t,\theta) = \frac{\sqrt{\pi}\,\Gamma(\nu+1)}{\Gamma(\nu+\frac{3}{2})}\,(1-t^2)^{\nu+\frac{1}{2}}\,\bigsqcap(t) \quad \forall\,(t,\theta) \in \mathbb{R}\times[0,\pi).$$

Proof. For fixed $(t,\theta) \in \mathbb{R}\times[0,\pi)$ the definition of the Radon transform $\mathcal{R}$ yields

$$\mathcal{R}p_\nu(t,\theta) = \int_{\mathbb{R}} p_\nu(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta))\,\mathrm{d}s,$$

where

$$p_\nu(t\cos(\theta) - s\sin(\theta), t\sin(\theta) + s\cos(\theta)) = \begin{cases} (1-t^2-s^2)^\nu & \text{for } t^2+s^2 \le 1 \\ 0 & \text{for } t^2+s^2 > 1 \end{cases}$$

by using (2.5). Therefore, for $|t| \ge 1$ we have

$$\mathcal{R}p_\nu(t,\theta) = 0$$

and, for $|t| < 1$,

$$\mathcal{R}p_\nu(t,\theta) = \int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}} (1-t^2-s^2)^\nu\,\mathrm{d}s = 2\int_0^{\sqrt{1-t^2}} (1-t^2-s^2)^\nu\,\mathrm{d}s.$$

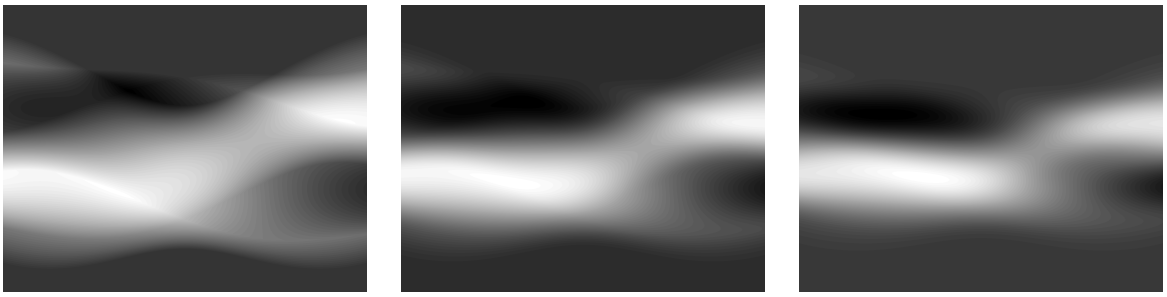
Substituting $s = \sqrt{1-t^2}\,\sigma$ gives $\mathrm{d}s = \sqrt{1-t^2}\,\mathrm{d}\sigma$ and, thus, we have

$$\mathcal{R}p_\nu(t,\theta) = 2\int_0^1 (1-t^2-(1-t^2)\sigma^2)^\nu\,\sqrt{1-t^2}\,\mathrm{d}\sigma = 2(1-t^2)^{\nu+\frac{1}{2}}\int_0^1 (1-\sigma^2)^\nu\,\mathrm{d}\sigma.$$

Further, by applying the transformation $\sigma = \sqrt{s}$ we obtain $\mathrm{d}\sigma = \frac{1}{2}\,s^{-\frac{1}{2}}\,\mathrm{d}s$ so that

$$2\int_0^1 (1-\sigma^2)^\nu\,d\sigma = \int_0^1 (1-s)^\nu\,s^{-\frac{1}{2}}\,\mathrm{d}s = B\Big(\frac{1}{2},\nu+1\Big) = \frac{\Gamma\big(\frac{1}{2}\big)\,\Gamma(\nu+1)}{\Gamma\big(\nu+\frac{3}{2}\big)} = \frac{\sqrt{\pi}\,\Gamma(\nu+1)}{\Gamma\big(\nu+\frac{3}{2}\big)},$$

where B denotes the beta function and we used its close relation to the gamma function Γ.



(a) $\nu = 0.5$ (b) $\nu = 1.5$ (c) $\nu = 3$

Figure 5.4: The sinogram of the smooth phantom of order $\nu \in \{0.5, 1.5, 3\}$.

Hence, for $|t| < 1$ follows that

$$\mathcal{R}p_\nu(t,\theta) = \frac{\sqrt{\pi}\,\Gamma(\nu+1)}{\Gamma(\nu+\frac{3}{2})}\,(1-t^2)^{\nu+\frac{1}{2}}$$

and the proof is complete. □

With Proposition 5.2.5 we are now prepared to determine the Radon transform $\mathcal{R}f_{\text{smooth}}$ of the smooth phantom of order $\nu > 0$. Using the linearity of the Radon transform $\mathcal{R}$ and following along the lines of the proof of Proposition 5.2.4, we obtain

$$\mathcal{R}f_{\text{smooth}}(t,\theta) = \mathcal{R}f_1(t,\theta) - \frac{3}{2}\,\mathcal{R}f_2(t,\theta) + \frac{3}{2}\,\mathcal{R}f_3(t,\theta) \quad \forall\,(t,\theta) \in \mathbb{R} \times [0,\pi),$$

where

$$\mathcal{R}f_j(t,\theta) = \frac{\sqrt{\pi}\,\Gamma(\nu+1)}{\Gamma(\nu+\frac{3}{2})}\,\frac{a_j b_j}{c^{2\nu+2}_{a_j,b_j,\varphi_j}(\theta)}\,\big(c^2_{a_j,b_j,\varphi_j}(\theta) - t^2_{h_j,k_j}(t,\theta)\big)^{\nu+\frac{1}{2}}\,\textstyle\prod_{c_{a_j,b_j,\varphi_j}(\theta)}(t_{h_j,k_j}(t,\theta))$$

with

$$c_{a_j,b_j,\varphi_j}(\theta) = \sqrt{a_j^2\cos^2(\theta-\varphi_j) + b_j^2\sin^2(\theta-\varphi_j)} \quad \text{for } \theta \in [0,\pi)$$

and

$$t_{h_j,k_j}(t,\theta) = t - h_j\cos(\theta) - k_j\sin(\theta) \quad \text{for } (t,\theta) \in \mathbb{R}\times[0,\pi).$$

The sinogram of the smooth phantom f_{smooth} of order $\nu \in \{0.5, 1.5, 3\}$ is shown in Figure 5.4.

5.3 Numerical investigation of the reconstruction error

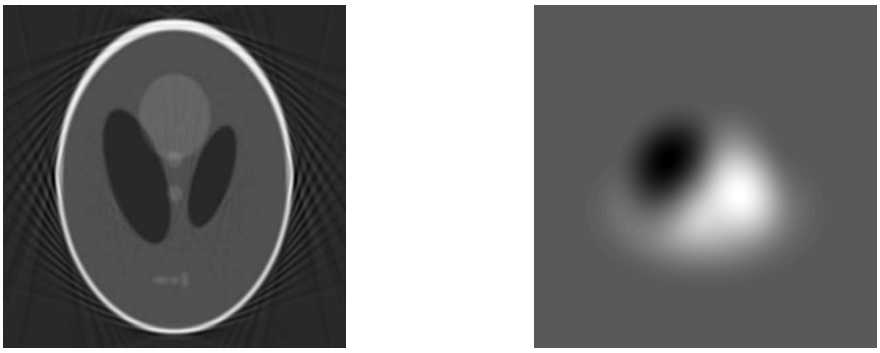
As announced in the introductory paragraph of this chapter, we henceforth investigate the reconstruction error of the FBP method numerically to validate our error theory from Chapter 4. To this end, we use the discrete image reconstruction Algorithm 1, which we derived in Section 5.1 and which is based on the discrete FBP reconstruction formula (5.3), i.e.,

$$f_{\text{FBP}} = \frac{1}{2}\,\mathcal{B}_D\Big(\mathcal{I}\big[\mathcal{F}^{-1}A_L *_D \mathcal{R}f\big]\Big).$$

As test cases we use the Shepp-Logan phantom and the smooth phantom of order $\nu = 3$ we introduced in the previous Section 5.2. For illustration, the FBP reconstructions of both phantoms are displayed in Figure 5.5, where we used linear interpolation for the Shepp-Logan and cubic spline interpolation for the smooth phantom. Further, we applied the Ram-Lak filter from Example 3.2.8 with window

$$W(S) = \textstyle\prod(S) \quad \text{for } S \in \mathbb{R}$$

and bandwidth $L = 50\pi$. This corresponds to $M = 50$ and $N = 150$ so that $(2M+1)N = 15150$ equally distributed Radon samples are taken.



(a) Shepp-Logan phantom (b) Smooth phantom with $\nu = 3$

Figure 5.5: FBP reconstructions of the phantoms with the Ram-Lak filter and $L = 50\pi$.

To measure the reconstruction error and, therefore, the approximation quality of the FBP reconstructions, we use the standard *root mean square error* (RMSE), which is defined for images with $J \times K$ pixels as

$$\mathrm{RMSE} = \sqrt{\frac{1}{J \cdot K} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(f_{j,k} - (f_{\mathrm{FBP}})_{j,k}\right)^2}.$$

In our numerical experiments, we evaluated the phantoms and their FBP reconstructions with different window functions and bandwidths on a square grid with 512×512 pixels.

5.3.1 Validation of the order of convergence

We start with numerically validating the rates of convergence for the L^2-norm of the inherent FBP reconstruction error $e_L = f - f_L$ we have proven in Section 4.3. As a representative result, we first recall that for target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$ and window functions $W \in \mathcal{AC}([-1,1])$ with $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $1 \leq j \leq k-1$, for $k \in \mathbb{N}$, and $W^{(k)} \in \mathrm{L}^p([-1,1])$ with $1 < p \leq \infty$ such that

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1$$

the L^2-norm of e_L is bounded above by

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \left(c_{\alpha,k,p}\, \|W^{(k)}\|_{\mathrm{L}^p([0,1])}\, L^{-\min\{k-1/p,\alpha\}} + L^{-\alpha}\right) \|f\|_\alpha \tag{5.6}$$

with some constant $c_{\alpha,k,p} > 0$ independent of W and f, see Corollaries 4.3.9 and 4.3.11. In particular, our theory predicts that the decay rate of the error is of order $\mathcal{O}(L^{-\alpha})$ for $\alpha \leq k - \frac{1}{p}$, but saturates at (possibly) fractional order $\mathcal{O}(L^{-(k-1/p)})$ for $\alpha > k - \frac{1}{p}$.

In our first numerical simulations we have employed the three commonly used low-pass filters $A_L(S) = |S| W(S/L)$ from Examples 3.2.9, 3.2.11 and 3.2.12:

Name	$W(S)$ for $\lvert S\rvert \leq 1$	Parameter
Shepp-Logan	$\mathrm{sinc}(\pi S/2)$	–
Hamming	$\beta + (1-\beta)\cos(\pi S)$	$\beta \in [1/2, 1]$
Gaussian	$\exp\left(-(\pi S/\beta)^2\right)$	$\beta > 1$

Recall that each window function W is even and compactly supported with $\mathrm{supp}(W) = [-1,1]$. Further, W is twice continuously differentiable on the interval $[-1,1]$, $W \in \mathscr{C}^2([-1,1])$, with

$$W(0) = 1, \quad W'(0) = 0, \quad W''(0) \neq 0.$$

Consequently, these windows satisfy the assumptions of Corollary 4.3.9 with $k = 2$ and our error theory states that, for any function $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with smoothness $\alpha > 0$, the L^2-norm of the inherent FBP reconstruction error is bounded above by

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \Big(C_\alpha \, \|W''\|_{\mathrm{L}^\infty([0,1])} + 1\Big) \, L^{-\min\{2,\alpha\}} \, \|f\|_\alpha. \tag{5.7}$$

Figure 5.6 shows the RMSE of the FBP reconstruction of the Shepp-Logan phantom f_{SL} as a function of the bandwidth L in logarithmic scales for different window functions. In addition to the popular Shepp-Logan filter (Figure 5.6(a)), we applied the Hamming filter with parameter $\beta = 0.92$ (Figure 5.6(b)) and the Gaussian filter with parameter $\beta = 4.9$ (Figure 5.6(c)). These parameters were chosen such that the Hamming and Gaussian filters have the same value for the quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$ as the Shepp-Logan filter, see Table 5.2 below. Hence, the corresponding reconstruction errors should behave similarly due to our error estimate (5.7).

As expected, we see that the RMSE for the Shepp-Logan filter, the Hamming filter with $\beta = 0.92$ and the Gaussian filter with $\beta = 4.9$ are nearly the same. Moreover, in all three cases we observe a decrease of the RMSE with rate $L^{-0.5}$. This is exactly the behaviour we expected due to our L^2-error estimate (5.7), since $f_{\mathrm{SL}} \in \mathrm{H}^\alpha(\mathbb{R}^2)$ for all $\alpha < \frac{1}{2}$.

Figure 5.7 now shows the RMSE of the FBP reconstruction for the smooth phantom f_{smooth} of order $\nu = 3$, which belongs to $\mathrm{H}^\alpha(\mathbb{R}^2)$ for all $\alpha < \frac{7}{2}$. Hence, according to the estimate (5.7), the convergence rate of the RMSE should saturate at order L^{-2}. Indeed, this behaviour can be observed in our numerical results, see Figures 5.7(a)–5.7(c). Furthermore, the RMSE for the Shepp-Logan filter again coincides with the RMSE of the Hamming filter with $\beta = 0.92$ and of the Gaussian filter with $\beta = 4.9$. Note that this behaviour is more pronounced for the smooth phantom f_{smooth} than for the Shepp-Logan phantom f_{SL}.

In conclusion, our numerical results for $\mathscr{C}^2$-windows totally comply with our L^2-error theory with $k = 2$ and $p = \infty$, although the inevitable discretization errors are not covered.

In our second set of numerical simulations we considered the generalized polynomial filter $A_L(S) = |S| W(S/L)$ with window function

$$W(S) = \begin{cases} 1 - (1-\beta)\,|S|^\nu & \text{for } |S| \leq 1 \\ 0 & \text{for } |S| > 1 \end{cases}$$

of order $\nu \in \mathbb{R}_{>0}$ and with jump height $\beta \in [0, 1)$, see Example 4.3.12. Recall that for this filter our error theory from Corollaries 4.3.9 and 4.3.11 states that, for any $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$, the L^2-norm of the inherent FBP reconstruction error is bounded above by

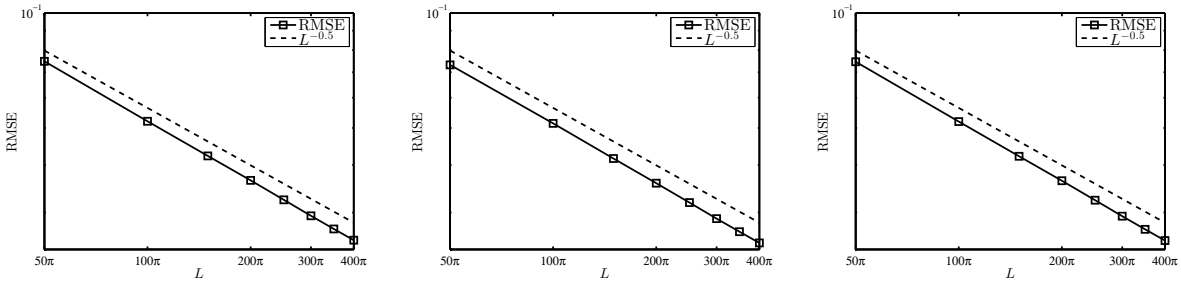
$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \Big(C_{\alpha,\nu,\beta} + 1\Big) \, L^{-\min\{\nu,\alpha\}} \, \|f\|_\alpha, \tag{5.8}$$

where, for fixed α and ν, the constant $C_{\alpha,\nu,\beta} > 0$ decreases with increasing values for $0 \leq \beta < 1$. In particular,

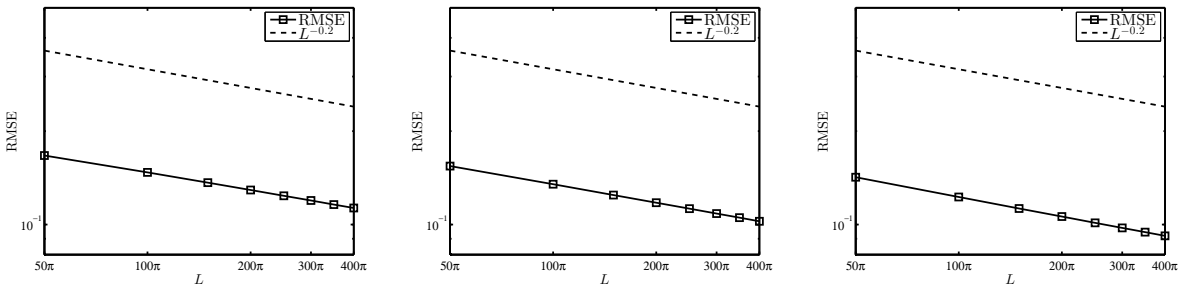
$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}\Big(L^{-\min\{\nu,\alpha\}}\Big) \quad \text{for} \quad L \longrightarrow \infty$$

and the rate of convergence saturates at fractional order $L^{-\nu}$.

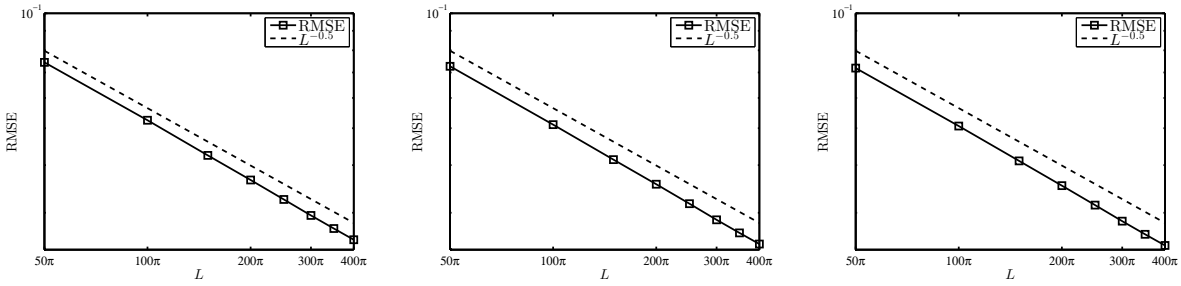
The numerical results for the reconstruction of the Shepp-Logan phantom f_{SL} are displayed in Figures 5.6(d)–5.6(i) and can be summarized as follows. For $\nu = 0.2$, the convergence rate of the RMSE saturates at fractional order $L^{-0.2}$. Moreover, when increasing the parameter β from $\beta = 0$ to $\beta = 0.2$, the RMSE decreases (Figures 5.6(d)–5.6(f)). But for $\nu \in \{0.9, 1.8, 2.7\}$, the RMSE behaves like $L^{-0.5}$, see Figures 5.6(g)–5.6(i), where we always chose the jump height $\beta = 0.8$. Since $f_{\mathrm{SL}} \in \mathrm{H}^\alpha(\mathbb{R}^2)$ for all $\alpha < \frac{1}{2}$, this is exactly the behaviour we expected due to our L^2-error estimate (5.8). In particular, for $\nu \in \{0.9, 1.8, 2.7\}$ the rate of convergence is given by the smoothness of the target function f_{SL}.



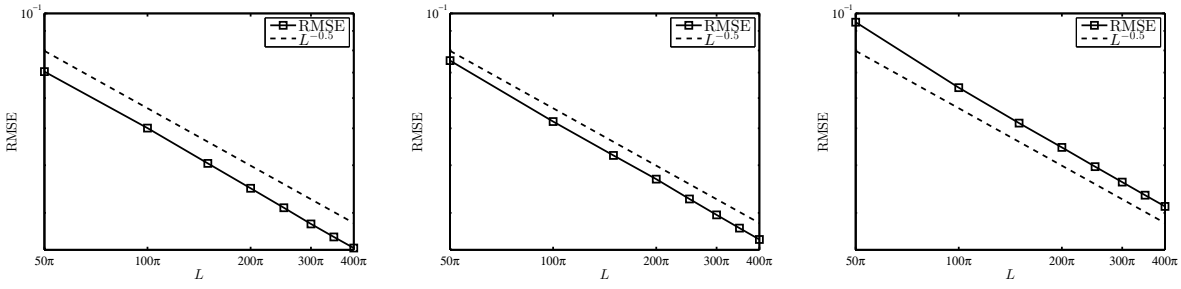
(a) Shepp-Logan (b) Hamming ($\beta = 0.92$) (c) Gaussian ($\beta = 4.9$)



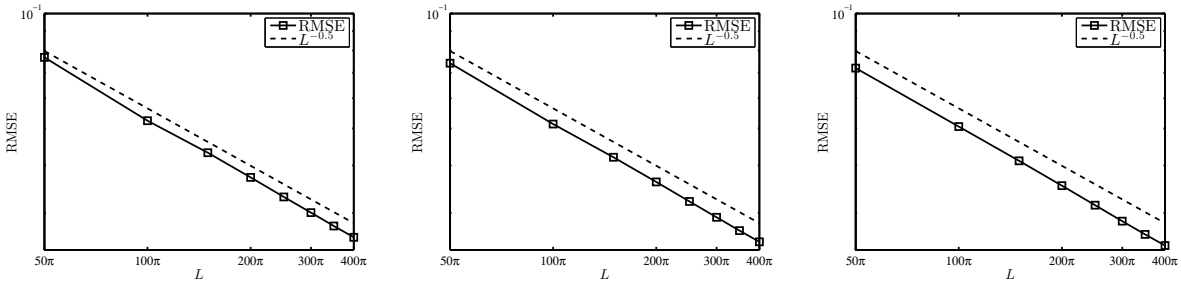
(d) Poly ($\nu = 0.2, \beta = 0$) (e) Poly ($\nu = 0.2, \beta = 0.1$) (f) Poly ($\nu = 0.2, \beta = 0.2$)



(g) Poly ($\nu = 0.9, \beta = 0.8$) (h) Poly ($\nu = 1.8, \beta = 0.8$) (i) Poly ($\nu = 2.7, \beta = 0.8$)

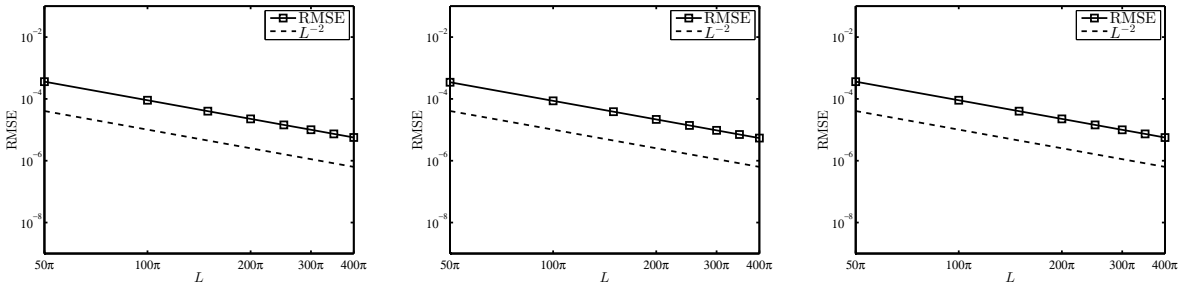


(j) Ram-Lak (k) Ramp ($\beta = 0.7, \lambda = 0$) (l) Ramp ($\beta = 0.1, \lambda = 0$)

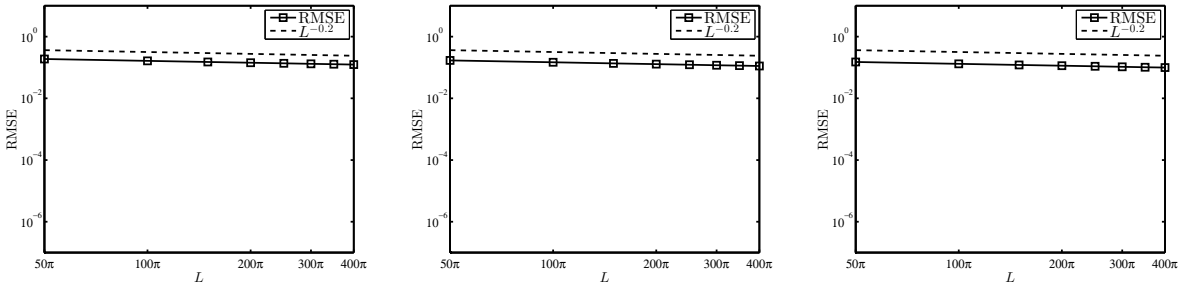


(m) Ramp ($\beta = 0.5, \lambda = 0.25$) (n) Ramp ($\beta = 0.5, \lambda = 0.5$) (o) Ramp ($\beta = 0.5, \lambda = 0.75$)

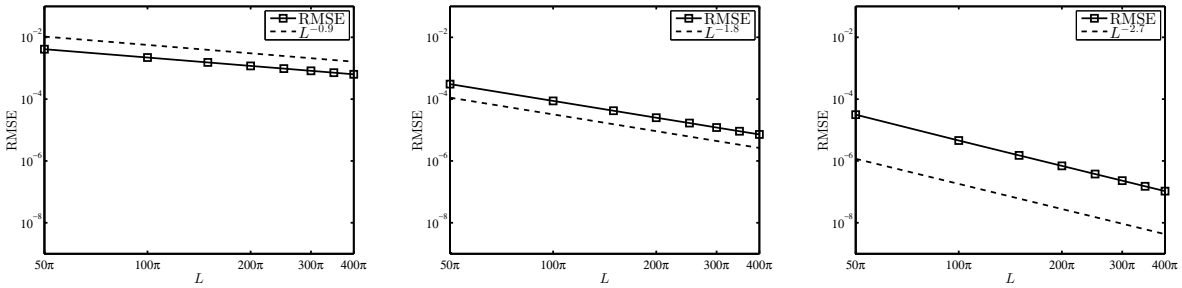
Figure 5.6: Decay rate of the discrete L^2-error for the Shepp-Logan phantom.



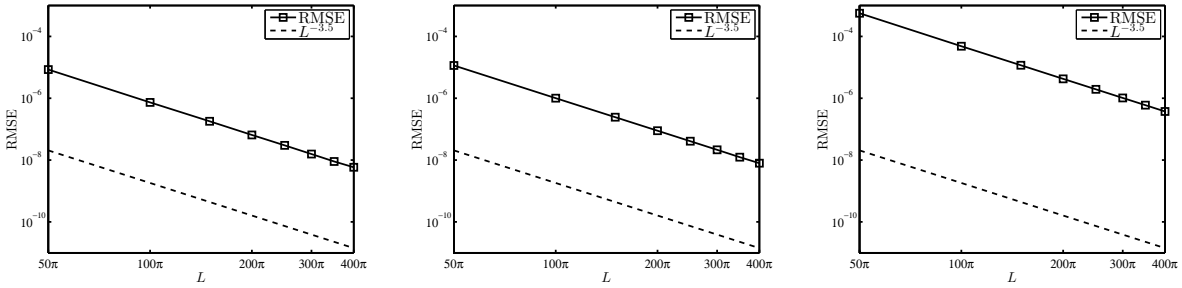
(a) Shepp-Logan (b) Hamming ($\beta = 0.92$) (c) Gaussian ($\beta = 4.9$)



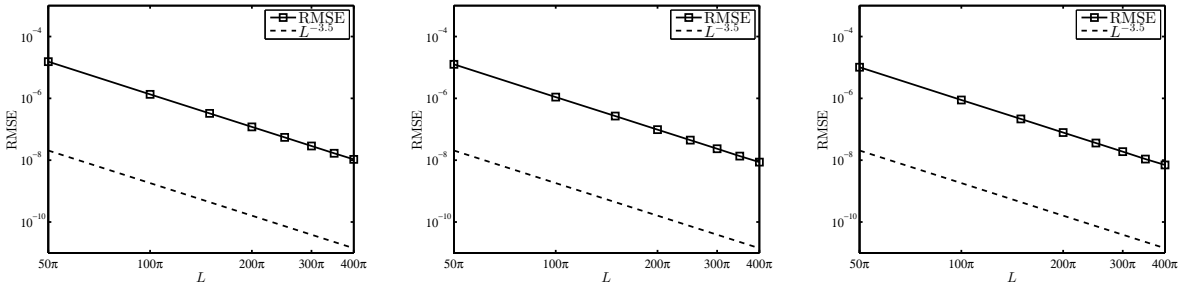
(d) Poly ($\nu = 0.2, \beta = 0$) (e) Poly ($\nu = 0.2, \beta = 0.1$) (f) Poly ($\nu = 0.2, \beta = 0.2$)



(g) Poly ($\nu = 0.9, \beta = 0.8$) (h) Poly ($\nu = 1.8, \beta = 0.8$) (i) Poly ($\nu = 2.7, \beta = 0.8$)



(j) Ram-Lak (k) Ramp ($\beta = 0.7, \lambda = 0$) (l) Ramp ($\beta = 0.1, \lambda = 0$)



(m) Ramp ($\beta = 0.5, \lambda = 0.25$) (n) Ramp ($\beta = 0.5, \lambda = 0.5$) (o) Ramp ($\beta = 0.5, \lambda = 0.75$)

Figure 5.7: Decay rate of the discrete L^2-error for the smooth phantom with $\nu = 3$.

In contrast to that, our numerical results for the FBP reconstruction of the smooth phantom f_{smooth} of order $\nu = 3$ show that the rate of convergence saturates for all our choices of ν, see Figures 5.7(d)–5.7(i). Indeed, for $\nu = 0.2$, the convergence rate of the RMSE again saturates at fractional order $L^{-0.2}$ and, further, increasing the parameter β from $\beta = 0$ to $\beta = 0.2$ decreases the RMSE (Figures 5.7(d)–5.7(f)). But also for $\nu \in \{0.9, 1.8, 2.7\}$, the RMSE behaves like $L^{-\nu}$, see Figures 5.7(g)–5.7(i). However, this was expected, since we have $f_{\text{smooth}} \in \mathrm{H}^\alpha(\mathbb{R}^2)$ for $\alpha < \frac{7}{2}$.

Consequently, our numerical results again totally comply with our L^2-error theory and, especially, the saturation of the convergence rate at fractional order is observable.

In our third and last sequence of numerical simulations we selected the generalized ramp filter $A_L(S) = |S| W(S/L)$ with window function

$$W(S) = \begin{cases} 1 & \text{for } |S| \leq \beta \\ \frac{1-\beta\lambda}{1-\beta} - \frac{1-\lambda}{1-\beta}\,|S| & \text{for } \beta < |S| \leq 1 \\ 0 & \text{for } |S| > 1 \end{cases}$$

of width $\beta \in (0,1)$ and with jump height $\lambda \in [0,1]$, see Example 4.3.2. Recall that choosing $\lambda = 1$ gives the classical Ram-Lak filter. Further, these window functions satisfy Assumption (A) and our error theory from Theorem 4.3.1 states that, for any $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$, the L^2-norm of the inherent FBP reconstruction error is bounded above by

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \left((1-\lambda)\,\beta^{-\alpha} + 1\right) L^{-\alpha}\,\|f\|_\alpha. \tag{5.9}$$

In particular,

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} = \mathcal{O}(L^{-\alpha}) \quad \text{for} \quad L \longrightarrow \infty$$

and the rate of convergence is always determined by the smoothness α of the target function f. Further, for fixed L and f, we see that the L^2-error decreases when increasing the window's width $\beta \in (0,1)$ or jump height $\lambda \in [0,1]$.

In all our numerical simulations for the reconstruction of the Shepp-Logan phantom f_{SL} (Figures 5.6(j)–5.6(o)) we observe that the RMSE behaves like $L^{-0.5}$, as predicted by the theory. Moreover, increasing the width β of the window W leads to an decrease of the RMSE, see Figures 5.6(j)–5.6(l), where we chose the Ram-Lak filter and the ramp filter with $\beta \in \{0.1, 0.7\}$ and $\lambda = 0$. This can also be seen when fixing the width $\beta = 0.5$ and increasing the jump height from $\lambda = 0.25$ to $\lambda = 0.75$ (Figures 5.6(m)–5.6(o)). Consequently, we exactly observe the behaviour predicted by the L^2-error estimate (5.9).

When considering the FBP reconstruction of the smooth phantom f_{smooth} of order $\nu = 3$, we see that for all choices of the parameters β and λ the RMSE behaves like $L^{-3.5}$, as observed in Figures 5.7(j)–5.7(o). Thus, the rate of convergence is determined by the smoothness of the target function and the numerical observations comply with our L^2-error estimate (5.9). Further, as for the Shepp-Logan phantom f_{SL}, increasing the width β results in a decrease of the RMSE (Figures 5.7(j)–5.7(l)). The same holds true when fixing β and increasing the jump height λ, see Figures 5.7(m)–5.7(o).

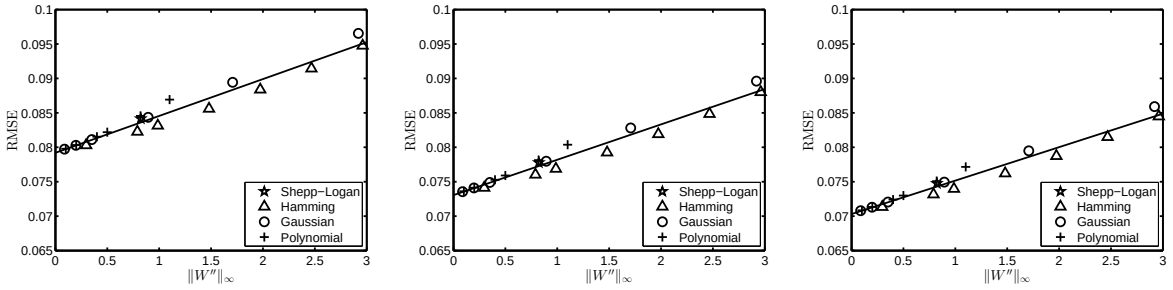
In conclusion, our numerical results totally comply with our L^2-error theory. In particular, we observe that the decay rate of the L^2-error is indeed determined by the smoothness α of the target function f if Assumption (A) is satisfied. Finally, we remark that for the two phantoms the RMSE is minimal for the Ram-Lak filter. This also complies with our error theory, since in our L^2-error estimate

$$\|f - f_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \left(\Phi_{\alpha,W}^{1/2}(L) + L^{-\alpha}\right) \|f\|_\alpha$$

from Theorem 4.2.3 (with $\sigma = 0$) the error term

$$\Phi_{\alpha,W}(L) = \operatorname*{ess\,sup}_{S \in [-1,1]} \frac{(1 - W(S))^2}{(1 + L^2 S^2)^\alpha} \quad \text{for } L > 0$$

only vanishes if $W \equiv 1$ on $[-1,1]$, i.e., when choosing the Ram-Lak filter.



(a) $L = 40\pi$ (b) $L = 46\pi$ (c) $L = 50\pi$

Figure 5.8: Affine-linear behaviour of the discrete L^2-error with respect to the quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$ for the Shepp-Logan phantom and different bandwidths $L > 0$.

5.3.2 Validation of the affine-linear behaviour

To close this chapter on numerical experiments we finally validate the affine-linear behaviour of the inherent FBP reconstruction error $e_L = f - f_L$ with respect to different quality indicators for the low-pass filter's window W. Summarizing Corollaries 4.3.9 and 4.3.11, we have seen in (5.6) that for target functions $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $\alpha > 0$ the L^2-norm of e_L is bounded above by

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \Big(c_{\alpha,k,p} \, \|W^{(k)}\|_{\mathrm{L}^p([0,1])} + 1 \Big) \, L^{-\min\{k-1/p,\alpha\}} \, \|f\|_\alpha,$$

provided that the window W satisfies $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \leq j \leq k-1$ with $k \in \mathbb{N}$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ with $1 < p \leq \infty$ such that

$$W(0) = 1 \quad \text{and} \quad W^{(j)}(0) = 0 \quad \forall\, 1 \leq j \leq k-1.$$

In particular, we observe that, for fixed target function f and bandwidth L, the performance of the utilized low-pass filter is governed by the L^p-norm of its window's k-th derivative $W^{(k)}$ on the interval $[0,1]$. Moreover, our theory predicts that the L^2-norm of the error e_L is affine-linear with respect to the quality indicator $\|W^{(k)}\|_{\mathrm{L}^p([0,1])}$.

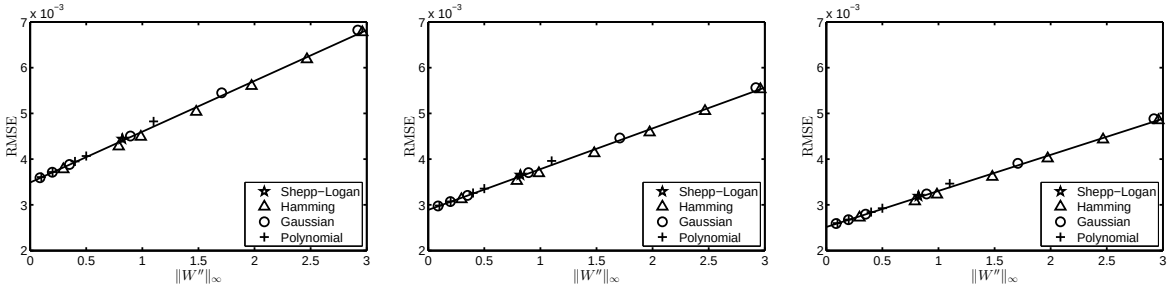
In our first numerical experiments we again consider the reconstruction of the Shepp-Logan phantom. In addition to the classical Shepp-Logan filter, we also applied the Hamming filter for different choices of the parameter $\beta \in [\frac{1}{2}, 1]$, namely for $\beta \in \{0.7, 0.75, 0.8, 0.85, 0.9, 0.92, 0.97\}$, and the Gaussian filter for various parameters $\beta > 1$, here for $\beta \in \{2.6, 3.4, 4.7, 7.5, 10, 15\}$. Moreover, we chose the generalized polynomial filter of order $\nu = 2$ from Example 4.3.12 for different jump heights $\beta \in [0,1)$, more precisely for $\beta \in \{0.45, 0.59, 0.75, 0.8, 0.9, 0.95\}$.

Note that the corresponding window functions satisfy the assumptions of our theory for $k = 2$ and $p = \infty$ so that the FBP reconstruction error e_L is predicted to behave affine-linearly with respect to the quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$, i.e., the L^∞-norm of the window's second derivative W''. The values of $\|W''\|_{\mathrm{L}^\infty([0,1])}$ for the filters mentioned above are listed in Table 5.2.

Figure 5.8 now shows the RMSE of the FBP reconstruction of the Shepp-Logan phantom as a function of the quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$ for different choices of the bandwidth $L > 0$.

Name	$W(S)$ for $\lvert S\rvert \leq 1$	Parameter	$\|W''\|_{\mathrm{L}^\infty([0,1])}$
Shepp-Logan	$\operatorname{sinc}(\pi S/2)$	-	$\pi^2/12$
Hamming	$\beta + (1-\beta)\cos(\pi S)$	$\beta \in [1/2, 1]$	$(1-\beta)\,\pi^2$
Gaussian	$\exp(-(\pi S/\beta)^2)$	$\beta > 1$	$2\pi^2/\beta^2$
Polynomial	$1 - (1-\beta)\,S^2$	$\beta \in [0,1)$	$2(1-\beta)$

Table 5.2: Quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$ for different low-pass filters.



(a) $L = 40\pi$ (b) $L = 46\pi$ (c) $L = 50\pi$

Figure 5.9: Affine-linear behaviour of the discrete L^2-error with respect to the quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$ for the smooth phantom of order $\nu = 1$ and different bandwidths $L > 0$.

Firstly, we observe an increasing RMSE when increasing the quality indicator $\|W''\|_{\mathrm{L}^\infty([0,1])}$ in all of our numerical experiments. This is exactly the behaviour we expected due to the first term in our L^2-error estimate, which is in this case given as

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \left(\frac{1}{2}\|W''\|_{\mathrm{L}^\infty([0,1])} + 1\right) L^{-\alpha}\|f\|_\alpha,$$

since the attenuation function f_{SL} of the Shepp-Logan phantom satisfies $f_{\mathrm{SL}} \in \mathrm{H}^\alpha(\mathbb{R}^2)$ for all $\alpha < \frac{1}{2}$ so that, in particular, we are in the case $\alpha < k$. Moreover, the predicted affine-linear behaviour of the RMSE with respect to $\|W''\|_{\mathrm{L}^\infty([0,1])}$ is clearly visible.

Secondly, we see that the RMSE decreases at increasing bandwidth L. This behaviour complies with the second term in our L^2-error estimate,

$$L^{-\alpha}\|f\|_\alpha \quad \text{for } \alpha > 0.$$

In more details, comparing the numerical results for the bandwidths $L = 40\pi$ (Figure 5.8(a)), $L = 46\pi$ (Figure5.8(b)) and $L = 50\pi$ (Figure 5.8(c)) we see that the RMSE decreases for all filters, as predicted by our error estimate. Moreover, the affine-linear behaviour of the error e_L with respect to $\|W''\|_{\mathrm{L}^\infty([0,1])}$ is more distinct for a larger bandwidth.

In our second sequence of numerical experiments we utilize the same low-pass filters as well as linear interpolation for reconstructing the smooth phantom of order $\nu = 1$, whose attenuation function f_{smooth} belongs to $\mathrm{H}^\alpha(\mathbb{R}^2)$ for all $\alpha < \frac{3}{2}$ so that we are again in the case $\alpha < k$.

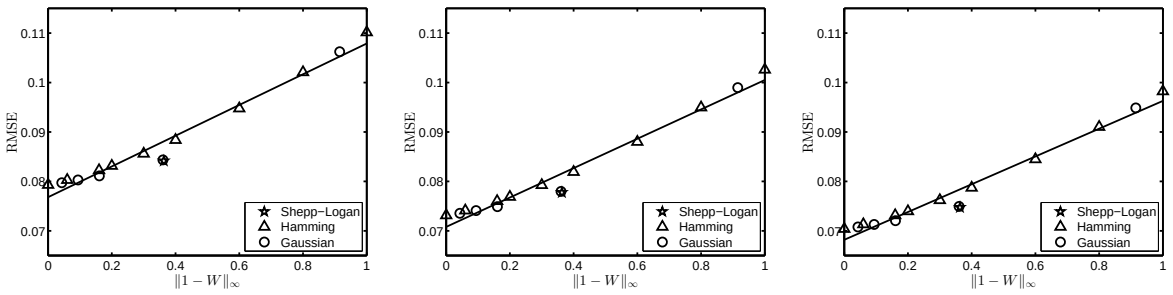
Summarizing our results, we observe exactly the same behaviour of the error as described before for the Shepp-Logan phantom, see Figure 5.9. However, the affine-linear behaviour of the discrete L^2-error is more pronounced when reconstructing the smoother phantom.

Recall that the window functions of the Shepp-Logan, Hamming and Gaussian filters satisfy Assumption (A) as long as the smoothness $\alpha > 0$ of the target function f is smaller than 2. This has been observed in our numerical experiments, as reported exemplary for the Shepp-Logan filter in Section 4.3. Therefore, our error theory from Theorem 4.3.1 states that, for any function $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{H}^\alpha(\mathbb{R}^2)$ with $0 < \alpha < 2$, the L^2-norm of the error e_L is bounded above by

$$\|e_L\|_{\mathrm{L}^2(\mathbb{R}^2)} \leq \left(c_{\alpha,W}\|1 - W\|_{\mathrm{L}^\infty([0,1])} + 1\right) L^{-\alpha}\|f\|_\alpha.$$

Name	$W(S)$ for $\lvert S\rvert \leq 1$	Parameter	$\|1 - W\|_{\mathrm{L}^\infty([0,1])}$
Shepp-Logan	$\mathrm{sinc}(\pi S/2)$	-	$1 - 2/\pi$
Hamming	$\beta + (1-\beta)\cos(\pi S)$	$\beta \in [1/2, 1]$	$2(1-\beta)$
Gaussian	$\exp(-(\pi S/\beta)^2)$	$\beta > 1$	$1 - \exp(-\pi^2/\beta^2)$

Table 5.3: Quality indicator $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ for different low-pass filters.



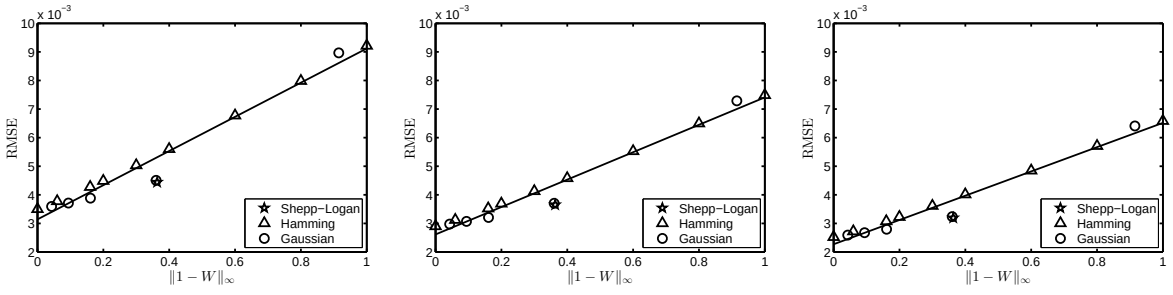
(a) $L = 40\pi$ (b) $L = 46\pi$ (c) $L = 50\pi$

Figure 5.10: Affine-linear behaviour of the discrete L^2-error with respect to the quality indicator $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ for the Shepp-Logan phantom and different bandwidths $L > 0$.

Note that the smoothness α of both the Shepp-Logan phantom and the smooth phantom of order $\nu = 1$ satisfy the assumption $\alpha < 2$. Thus, in both cases our theory now predicts that the L^2-norm of the error e_L is affine-linear with respect to the quality indicator $\|1-W\|_{\mathrm{L}^\infty([0,1])}$. To validate this behaviour numerically, we applied the Shepp-Logan filter, the Hamming filter with parameters $\beta \in \{0.5, 0.6, 0.7, 0.8, 0.85, 0.9, 0.92, 0.97, 1\}$ and the Gaussian filter with parameters $\beta \in \{2, 4.7, 7.5, 10, 15\}$. The corresponding values for $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ are listed in Table 5.3.

Figure 5.10 shows the RMSE of the FBP reconstruction of the Shepp-Logan phantom as a function of the quality indicator $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ for different choices of the bandwidth $L > 0$. As predicted, for each choice of $L \in \{40\pi, 46\pi, 50\pi\}$ the discrete L^2-norm of the error e_L behaves affine-linearly with respect to $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$. Further, increasing L decreases the RMSE and the affine-linear behaviour becomes more distinct. The same observations remain valid when considering the FBP reconstruction of the smooth phantom of order $\nu = 1$, see Figure 5.11. However, as reported before, the affine-linear behaviour of the reconstruction error is more pronounced for the smoother phantom.

In conclusion, our numerical results for both the FBP reconstruction of the Shepp-Logan phantom and the smooth phantom of order $\nu = 1$ totally comply with our L^2-error theory. On the one hand, under Assumption (A) we observe an affine-linear behaviour of the L^2-error with respect to $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ as predicted by the estimate in Theorem 4.3.1. On the other hand, for window functions W with vanishing derivatives at the origin up to the order $k-1$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ we observe that the L^2-error is affine-linear with respect to $\|W^{(k)}\|_{\mathrm{L}^p([0,1])}$, exactly as in the estimates in Corollaries 4.3.9 and 4.3.11. Here, we reported the numerical results only for two different phantoms and the case $k = 2$ and $p = \infty$. However, the affine-linear behaviour is also observable for other test functions and values of $k \in \mathbb{N}$ and $1 < p \leq \infty$.



(a) $L = 40\pi$ (b) $L = 46\pi$ (c) $L = 50\pi$

Figure 5.11: Affine-linear behaviour of the discrete L^2-error with respect to the quality indicator $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ for the smooth phantom of order $\nu = 1$ and different bandwidths $L > 0$.

Chapter 6

Summary and outlook

This thesis was concerned with the approximation of bivariate functions from the knowledge of their Radon data by utilizing the filtered back projection method from computerized tomography. The focus was on the analysis of the inherent FBP reconstruction error being incurred by the application of a low-pass filter with finite bandwidth and compactly supported window function.

Under suitable assumptions, the reconstruction of a bivariate target function $f \equiv f(x,y)$ from given Radon data

$$\{\mathcal{R}f(t,\theta) \mid t \in \mathbb{R},\, \theta \in [0,\pi)\}$$

can be achieved by applying the classical *filtered back projection formula*, which states that

$$f(x,y) = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}[|S|\mathcal{F}(\mathcal{R}f)(S,\theta)])(x,y) \quad \forall\,(x,y) \in \mathbb{R}^2.$$

Unfortunately, the FBP formula is highly sensitive with respect to noise and, thus, cannot be used in practice. A standard approach for stabilizing the FBP formula consists in incorporating a low-pass filter

$$A_L(S) = |S|\, W(S/L) \quad \text{ for } S \in \mathbb{R}$$

with finite bandwidth $L > 0$ and an even window function $W \in \mathrm{L}^\infty(\mathbb{R})$ with $\operatorname{supp}(W) \subseteq [-1,1]$. This reduces the noise sensitivity, but only leads to an approximate FBP method.

In this work, we rigorously defined the approximate FBP reconstruction f_L and showed that $f_L \in \mathrm{L}^2(\mathbb{R}^2)$ can be expressed in standard form as

$$f_L = \frac{1}{2}\,\mathcal{B}(\mathcal{F}^{-1}A_L * \mathcal{R}f) = f * K_L,$$

where the target functions f is only required to satisfy $f \in \mathrm{L}^1(\mathbb{R}^2) \cap \mathrm{L}^2(\mathbb{R}^2)$ and the convolution kernel K_L is not compactly supported and not necessarily integrable on $\mathbb{R}^2$.

In the main part of this thesis we have then considered the inherent FBP reconstruction error $e_L = f - f_L$ which is incurred by the application of the low-pass filter A_L. More precisely, for target functions f from Sobolev spaces $\mathrm{H}^\alpha(\mathbb{R}^2)$ of fractional order $\alpha > 0$ we have estimated the H^σ-norm of e_L for all $0 \le \sigma \le \alpha$ subject to the filter's bandwidth L and window function W. In particular, we have proven convergence rates for the approximate FBP reconstruction f_L as the bandwidth L goes to infinity, where we have observed saturation of the order of convergence at fractional rates depending on smoothness properties of the window W. For that purpose, we have considered various scenarios. Here, we wish to highlight two special cases.

If the window function W satisfies $W^{(j)} \in \mathcal{AC}([-1,1])$ for all $0 \le j \le k-1$ with $k \in \mathbb{N}$ and $W^{(k)} \in \mathrm{L}^p([-1,1])$ for some $1 < p \le \infty$ such that

$$W(0) = 1 \quad \text{ and } \quad W^{(j)}(0) = 0 \quad \forall\, 1 \le j \le k-1,$$

we have proven that the H^σ-norm of e_L is bounded above by

$$\|e_L\|_\sigma \le \left(C_{\alpha-\sigma,k,p}\,\|W^{(k)}\|_{\mathrm{L}^p([0,1])}\,L^{-\min\{k-1/p,\alpha-\sigma\}} + L^{\sigma-\alpha}\right)\|f\|_\alpha. \tag{6.1}$$

Thus, under these assumptions the order of convergence is predicted to saturate at fractional rate $L^{-(k-1/p)}$. Further, the error estimate indicates that, for fixed function f, the performance of the low-pass filter is governed by the L^p-norm of its window's k-th derivative. On top of that, the estimate predicts an affine-linear behaviour of the H^σ-norm of e_L with respect to $\|W^{(k)}\|_{\mathrm{L}^p([0,1])}$ so that this quantity acts as a quality indicator for the utilized low-pass filter.

On the other hand, if the window W satisfies Assumption (A), e.g., if, for some $0 < \varepsilon < 1$,

$$W(S) = 1 \quad \forall\, S \in (-\varepsilon, \varepsilon),$$

the H^σ-norm of e_L is proven to be bounded above by

$$\|e_L\|_\sigma \leq \Big(C_{\alpha-\sigma,W}\, \|1 - W\|_{\mathrm{L}^\infty([0,1])} + 1\Big)\, L^{\sigma-\alpha}\, \|f\|_\alpha. \tag{6.2}$$

Therefore, in this case the order of convergence is always determined by the difference between the smoothness α of the target function f and the order σ of the Sobolev norm in which the reconstruction error is measured. In addition, the error estimate now predicts an affine-linear behaviour of the H^σ-norm of e_L with respect to the quality indicator $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$ so that this quantity can be used to evaluate the approximation quality of the chosen low-pass filter.

From both our results we conclude that the *flatness* of the low-pass filter's window function at the origin determines the convergence rate of the H^σ-error bounds for the inherent FBP reconstruction error $e_L = f - f_L$. Moreover, the estimates provided for the approximation error can be combined with estimates for the data error to obtain convergence rates for noisy data.

Finally, we described a standard discretization of the FBP method and provided numerical experiments for the L^2-case, where we applied the standard root mean square error and considered two phantoms of different smoothness. In our first set of numerical experiments we investigated the convergence rate of the RMSE for various low-pass filters. Here, we observed exactly the decay rates predicted by our estimates (6.1) and (6.2). In particular, the saturation of the order of convergence at *fractional* rates was observable. In our second set of numerical experiments we analysed the behaviour of the RMSE with respect to different quality indicators. As predicted by (6.1) with $k = 2$ and $p = \infty$, we clearly observed an affine-linear behaviour of the error with respect to $\|W''\|_{\mathrm{L}^\infty([0,1])}$. In addition, under Assumption (A) we saw that the RMSE behaves affine-linearly with respect to $\|1 - W\|_{\mathrm{L}^\infty([0,1])}$, as prognosticated in (6.2). All in all, our numerical results totally comply with our L^2-theory and validate our theoretical results.

This thesis leaves room for further investigations and improvements. First of all, in practical applications, only finitely many Radon samples are given. Therefore, the approximate FBP reconstruction formula has to be discretized in order to be applicable. This introduces inevitable discretization error that are not covered by our error theory, since we consider the continuous setting. In future considerations, one could include these discretization errors, for which a rich literature is available. Moreover, in practice, the Radon data is not only discrete, but also corrupted with noise. In this thesis, we briefly discussed the noisy case, where we analysed the worst case overall reconstruction error from noisy data. For future works, it would be interesting to improve the derived error estimates when a concrete distribution of the noise is known. In particular, the question for an optimal low-pass filter for a given noise distribution is of interest.

Secondly, by deriving asymptotic error estimates, we observed that assuming differentiability of the filter's window only at zero allows us to derive saturation rates for the FBP reconstruction. Indeed, if W is k-times differentiable at 0 such that its first $k-1$ derivatives vanish at 0, the convergence order saturates at rate L^{-k} and the quantity $|W^{(k)}(0)|$ dominates the error bound. For the future, it would be interesting to generalize these results so that also fractional saturation rates can be predicted, while not assuming regularity of the window on its entire support.

Finally, it might be worthwhile to generalize our results to other notions of smoothness for the target function. Instead of fractional Sobolev space, one could consider Besov spaces or even more general smoothness concepts, as common in regularization theory of ill-posed problems.

Appendix A

Mathematical tools

In this chapter we summarize some basic facts about the Fourier transform and introduce Sobolev spaces of fractional order. To this end, we also explain the notion of (tempered) distributions and define the distributional Fourier transform. Finally, we list some general mathematical tools that are used throughout this work. For a comprehensive treatment of Fourier analysis and distribution theory we refer the reader to the monographs [27], [46], [108], [119]. Further details on Sobolev spaces can be found in the textbooks [2], [16].

A.1 Fourier transform and convolution

The Fourier transform is a basic tool in the mathematics of computerized tomography and is used extensively in this work. In this section we define the n-dimensional Fourier transform and collect some important properties. Further, we introduce the convolution product and describe its interplay with the Fourier transform.

The Fourier transform

We start with the definition of the Fourier transform on the space $\mathrm{L}^1(\mathbb{R}^n)$ of integrable functions.

Definition A.1.1 (Fourier transform)**.** *The* Fourier transform $\mathcal{F}f$ *of a function* $f \in \mathrm{L}^1(\mathbb{R}^n)$ *is defined as*

$$\mathcal{F}f(\omega) = \int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-\mathrm{i}x^T\omega}\, \mathrm{d}x \quad \text{for } \omega \in \mathbb{R}^n. \tag{A.1}$$

We remark that the Fourier transform $\mathcal{F}f$ of a function $f \in \mathrm{L}^1(\mathbb{R}^n)$ is well-defined on $\mathbb{R}^n$. The first important observation is that in this case the Fourier transform $\mathcal{F}f$ is even continuous and, in particular, its point evaluation makes sense.

Lemma A.1.2 (Riemann-Lebesgue lemma)**.** *For* $f \in \mathrm{L}^1(\mathbb{R}^n)$*, its Fourier transform* $\mathcal{F}f$ *is uniformly continuous on* $\mathbb{R}^n$ *and satisfies*

$$|\mathcal{F}f(\omega)| \longrightarrow 0 \quad \text{for} \quad \|\omega\|_{\mathbb{R}^n} \longrightarrow \infty.$$

Proof. See, for example, [119, Theorem I.1.1] and [119, Theorem I.1.2]. □

Let $\mathscr{C}_0(\mathbb{R}^n)$ denote the space of continuous functions vanishing at infinity, i.e.,

$$\mathscr{C}_0(\mathbb{R}^n) = \{f \in \mathscr{C}(\mathbb{R}^n) \mid f(x) \longrightarrow 0 \text{ for } \|x\|_{\mathbb{R}^n} \longrightarrow \infty\},$$

which is equipped with the norm $\|\cdot\|_\infty$ given by

$$\|f\|_\infty = \sup_{x\in\mathbb{R}^n} |f(x)| \quad \text{for } f \in \mathscr{C}_0(\mathbb{R}^n).$$

Then, we have the following continuity result for the Fourier transform.

Theorem A.1.3. *The Fourier transform $\mathcal{F} : \mathrm{L}^1(\mathbb{R}^n) \longrightarrow \mathscr{C}_0(\mathbb{R}^n)$ is a continuous linear operator with norm $\|\mathcal{F}\| \leq 1$, i.e.,*

$$\|\mathcal{F}f\|_\infty \leq \|f\|_{\mathrm{L}^1(\mathbb{R}^n)} \quad \forall\, f \in \mathrm{L}^1(\mathbb{R}^n).$$

Proof. See, for example, [124, Theorem V.2.2]. □

We now define the inverse Fourier transform on $\mathrm{L}^1(\mathbb{R}^n)$.

Definition A.1.4 (Inverse Fourier transform)**.** *For $f \in \mathrm{L}^1(\mathbb{R}^n)$, the* inverse Fourier transform *$\mathcal{F}^{-1}f$ is defined as*

$$\mathcal{F}^{-1}f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} f(\omega)\, \mathrm{e}^{\mathrm{i}x^T\omega}\, \mathrm{d}\omega \quad \text{for } x \in \mathbb{R}^n. \tag{A.2}$$

Although, by Lemma A.1.2, the Fourier transform $\mathcal{F}f$ of $f \in \mathrm{L}^1(\mathbb{R}^n)$ vanishes at infinity, this does not necessarily imply that $\mathcal{F}f \in \mathrm{L}^1(\mathbb{R}^n)$. Thus, in order to apply the inverse Fourier transform to $\mathcal{F}f$, we have to assume that $\mathcal{F}f \in \mathrm{L}^1(\mathbb{R}^n)$ is satisfied. In this case we indeed get the following inverse relationship.

Theorem A.1.5 (Fourier inversion)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ with $\mathcal{F}f \in \mathrm{L}^1(\mathbb{R}^n)$. Then, the identity*

$$\mathcal{F}^{-1}(\mathcal{F}f)(x) = f(x) = \mathcal{F}(\mathcal{F}^{-1}f)(x)$$

holds for almost all $x \in \mathbb{R}^n$ with equality in every continuity point of f.

Proof. See, for example, [119, Corollary I.1.21]. □

As a corollary we get the injectivity of the Fourier transform on $\mathrm{L}^1(\mathbb{R}^n)$.

Corollary A.1.6 (Injectivity of $\mathcal{F}$)**.** *For $f \in \mathrm{L}^1(\mathbb{R}^n)$ we have*

$$\mathcal{F}f = 0 \quad \Longrightarrow \quad f = 0,$$

i.e., the Fourier transform $\mathcal{F}$ is injective on $\mathrm{L}^1(\mathbb{R}^n)$. □

We have the following variant of the Fourier inversion theorem.

Theorem A.1.7 (Fourier inversion formula)**.** *For $f \in \mathrm{L}^1(\mathbb{R}^n)$, the* Fourier inversion formula

$$\mathcal{F}^{-1}(\mathcal{F}f)(x) = f(x) = \mathcal{F}(\mathcal{F}^{-1}f)(x) \tag{A.3}$$

holds for almost all $x \in \mathbb{R}^n$, where the (inverse) Fourier transforms have to be defined as

$$\mathcal{F}f(x) = \lim_{\epsilon \to 0} \int_{\mathbb{R}^n} f(\omega)\, \mathrm{e}^{-\mathrm{i}x^T\omega}\, \mathrm{e}^{-\epsilon\|\omega\|_{\mathbb{R}^n}^2}\, \mathrm{d}\omega \quad \text{for } x \in \mathbb{R}^n \tag{A.4}$$

and

$$\mathcal{F}^{-1}f(\omega) = (2\pi)^{-n} \lim_{\epsilon \to 0} \int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{\mathrm{i}x^T\omega}\, \mathrm{e}^{-\epsilon\|x\|_{\mathbb{R}^n}^2}\, \mathrm{d}x \quad \text{for } \omega \in \mathbb{R}^n. \tag{A.5}$$

Further, for $f \in \mathrm{L}^1(\mathbb{R}^n) \cap \mathscr{C}(\mathbb{R}^n)$ the identity (A.3) *holds for all $x \in \mathbb{R}^n$.*

Proof. See, for example, [34, Chapter 7.5]. □

Remark A.1.8. *For $f \in \mathrm{L}^1(\mathbb{R}^n)$ the definitions* (A.1) *and* (A.4) *as well as* (A.2) *and* (A.5) *coincide due to Lebesgue's theorem on dominated convergence.*

We remark that the inverse Fourier transform $\mathcal{F}^{-1}$ can be expressed in terms of the Fourier transform $\mathcal{F}$ and the parity operator $^* : \mathrm{L}^p(\mathbb{R}^n) \longrightarrow \mathrm{L}^p(\mathbb{R}^n)$ for $1 \leq p \leq \infty$, which is defined as

$$f^*(x) = f(-x) \quad \text{for } x \in \mathbb{R}^n. \tag{A.6}$$

Remark A.1.9. *For $f \in \mathrm{L}^1(\mathbb{R}^n)$ we have $\mathcal{F}^{-1} f = (2\pi)^{-n}\, \mathcal{F} f^* = (2\pi)^{-n}\, (\mathcal{F} f)^*$.*

We now list some basic properties of the Fourier transform.

Proposition A.1.10. *For $f \in \mathrm{L}^1(\mathbb{R}^n)$ the following properties hold true.*

(i) Translation: *For $y \in \mathbb{R}^n$ we consider the function*

$$g(x) = f(x - y) \quad \text{for } x \in \mathbb{R}^n.$$

Then,

$$\mathcal{F} g(x) = \mathrm{e}^{-\mathrm{i} x^T y}\, \mathcal{F} f(x) \quad \forall\, x \in \mathbb{R}^n.$$

(ii) Scaling: *For $a > 0$ we consider the function*

$$g(x) = f(ax) \quad \text{for } x \in \mathbb{R}^n.$$

Then,

$$\mathcal{F} g(x) = a^{-n}\, \mathcal{F} f(a^{-1} x) \quad \forall\, x \in \mathbb{R}^n.$$

(iii) Modulation: *For $y \in \mathbb{R}^n$ we consider the function*

$$g(x) = \mathrm{e}^{\mathrm{i} x^T y}\, f(x) \quad \text{for } x \in \mathbb{R}^n.$$

Then,

$$\mathcal{F} g(x) = \mathcal{F} f(x - y) \quad \forall\, x \in \mathbb{R}^n.$$

(iv) Let $\alpha \in \mathbb{N}_0^n$ be a multi-index and $\mathrm{D}^\alpha = \frac{\partial^\alpha}{\partial x^\alpha}$. If $\mathrm{D}^\alpha f$ exists and is in $\mathrm{L}^1(\mathbb{R}^n)$, then

$$\mathcal{F}(\mathrm{D}^\alpha f) = \mathrm{i}^{|\alpha|}\, x^\alpha\, \mathcal{F} f,$$

whereas, if $x^\alpha f$ is integrable on $\mathbb{R}^n$, then

$$\mathcal{F}(x^\alpha f) = \mathrm{i}^{|\alpha|}\, \mathrm{D}^\alpha (\mathcal{F} f).$$

Proof. See, for example, [34, Theorem 7.8]. □

Another important property of the Fourier transform is *Parseval's identity.*

Theorem A.1.11 (Parseval's identity)**.** *For $f, g \in \mathrm{L}^1(\mathbb{R}^n)$ we have*

$$\int_{\mathbb{R}^n} \mathcal{F} f(x)\, g(x)\, \mathrm{d}x = \int_{\mathbb{R}^n} f(x)\, \mathcal{F} g(x)\, \mathrm{d}x.$$

Proof. See, for example, [119, Theorem I.1.15]. □

The next theorem is the classical Rayleigh-Plancherel theorem, which shows that the Fourier transform preserves the L^2-norm up to a multiplicative constant.

Theorem A.1.12 (Rayleigh-Plancherel)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $f \in \mathrm{L}^2(\mathbb{R}^n)$ or $\mathcal{F} f \in \mathrm{L}^2(\mathbb{R}^n)$. Then, we have $\mathcal{F} f \in \mathrm{L}^2(\mathbb{R}^n)$ or $f \in \mathrm{L}^2(\mathbb{R}^n)$, respectively, and*

$$\|f\|_{\mathrm{L}^2(\mathbb{R}^n)} = (2\pi)^{-n/2}\, \|\mathcal{F} f\|_{\mathrm{L}^2(\mathbb{R}^n)}.$$

More generally, for $f, g \in \mathrm{L}^1(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n)$ we have

$$(f, g)_{\mathrm{L}^2(\mathbb{R}^n)} = (2\pi)^{-n}\, (\mathcal{F} f, \mathcal{F} g)_{\mathrm{L}^2(\mathbb{R}^n)}.$$

Proof. See, for example, [119, Theorem I.2.1]. □

Since $\mathrm{L}^1(\mathbb{R}^n) \cap \mathrm{L}^2(\mathbb{R}^n) \subset \mathrm{L}^2(\mathbb{R}^n)$ is dense, the Rayleigh-Plancherel Theorem A.1.12 shows that the Fourier transform can be continuously extended to an operator

$$\mathcal{F} : \mathrm{L}^2(\mathbb{R}^n) \longrightarrow \mathrm{L}^2(\mathbb{R}^n),$$

which is an isometry up to a multiplicative constant. Further, the extended operator $\mathcal{F}$ is bijective on $\mathrm{L}^2(\mathbb{R}^n)$ and its inverse $\mathcal{F}^{-1}$ is the continuous extension of the inverse Fourier transform. In this work, however, we will not distinguish between the regular Fourier transform and its extension.

Consequently, the Fourier transform and its inverse are now defined on the whole of $\mathrm{L}^2(\mathbb{R}^n)$. But for $f \in \mathrm{L}^2(\mathbb{R}^n)$, the point evaluation of $\mathcal{F}f$ makes sense only almost everywhere and the Fourier inversion formula holds in L^2-sense.

Corollary A.1.13 (Fourier inversion in $\mathrm{L}^2(\mathbb{R}^n)$)**.** *For $f \in \mathrm{L}^2(\mathbb{R}^n)$ the Fourier inversion formula*

$$\mathcal{F}^{-1}(\mathcal{F}f) = f = \mathcal{F}(\mathcal{F}^{-1}f)$$

holds in L^2-sense and, in particular, almost everywhere on $\mathbb{R}^n$. □

We close this paragraph on the Fourier transform with a variant of the classical *Paley-Wiener theorem*, which characterizes the Fourier transform of compactly supported functions.

Theorem A.1.14 (Paley-Wiener)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^n) \setminus \{0\}$ be compactly supported. Then, its Fourier transform $\mathcal{F}f$ is analytic and cannot have compact support.*

Proof. See, for example, [108, Lemma 7.21] and [108, Theorem 7.23]. □

The convolution product

We now define the convolution product of functions in $\mathrm{L}^1(\mathbb{R}^n)$ and investigate its interaction with the Fourier transform.

Definition A.1.15 (Convolution product)**.** *The* convolution product $f * g$ *of two functions $f, g \in \mathrm{L}^1(\mathbb{R}^n)$ is defined as*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)\, g(y)\, \mathrm{d}y \quad \text{for } x \in \mathbb{R}^n.$$

We remark that the convolution product of $f, g \in \mathrm{L}^1(\mathbb{R}^n)$ exists and is again in $\mathrm{L}^1(\mathbb{R}^n)$. More generally, we have the following result.

Theorem A.1.16 (Young's inequality)**.** *Let $f \in \mathrm{L}^p(\mathbb{R}^n)$ and $g \in \mathrm{L}^q(\mathbb{R}^n)$ with $1 \leq p, q \leq \infty$. Then, we have $f * g \in \mathrm{L}^r(\mathbb{R}^n)$ with $1 \leq r \leq \infty$ satisfying*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$$

and Young's inequality

$$\|f * g\|_{\mathrm{L}^r(\mathbb{R}^n)} \leq \|f\|_{\mathrm{L}^p(\mathbb{R}^n)}\, \|g\|_{\mathrm{L}^q(\mathbb{R}^n)}$$

holds with equality if f and g are non-negative almost everywhere on $\mathbb{R}^n$.

Proof. See, for example, [14, Theorem 3.9.4]. □

A special situation occurs if $f \in \mathrm{L}^p(\mathbb{R}^n)$ and $g \in \mathrm{L}^q(\mathbb{R}^n)$ with *dual* exponents $1 \leq p, q \leq \infty$, i.e.,

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Theorem A.1.17. *Let $f \in \mathrm{L}^p(\mathbb{R}^n)$ and $g \in \mathrm{L}^q(\mathbb{R}^n)$ with $1 \leq p, q \leq \infty$ satisfying*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then, $f * g$ is bounded and continuous on $\mathbb{R}^n$, i.e., $f * g \in \mathscr{C}_b(\mathbb{R}^n)$, where

$$\mathscr{C}_b(\mathbb{R}^n) = \{f \in \mathscr{C}(\mathbb{R}^n) \mid \|f\|_\infty < \infty\}.$$

If we further have $1 < p, q < \infty$, then $f * g$ vanishes at infinity, i.e., $f * g \in \mathscr{C}_0(\mathbb{R}^n)$.

Proof. See, for example, [16, Theorem 3.14]. □

We now list some basic properties of the convolution product.

Proposition A.1.18. *The convolution product satisfies the following properties.*

(i) Commutativity:

$$f * g = g * f \quad \forall\, f, g \in \mathrm{L}^1(\mathbb{R}^n)$$

(ii) Linearity:

$$f * (\alpha\, g + \beta\, h) = \alpha\, (f * g) + \beta\, (f * h) \quad \forall\, \alpha, \beta \in \mathbb{R},\ f, g, h \in \mathrm{L}^1(\mathbb{R}^n)$$

(iii) Integration:

$$\int_{\mathbb{R}^n} (f * g)(x)\, \mathrm{d}x = \left(\int_{\mathbb{R}^n} f(x)\, \mathrm{d}x\right) \left(\int_{\mathbb{R}^n} g(x)\, \mathrm{d}x\right) \quad \forall\, f, g \in \mathrm{L}^1(\mathbb{R}^n)$$

(iv) Translation: *For $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $a \in \mathbb{R}^n$ we consider the function*

$$f_a(x) = f(x - a) \quad \text{for } x \in \mathbb{R}^n.$$

Then,

$$f_a * g = (f * g)_a \quad \forall\, g \in \mathrm{L}^1(\mathbb{R}^n).$$

(v) Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and let $g \in \mathscr{C}^k(\mathbb{R}^n)$, $k \in \mathbb{N}$, be bounded such that its derivatives $\mathrm{D}^\alpha g$ are also bounded for all multi-indices $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$. Then, we have $f * g \in \mathscr{C}^k(\mathbb{R}^n)$ and

$$\mathrm{D}^\alpha (f * g) = f * \mathrm{D}^\alpha g.$$

Proof. See, for example, [16, Chapter 3.3] and [55, Chapter 10.1]. □

We finish this section by stating the most important property of the convolution product, which is given by the classical *Fourier convolution theorem* and describes the interaction between the convolution product and the Fourier transform.

Theorem A.1.19 (Fourier convolution theorem)**.** *Let $f, g \in \mathrm{L}^1(\mathbb{R}^n)$ be given functions. Then, we have*

$$\mathcal{F}(f * g) = \mathcal{F}f \cdot \mathcal{F}g$$

and

$$\mathcal{F}^{-1}(f * g) = (2\pi)^n\, \mathcal{F}^{-1}f \cdot \mathcal{F}^{-1}g.$$

Additionally, if $\mathcal{F}f, \mathcal{F}g \in \mathrm{L}^1(\mathbb{R}^n)$, then

$$\mathcal{F}(f \cdot g) = (2\pi)^{-n}\, (\mathcal{F}f) * (\mathcal{F}g).$$

Proof. See, for example, [119, Theorem I.1.4]. □

A.2 Distributions

Distributions or, to be more precise, *tempered* distributions play an important role in the definition of Sobolev spaces of fractional order. Thus, in this section we introduce distributions as generalized functions and extend the Fourier transform to the space of tempered distributions. Here, we restrict ourselves to the most basic facts, which can be found in [16], [27], [46], [108].

The space of distributions is given by the topological dual of the space of test functions, which is defined as follows.

Definition A.2.1 (Space of test functions)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain in $\mathbb{R}^n$. Then, the space of* test functions *on Ω is defined as*

$$\mathscr{D}(\Omega) = \{f \in \mathscr{C}^\infty(\Omega) \mid \operatorname{supp}(f) \subseteq \Omega \text{ compact}\}.$$

The following lemma explains the expression 'test function'.

Lemma A.2.2 (Fundamental lemma of variational calculus)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain in $\mathbb{R}^n$ and $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ be locally integrable. Then, we have*

$$f \equiv 0 \quad \text{a.e. on } \Omega \qquad \Longleftrightarrow \qquad \int_\Omega f(x)\,\phi(x)\,\mathrm{d}x = 0 \quad \forall\, \phi \in \mathscr{D}(\Omega).$$

Proof. See, for example, [16, Lemma 2.75]. □

Calculating the integral $\int_\Omega f(x)\,\phi(x)\,\mathrm{d}x$ is also called *testing* the function $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ with $\phi \in \mathscr{D}(\Omega)$. Thus, the fundamental lemma of variational calculus, Lemma A.2.2, states that $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ is almost everywhere uniquely determined by testing with all functions $\phi \in \mathscr{D}(\Omega)$.

Example A.2.3. *The function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ with*

$$f(x) = \begin{cases} \exp\left(-\frac{1}{1-\|x\|^2_{\mathbb{R}^n}}\right) & \text{for } \|x\|_{\mathbb{R}^n} < 1 \\ 0 & \text{for } \|x\|_{\mathbb{R}^n} \geq 1 \end{cases}$$

is a test function on $\mathbb{R}^n$, i.e., it satisfies $f \in \mathscr{D}(\mathbb{R}^n)$.

We now introduce the notion of a distribution.

Definition A.2.4 (Distribution)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain in $\mathbb{R}^n$. The topological dual space of $\mathscr{D}(\Omega)$ with respect to the natural topology, denoted by $\mathscr{D}'(\Omega)$, is called the space of* distributions*.*

We note that, for a domain $\Omega \subseteq \mathbb{R}^n$, each function $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ induces a distribution $T_f \in \mathscr{D}'(\Omega)$ via

$$T_f(\phi) = \int_\Omega f(x)\,\phi(x)\,\mathrm{d}x \quad \text{for } \phi \in \mathscr{D}(\Omega).$$

In this sense we have $\mathrm{L}^1_{\mathrm{loc}}(\Omega) \subset \mathscr{D}'(\Omega)$ and distributions of the form T_f are called *regular*. Because of the relation between f and T_f, distributions are also called *generalized functions*.

However, there also exist distributions that are not regular. One example is the well-known *Dirac distribution* $\delta \in \mathscr{D}'(\mathbb{R}^n)$ with

$$\delta(f) = f(0) \quad \text{for } f \in \mathscr{D}(\mathbb{R}^n)$$

or, for fixed $x_0 \in \mathbb{R}^n$, the *shifted Dirac distribution* $\delta_{x_0} \in \mathscr{D}'(\mathbb{R}^n)$ with

$$\delta_{x_0}(f) = f(x_0) \quad \text{for } f \in \mathscr{D}(\mathbb{R}^n).$$

In what follows, we denote the action of a distribution $T \in \mathscr{D}'(\Omega)$ on a test function $\phi \in \mathscr{D}(\Omega)$ by the *duality pairing*

$$\langle T, \phi \rangle = T(\phi).$$

With this we define the derivative of a distribution as follows.

Definition A.2.5 (Derivative of a distribution)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain and $T \in \mathscr{D}'(\Omega)$. For $\alpha \in \mathbb{N}_0^n$, we define the* derivative $\mathrm{D}^\alpha T \in \mathscr{D}'(\Omega)$ *of T via*

$$\langle \mathrm{D}^\alpha T, \phi \rangle = (-1)^{|\alpha|} \langle T, \mathrm{D}^\alpha \phi \rangle \quad \text{for } \phi \in \mathscr{D}(\Omega).$$

We use the same technique to define the multiplication of a distribution with a $\mathscr{C}^\infty$-function. To this end, we note that the product $g \cdot \phi$ of $g \in \mathscr{C}^\infty(\Omega)$ and $\phi \in \mathscr{D}(\Omega)$ is again in $\mathscr{D}(\Omega)$.

Definition A.2.6 (Multiplication with a $\mathscr{C}^\infty$-function)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain. Further, let $T \in \mathscr{D}'(\Omega)$ be a distribution and $g \in \mathscr{C}^\infty(\Omega)$. Then, the distribution $g \cdot T \in \mathscr{D}'(\Omega)$ is defined as*

$$\langle g \cdot T, \phi \rangle = \langle T, g \cdot \phi \rangle \quad \text{for } \phi \in \mathscr{D}(\Omega).$$

We would like to use this technique to also define the Fourier transform of a distribution. However, the Fourier transform of a non-trivial test function is not a test function, since it cannot have compact support due to the Paley-Wiener Theorem A.1.14. To resolve this problem, we have to restrict ourselves to a smaller subspace of distributions, the so called *tempered distributions*. This space is given by the topological dual of the so called *Schwartz space* of rapidly decreasing functions, which is defined as follows.

Definition A.2.7 (Schwartz space)**.** *The* Schwartz space $\mathcal{S}(\mathbb{R}^n)$ *of rapidly decaying functions is defined as*

$$\mathcal{S}(\mathbb{R}^n) = \left\{ f \in \mathscr{C}^\infty(\mathbb{R}^n) \mid \forall\, \alpha, \beta \in \mathbb{N}_0^n : \ |f|_{\alpha,\beta} < \infty \right\},$$

where

$$|f|_{\alpha,\beta} = \sup_{x \in \mathbb{R}^n} |x^\alpha \, \mathrm{D}^\beta f(x)| \quad \text{for } \alpha, \beta \in \mathbb{N}_0^n.$$

The space of Schwartz functions plays a central role in the theory of Fourier transforms.

Theorem A.2.8. *The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ is an automorphism of $\mathcal{S}(\mathbb{R}^n)$. In particular,*

$$\mathcal{F}^{-1}(\mathcal{F} f) = f = \mathcal{F}(\mathcal{F}^{-1} f) \quad \forall\, f \in \mathcal{S}(\mathbb{R}^n).$$

Proof. See, for example, [16, Theorem 4.15]. □

Since $\mathscr{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, we have $\mathcal{S}'(\mathbb{R}^n) \subset \mathscr{D}'(\mathbb{R}^n)$ and the dual space of $\mathcal{S}(\mathbb{R}^n)$ consists of a special class of distributions. These are called *tempered distributions*.

Definition A.2.9 (Tempered distribution)**.** *The topological dual space of $\mathcal{S}(\mathbb{R}^n)$ with respect to the natural topology, denoted by $\mathcal{S}'(\mathbb{R}^n)$, is called the space of* tempered distributions.

Since the Fourier transform of a Schwartz function is again a Schwartz function, we can now define the Fourier transform of tempered distributions.

Definition A.2.10 (Fourier transform of tempered distributions)**.** *Let $T \in \mathcal{S}'(\mathbb{R}^n)$ be tempered. Then, its Fourier transform $\mathcal{F}T \in \mathcal{S}'(\mathbb{R}^n)$ is defined via*

$$\langle \mathcal{F}T, f \rangle = \langle T, \mathcal{F} f \rangle \quad \text{for } f \in \mathcal{S}(\mathbb{R}^n).$$

Analogously, its inverse Fourier transform $\mathcal{F}^{-1}T \in \mathcal{S}'(\mathbb{R}^n)$ is given by

$$\langle \mathcal{F}^{-1}T, f \rangle = \langle T, \mathcal{F}^{-1} f \rangle \quad \text{for } f \in \mathcal{S}(\mathbb{R}^n).$$

We remark that, if $T \in \mathcal{S}'(\mathbb{R}^n)$ is regular and given by $T = T_f$ for some function $f \in \mathrm{L}^1(\mathbb{R}^n)$, we have

$$\mathcal{F}T_f = T_{\mathcal{F}f}$$

due to Parseval's identity, cf. Theorem A.1.11. Hence, the definition of the classical and the distributional Fourier transform coincide on $\mathrm{L}^1(\mathbb{R}^n)$.

Remark A.2.11. *We have* $\mathrm{L}^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ *for all* $1 \leq p \leq \infty$ *in the sense that the functional* $T_f : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathbb{R}$,

$$\langle T_f, \phi \rangle = \int_{\mathbb{R}^n} f(x)\, \phi(x)\, \mathrm{d}x \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}^n),$$

is a tempered distribution. This observation implies that the Fourier transform is now defined for all L^p*-spaces. However, the Fourier transform of* $f \in \mathrm{L}^p(\mathbb{R}^n)$ *with* $p > 2$ *is in general not a function, but only a distribution, in contrast to the case of* $p \leq 2$.

Like the Schwartz space $\mathcal{S}(\mathbb{R}^n)$, also the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ plays a central role in Fourier analysis.

Theorem A.2.12. *The Fourier transform* $\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n)$ *is an automorphism of* $\mathcal{S}'(\mathbb{R}^n)$ *with respect to the weak topology. In particular, we have*

$$\mathcal{F}^{-1}(\mathcal{F}f) = f = \mathcal{F}(\mathcal{F}^{-1}f) \quad \forall\, f \in \mathcal{S}'(\mathbb{R}^n).$$

Proof. See, for example, [16, Theorem 4.25]. □

Many properties of the regular Fourier transform carry over to the distributional Fourier transform. We start with the property that $\mathcal{F}$ translates differentiation into multiplication. To this end, we first remark that for $T \in \mathcal{S}'(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ the distributional derivative $\mathrm{D}^\alpha T \in \mathcal{S}'(\mathbb{R}^n)$ is again tempered. Further, the multiplication with a function $f \in \mathscr{C}^\infty(\mathbb{R}^n)$ of at most polynomial growth is well-defined via

$$\langle f \cdot T, \phi \rangle = \langle T, f \cdot \phi \rangle \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Proposition A.2.13. *For* $T \in \mathcal{S}'(\mathbb{R}^n)$ *and* $\alpha \in \mathbb{N}_0^n$*, we have*

$$\mathcal{F}(\mathrm{D}^\alpha T) = \mathrm{i}^{|\alpha|}\, x^\alpha\, \mathcal{F}T.$$

Proof. See, for example, [124, Theorem VIII.5.12]. □

We proceed to the distributional definition of the convolution product. For this, recall the definition of the parity operator $^* : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ in (A.6).

Definition A.2.14 (Convolution of a tempered distribution with a function)**.** *For* $T \in \mathcal{S}'(\mathbb{R}^n)$ *and* $f \in \mathcal{S}(\mathbb{R}^n)$ *the convolution product* $f * T \in \mathcal{S}'(\mathbb{R}^n)$ *is defined as*

$$\langle f * T, \phi \rangle = \langle T, f^* * \phi \rangle \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}^n).$$

We close this section on distributions by stating the Fourier convolution theorem for the distributional Fourier transform.

Theorem A.2.15 (Fourier convolution theorem)**.** *Let* $T \in \mathcal{S}'(\mathbb{R}^n)$ *and* $f \in \mathcal{S}(\mathbb{R}^n)$ *be given. Then, we have*

$$\mathcal{F}(f * T) = \mathcal{F}f \cdot \mathcal{F}T.$$

Proof. See, for example, [46, Theorem 7.1.18]. □

A.3 Sobolev spaces

Sobolev spaces play an important role in our error analysis of the filtered back projection method in Chapter 4 and are of particular interest for understanding of the ill-posedness of the CT reconstruction problem, see Section 2.2.3. For this reason, we now define Sobolev spaces of fractional order and list some basic properties. This is based on a characterization of regular Sobolev spaces by means of the Fourier transform. For further details we refer to the textbook [2].

We begin with the standard definition of Sobolev spaces. To this end, let us recall that for a domain $\Omega \subseteq \mathbb{R}^n$, a multi-index $\alpha \in \mathbb{N}_0^n$ and a distribution $T \in \mathscr{D}'(\Omega)$ the derivative $\mathrm{D}^\alpha f \in \mathscr{D}'(\Omega)$ is defined via

$$\langle \mathrm{D}^\alpha T, \phi \rangle = (-1)^{|\alpha|} \langle T, \mathrm{D}^\alpha \phi \rangle \quad \text{for } \phi \in \mathscr{D}(\Omega).$$

In general, the distributional derivative $\mathrm{D}^\alpha f$ of a function $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega) \subset \mathscr{D}'(\Omega)$ is not a function itself. But in case it is, $\mathrm{D}^\alpha f$ is called *weak derivative* of f.

Definition A.3.1 (Weak derivative)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain, $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ be locally integrable and $\alpha \in \mathbb{N}_0^n$. If there exists a function $g \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ with*

$$\int_\Omega g(x)\, \phi(x)\, \mathrm{d}x = (-1)^{|\alpha|} \int_\Omega f(x)\, \mathrm{D}^\alpha \phi(x)\, \mathrm{d}x \quad \forall\, \phi \in \mathscr{D}(\Omega),$$

then f is called weakly differentiable *on Ω with weak derivative $\mathrm{D}^\alpha f = g$. If the weak derivatives $\mathrm{D}^\alpha f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ exist for all $|\alpha| \leq k$ with $k \in \mathbb{N}$, then f is called k-*times weakly differentiable.

Remark A.3.2. *Weak derivatives are uniquely determined almost everywhere on Ω according to the fundamental lemma of variational calculus, Lemma A.2.2.*

Now, the common Sobolev spaces are defined as spaces of functions whose weak derivatives belong to certain L^p-spaces.

Definition A.3.3 (Sobolev space)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be a domain, $1 \leq p \leq \infty$ and $k \in \mathbb{N}_0$. Then, the* Sobolev space $\mathrm{H}^{k,p}(\Omega)$ *is defined as*

$$\mathrm{H}^{k,p}(\Omega) = \{ f \in \mathrm{L}^p(\Omega) \mid \forall\, |\alpha| \leq k :\ \mathrm{D}^\alpha f \in \mathrm{L}^p(\Omega) \}$$

and equipped with the Sobolev norm

$$\|f\|_{\mathrm{H}^{k,p}(\Omega)} = \begin{cases} \left(\sum_{|\alpha| \leq k} \|\mathrm{D}^\alpha f\|^p_{\mathrm{L}^p(\Omega)} \right)^{1/p} & \text{for } p < \infty \\ \max_{|\alpha| \leq k} \|\mathrm{D}^\alpha f\|_{\mathrm{L}^\infty(\Omega)} & \text{for } p = \infty. \end{cases}$$

Remark A.3.4. *For $p = 2$ we simply write $\mathrm{H}^k(\Omega) \equiv \mathrm{H}^{k,2}(\Omega)$ and these spaces are Hilbert spaces with the inner product*

$$(f, g)_{\mathrm{H}^k(\Omega)} = \sum_{|\alpha| \leq k} (\mathrm{D}^\alpha f, \mathrm{D}^\alpha g)_{\mathrm{L}^2(\Omega)} \quad \text{for } f, g \in \mathrm{H}^k(\Omega).$$

Since $\mathrm{H}^{k,p}(\mathbb{R}^n) \subset \mathrm{L}^p(\mathbb{R}^n)$, we further have $\mathrm{H}^{k,p}(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ in virtue of Remark A.2.11. In particular, we get the following variant of Proposition A.2.13.

Lemma A.3.5. *Let $f \in \mathrm{L}^2(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ so that the weak derivative $\mathrm{D}^\alpha f$ is also in $\mathrm{L}^2(\mathbb{R}^n)$. Then, we have*

$$\mathcal{F}(\mathrm{D}^\alpha f) = \mathrm{i}^{|\alpha|}\, x^\alpha\, \mathcal{F} f$$

and, if $x^\alpha f \in \mathrm{L}^2(\mathbb{R}^n)$,

$$\mathcal{F}(x^\alpha f) = \mathrm{i}^{|\alpha|}\, \mathrm{D}^\alpha \mathcal{F} f.$$

Proof. See, for example, [16, Lemma 4.28]. □

This lemma yields the following characterization of regular Sobolev spaces by means of the Fourier transform.

Theorem A.3.6 (Characterization of $\mathrm{H}^k(\mathbb{R}^n)$)**.** *For $k \in \mathbb{N}$ we have*

$$f \in \mathrm{H}^k(\mathbb{R}^n) \quad \Longleftrightarrow \quad \int_{\mathbb{R}^n} \left(1 + \|\omega\|^2_{\mathbb{R}^n}\right)^k |\mathcal{F} f(\omega)|^2\, \mathrm{d}\omega < \infty.$$

Proof. See, for example, [16, Theorem 4.29]. □

Observe that Theorem A.3.6 relates the weak differentiability of a function $f \in \mathrm{L}^2(\mathbb{R}^n)$ to the decay properties of its Fourier transform $\mathcal{F}f$. This characterization does not only give a useful tool to investigate the smoothness of a function but also offers a possibility to generalize the definition of Sobolev spaces $\mathrm{H}^k(\mathbb{R}^n)$ of integer order to spaces $\mathrm{H}^\alpha(\mathbb{R}^n)$ of arbitrary smoothness order $\alpha \in \mathbb{R}$. However, if $\alpha < 0$, we have to enlarge the basic set from $\mathrm{L}^2(\mathbb{R}^n)$ to the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$.

Definition A.3.7 (Sobolev space of fractional order)**.** *The* Sobolev space $\mathrm{H}^\alpha(\mathbb{R}^n)$ of fractional order $\alpha \in \mathbb{R}$ *is defined as*

$$\mathrm{H}^\alpha(\mathbb{R}^n) = \left\{ f \in \mathcal{S}'(\mathbb{R}^n) \mid \|f\|_{\mathrm{H}^\alpha(\mathbb{R}^n)} < \infty \right\},$$

where the Sobolev norm $\|\cdot\|_{\mathrm{H}^\alpha(\mathbb{R}^n)}$ *is given by*

$$\|f\|^2_{\mathrm{H}^\alpha(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \left(1 + \|\omega\|^2_{\mathbb{R}^n}\right)^\alpha |\mathcal{F}f(\omega)|^2 \,\mathrm{d}\omega.$$

Further, for an open subset $\Omega \subseteq \mathbb{R}^n$*, we define the Sobolev space* $\mathrm{H}^\alpha_0(\Omega)$ *by*

$$\mathrm{H}^\alpha_0(\Omega) = \left\{ f \in \mathrm{H}^\alpha(\mathbb{R}^n) \mid \mathrm{supp}(f) \subseteq \overline{\Omega} \right\},$$

where the support $\mathrm{supp}(f)$ *of a tempered distribution* $f \in \mathcal{S}'(\mathbb{R}^n)$ *is defined as the complement of the largest open set* $U \subset \mathbb{R}^n$ *for which* $\langle f, \phi \rangle = 0$ *for all* $\phi \in \mathcal{S}(\mathbb{R}^n)$ *with* $\mathrm{supp}(\phi) \subset U$*.*

For $\alpha \in \mathbb{N}_0$, Theorem A.3.6 shows that the space $\mathrm{H}^\alpha(\mathbb{R}^n)$ consists of those functions whose (distributional) derivatives up to order α are square-integrable. Therefore, Definition A.3.7 is compatible with the classical definition of Sobolev spaces. In particular, for $\alpha = 0$ we simply have

$$\mathrm{H}^0(\mathbb{R}^n) = \mathrm{L}^2(\mathbb{R}^n).$$

By defining the equivalent Sobolev norms $\|\cdot\|_\alpha$ on $\mathrm{H}^\alpha(\mathbb{R}^n)$ for $\alpha \in \mathbb{R}$ via

$$\|f\|^2_\alpha = (2\pi)^{-n} \int_{\mathbb{R}^n} \left(1 + \|\omega\|^2_{\mathbb{R}^n}\right)^\alpha |\mathcal{F}f(\omega)|^2 \,\mathrm{d}\omega \quad \text{for } f \in \mathrm{H}^\alpha(\mathbb{R}^n),$$

we further obtain

$$\|\cdot\|_0 = \|\cdot\|_{\mathrm{L}^2(\mathbb{R}^n)}$$

according to the Rayleigh-Plancherel Theorem A.1.12.

We close this section with some final remarks on Sobolev spaces.

(i) The Sobolev space $\mathrm{H}^\alpha(\mathbb{R}^n)$ is a Hilbert space with the inner product

$$(f, g)_{\mathrm{H}^\alpha(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \left(1 + \|\omega\|^2_{\mathbb{R}^n}\right)^\alpha \mathcal{F}f(\omega)\, \overline{\mathcal{F}g}(\omega) \,\mathrm{d}\omega \quad \text{for } f, g \in \mathrm{H}^\alpha(\mathbb{R}^n).$$

(ii) For $\alpha < \beta$ we have $\mathrm{H}^\beta(\mathbb{R}^n) \subset \mathrm{H}^\alpha(\mathbb{R}^n)$ and, in particular,

$$\mathrm{H}^\alpha(\mathbb{R}^n) \subset \mathrm{L}^2(\mathbb{R}^n) \quad \forall\, \alpha > 0.$$

Thus, for $\alpha > 0$, the Sobolev space $\mathrm{H}^\alpha(\mathbb{R}^n)$ can equivalently be defined as

$$\mathrm{H}^\alpha(\mathbb{R}^n) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^n) \mid \|f\|_\alpha < \infty \right\}.$$

(iii) The dual space of $\mathrm{H}^\alpha(\mathbb{R}^n)$ is topologically isomorphic to $\mathrm{H}^{-\alpha}(\mathbb{R}^n)$.

A.4 General mathematical tools

In this section we finally list some general tools that are used in the derivation of the approximate reconstruction formula in Chapter 3 and in our error analysis of the FBP method in Chapter 4.

We start with the following generalized Hölder inequality.

Theorem A.4.1 (Hölder inequality)**.** *Let $f_j \in \mathrm{L}^{p_j}(\mathbb{R}^n)$ with $1 \leq p_j \leq \infty$ for all $j = 1, \ldots, k$. Then, we have $\prod_{j=1}^k f_j \in \mathrm{L}^r(\mathbb{R}^n)$ with $1 \leq r \leq \infty$ satisfying*

$$\sum_{j=1}^{k} \frac{1}{p_j} = \frac{1}{r}$$

and

$$\left\| \prod_{j=1}^{k} f_j \right\|_{\mathrm{L}^r(\mathbb{R}^n)} \leq \prod_{j=1}^{k} \|f_j\|_{\mathrm{L}^{p_j}(\mathbb{R}^n)}.$$

Proof. See, for example, [14, Corollary 2.11.5]. □

We continue with Taylor's theorem in one variable.

Theorem A.4.2 (Taylor's theorem)**.** *For $k \in \mathbb{N}$, let the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ be k-times differentiable at a point $x_0 \in \mathbb{R}$. Then, there exists a function $h_k : \mathbb{R} \longrightarrow \mathbb{R}$ such that*

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + h_k(x) (x - x_0)^k \quad \forall\, x \in \mathbb{R}$$

and

$$\lim_{x \to x_0} h_k(x) = 0.$$

Further, if f is k-times differentiable on the open interval with $f^{(k-1)}$ continuous on the closed interval between x_0 and x, then there exists some $\xi \in \mathbb{R}$ between x_0 and x such that

$$f(x) = \sum_{j=0}^{k-1} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + \frac{f^{(k)}(\xi)}{k!} (x - x_0)^k.$$

Proof. See, for example, [42, Chapter VII]. □

We now turn to absolutely continuous functions on the real line, see [35], [107] for details.

Definition A.4.3 (Absolute continuity)**.** *A function $f : [a, b] \longrightarrow \mathbb{R}$ is* absolutely continuous *on the interval $[a, b] \subset \mathbb{R}$ if for every $\varepsilon > 0$ there exists some $\delta > 0$ such that for any finite set of pairwise disjoint open subintervals $\{(a_j, b_j)\}_{j \in J}$ of $[a, b]$ we have*

$$\sum_{j \in J} (b_j - a_j) < \delta \quad \Longrightarrow \quad \sum_{j \in J} |f(b_j) - f(a_j)| < \varepsilon.$$

We define the function spaces

$$\mathcal{AC}([a, b]) = \{f : [a, b] \longrightarrow \mathbb{R} \mid f \text{ is absolutely continuous on } [a, b]\}$$

and

$$\mathscr{C}^{0,1}([a, b]) = \{f : [a, b] \longrightarrow \mathbb{R} \mid f \text{ is Lipschitz continuous on } [a, b]\},$$

where a function $f : [a, b] \longrightarrow \mathbb{R}$ is *Lipschitz continuous* on $[a, b]$ if there exists a constant $C > 0$ such that

$$|f(x) - f(y)| \leq C\, |x - y| \quad \forall\, x, y \in [a, b].$$

We remark that we have

$$\mathscr{C}^1([a,b]) \subsetneq \mathscr{C}^{0,1}([a,b]) \subsetneq \mathcal{AC}([a,b]) \subsetneq \mathscr{C}([a,b]).$$

In addition, any absolutely continuous function $f \in \mathcal{AC}([a,b])$ is of bounded variation and its pointwise derivative exists almost everywhere on $[a,b]$, see [107, Lemma 4.11 & Corollary 4.12].

Further note that for any $f \in \mathrm{L}^1([a,b])$ the function $g : [a,b] \longrightarrow \mathbb{R}$ with

$$g(x) = \int_a^x f(t)\,\mathrm{d}t \quad \text{for } x \in [a,b]$$

is absolutely continuous on $[a,b]$, i.e., $g \in \mathcal{AC}([a,b])$, and differentiable almost everywhere with

$$g'(x) = f(x) \quad \text{for almost all } x \in [a,b].$$

In fact, much more holds true.

Theorem A.4.4 (Fundamental theorem of calculus of $\mathcal{AC}$-functions)**.** *For $g : [a,b] \longrightarrow \mathbb{R}$ the following statements are equivalent:*

(i) g is absolutely continuous on $[a,b]$, i.e., $g \in \mathcal{AC}([a,b])$.

(ii) There exists some function $f \in \mathrm{L}^1([a,b])$ such that

$$g(x) - g(a) = \int_a^x f(t)\,\mathrm{d}t \quad \forall\, x \in [a,b].$$

(iii) g is differentiable almost everywhere on $[a,b]$ with $g' \in \mathrm{L}^1([a,b])$ and

$$g(x) - g(a) = \int_a^x g'(t)\,\mathrm{d}t \quad \forall\, x \in [a,b].$$

Proof. See, for example, [35, Theorem 3.35]. □

An important fact is that integration by parts is valid for absolutely continuous functions.

Theorem A.4.5 (Integration by parts)**.** *Let $f, g \in \mathcal{AC}([a,b])$. Then,*

$$\int_a^b f(x)\,g'(x)\,\mathrm{d}x = f(b)\,g(b) - f(a)\,g(a) - \int_a^b f'(x)\,g(x)\,\mathrm{d}x.$$

Proof. See, for example, [35, Theorem 3.36]. □

We close this section with the following remark on Lipschitz continuous functions.

Remark A.4.6. *Let $f \in \mathscr{C}^{0,1}([a,b])$ be Lipschitz continuous on the interval $[a,b]$ with Lipschitz constant $C > 0$, i.e.,*

$$|f(x) - f(y)| \leq C\,|x - y| \quad \forall\, x, y \in [a,b].$$

Then, f is also absolutely continuous on $[a,b]$, $f \in \mathcal{AC}([a,b])$, and its pointwise derivative f', which exists almost everywhere on $[a,b]$, is essentially bounded with

$$|f'(x)| \leq C \quad \textit{for almost all } x \in [a,b].$$

Conversely, if $f \in \mathcal{AC}([a,b])$ is absolutely continuous on $[a,b]$ and has an essentially bounded derivative $f' \in \mathrm{L}^\infty([a,b])$, then f is also Lipschitz continuous on $[a,b]$, $f \in \mathscr{C}^{0,1}([a,b])$, and satisfies

$$|f(x) - f(y)| \leq \|f'\|_{\mathrm{L}^\infty([a,b])}\,|x - y| \quad \forall\, x, y \in [a,b].$$

Notations

List of symbols

Symbol	Description	Page
$\ell_{t,\theta}$	straight line in the plane parametrized by $(t,\theta) \in \mathbb{R} \times [0,\pi)$	6
$\mathcal{R}$	Radon transform	8
$\mathcal{B}$	back projection	14
$\mathcal{F}$	Fourier transform	133
	– for bivariate functions in Cartesian coordinates	13
	– for bivariate functions in polar coordinates	13
$*$	convolution product	136
	– for bivariate functions in Cartesian coordinates	17
	– for bivariate functions in polar coordinates	17
$\langle \cdot, \cdot \rangle$	duality pairing	138
χ_K	characteristic function of $K \subseteq \mathbb{R}^n$	12
$\operatorname{diam}(K)$	diameter of $K \subseteq \mathbb{R}^n$	11
$\sqcap_L$	characteristic function of $[-L,L] \subset \mathbb{R}$	38
W_L	window function with bandwidth $L > 0$	33
A_L	low-pass filter with window W and bandwidth L	33
q_L	convolving function (inverse Fourier transform of filter A_L)	34
K_L	convolution kernel (back projection of convolving function q_L)	35
f_L	approximate FBP reconstruction of the function f	33
e_L	inherent FBP reconstruction error	54
R_L	FBP regularization operator	94
g^δ	noisy measurements with noise level $\delta > 0$	54
f_L^δ	approximate FBP reconstruction from noisy measurements g^δ	54
e_L^δ	overall FBP reconstruction error	54
$\mathcal{B}_D$	discrete back projection	105
$*_D$	discrete convolution product	105
$\mathcal{I}$	interpolation method	105
f_{FBP}	discrete FBP reconstruction	105
RMSE	root mean square error	123

List of spaces

Bibliography

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, 9th ed. New York: Dover Publications, 1972.

[2] R. Adams and J. Fournier, *Sobolev spaces*, 2nd ed., ser. Pure and Applied Mathematics 140. Amserdam: Academic Press, 2003.

[3] J. Ahlberg, E. Nilson, and J. Walsh, *The Theory of Splines and Their Applications*, ser. Mathematics in Science and Engineering 38. New York: Academic Press, 1967.

[4] F. Arcadu, J. Vogel, M. Stampanoni, and F. Marone, "Improving analytical tomographic reconstructions through consistency conditions", *Fundamenta Informaticae*, vol. 155, no. 4, pp. 341–361, 2017.

[5] K. Batenburg and L. Plantagie, "Fast approximation of algebraic reconstruction methods for tomography", *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3648–3658, 2012.

[6] M. Beckmann, "Error estimates for filtered back projection", Master's thesis, University of Hamburg, 2014.

[7] M. Beckmann and A. Iske, "Error estimates for filtered back projection", in *IEEE 2015 International Conference on Sampling Theory and Applications (SampTA2015)*, 2015, pp. 553–557.

[8] ——, "Analysis of the inherent reconstruction error in filtered back projection", University of Hamburg, Hamburger Beiträge zur Angewandten Mathematik 2016-01, 2016.

[9] ——, "Error estimates and convergence rates for filtered back projection", University of Hamburg, Hamburger Beiträge zur Angewandten Mathematik 2016-06, 2016, accepted for publication in Mathematics of Computation.

[10] ——, "On the error behaviour of the filtered back projection", *Proceedings in Applied Mathematics and Mechanics (PAMM)*, vol. 16, no. 1, pp. 833–834, 2016.

[11] ——, "Approximation of bivariate functions from fractional Sobolev spaces by filtered back projection", University of Hamburg, Hamburger Beiträge zur Angewandten Mathematik 2017-05, 2017.

[12] ——, "Sobolev error estimates for filtered back projection reconstructions", in *IEEE 2017 International Conference on Sampling Theory and Applications (SampTA2017)*, 2017, pp. 251–255.

[13] M. Bertero, H. Lanteri, and L. Zanni, "Iterative image reconstruction: A point of view", in *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, Y. Censor, M. Jiang, and A. Louis, Eds., ser. CRM Series. Pisa: Edizioni della Normale, 2008, vol. 7, pp. 37–63.

[14] V. Bogachev, *Measure Theory: Volume I.* Berlin: Springer, 2007.

[15] C. de Boor, *A Practical Guide to Splines*, ser. Applied Mathematical Science 27. New York: Springer, 2001.

[16] K. Bredies and D. Lorenz, *Mathematische Bildverarbeitung: Einführung in Grundlagen und moderne Theorie*. Wiesbaden: Vieweg+Teubner, 2011.

[17] D. Calvetti and E. Somersalo, "Statistical methods in imaging", in *Handbook of Mathematical Methods in Imaging*, O. Scherzer, Ed., 2nd ed. New York: Springer, 2015, pp. 1343–1392.

[18] T. Chang and T. Gabor, "A scientific study of filter selection for a fan-beam convolution reconstruction algorithm", *SIAM Journal on Applied Mathematics*, vol. 39, no. 1, pp. 83–105, 1980.

[19] O. Christensen, *An Introduction to Frames and Riesz Bases*, 2nd ed., ser. Applied and Numerical Harmonic Analysis. Basel: Birkhäuser, 2016.

[20] A. Cormack, "Representation of a function by its line integrals, with some radiological applications. II", *Journal of Applied Physics*, vol. 35, no. 10, pp. 2908–2913, 1964.

[21] A. Cormack, "Representation of a function by its line integrals, with some radiological applications", *Journal of Applied Physics*, vol. 34, no. 9, pp. 2722–2727, 1963.

[22] G. Davis, P. Munshi, and J. Elliott, "An analysis of biological hard tissues using the tomographic reconstruction error formula", *Journal of X-Ray Science and Technology*, vol. 6, no. 1, pp. 63–76, 1996.

[23] M. Davison and F. Grunbaum, "Tomographic reconstruction with arbitrary directions", *Communications on Pure and Applied Mathematics*, vol. 34, no. 1, pp. 77–119, 1981.

[24] S. De Marchi, A. Iske, and G. Santin, "Image reconstruction from scattered Radon data by weighted positive definite kernel functions", University of Hamburg, Hamburger Beiträge zur Angewandten Mathematik 2017-02, 2017, accepted for publication in Calcolo.

[25] S. De Marchi, A. Iske, and A. Sironi, "Kernel-based image reconstruction from scattered Radon data", *Dolomites Research Notes on Approximation*, vol. 9, pp. 19–31, Special Issue 2016.

[26] S. Deans, *The Radon Transform and Some of Its Applications*. Mineola: Dover Publications, 2007.

[27] G. van Dijk, *Distribution Theory: Convolution, Fourier Transform, and Laplace Transform*. Berlin: De Gruyter, 2013.

[28] J. Dubourdieu, "Sur un theoreme de M.S. Bernstein relatif a la transformation de Laplace-Stieltjes", *Compositio Mathematica*, vol. 7, pp. 96–111, 1940.

[29] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, ser. Mathematics and Its Applications 375. Dordrecht: Kluwer Academic Publisher, 2000.

[30] C. Epstein, *Introduction to the mathematics of medical imaging*, 2nd ed. Philadelphia: SIAM, 2008.

[31] K. Erhard, M. Grass, S. Hitziger, A. Iske, and T. Nielsen, "Generalized filtered back-projection for digital breast tomosynthesis reconstruction", in *Proceedings of SPIE Medical Imaging 2012: Physics of Medical Imaging*, vol. 8313, 2012.

[32] A. Faridani and E. Ritman, "High-resolution computed tomography from efficient sampling", *Inverse Problems*, vol. 16, no. 3, pp. 635–650, 2000.

[33] T. Feeman, *The Mathematics of Medical Imaging*, 2nd ed., ser. Springer Undergraduate Texts in Mathematics and Technology (SUMAT). New York: Springer, 2015.

[34] G. Folland, *Fourier Analysis and Its Applications*, ser. Pure and Applied Undergraduate Texts 4. Providence: American Mathematical Society, 1992.

[35] ——, *Real Analysis*, 2nd ed. New York: Wiley, 1999.

[36] J. Frikel, "Reconstructions in limited angle X-ray tomography: Characterization of classical reconstructions and adapted curvelet sparse regularization", Ph.D. thesis, Technical University of Munich, 2013.

[37] L. Grafakos and C. Sansing, "Gabor frames and directional time-frequency analysis", *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 47–67, 2008.

[38] B. Hahn and A. Louis, "Reconstruction in the three-dimensional parallel scanning geometry with applications in synchrotron-based X-ray tomography", *Inverse Problems*, vol. 28, no. 4, p. 045 013, 2012.

[39] B. Hahn, A. Louis, and C. Schorr, "Combined reconstruction and edge detection in dimensioning", *Measurement Science and Technology*, vol. 24, no. 12, p. 125 601, 2013.

[40] K. Hämäläinen, L. Harhanen, A. Hauptmann, A. Kallonen, E. Niemi, and S. Siltanen, "Total variation regularization for large-scale X-ray tomography", *International Journal of Tomography and Simulation*, vol. 25, no. 1, pp. 1–25, 2014.

[41] Y. Han, B. Yan, L. Li, X. Xi, and G. Hu, "Rebinned filtered back-projection reconstruction from truncated data for half-covered helical cone-beam computed tomography", *IEEE Transactions on Nuclear Science*, vol. 61, no. 5, pp. 2753–2763, 2014.

[42] G. Hardy, *A Course of Pure Mathematics*, 10th ed. Cambridge: Cambridge University Press, 1958.

[43] S. Helgason, "The Radon transform on Euclidean spaces, compact two-point homogeneous spaces and Grassmann manifolds", *Acta Mathematica*, vol. 113, no. 1, pp. 153–180, 1965.

[44] ——, *The Radon Transform*, 2nd ed., ser. Progress in Mathematics 5. Boston: Birkhäuser, 1999.

[45] G. Herman, *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*, 2nd ed., ser. Advances in Pattern Recognition. London: Springer, 2009.

[46] L. Hörmander, *The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis*, ser. Grundlehren der mathematischen Wissenschaften 256. Berlin: Springer, 1983.

[47] G. Hounsfield, "Computerized transverse axial scanning (tomography): Part 1. Description of system", *British Journal of Radiology*, vol. 46, no. 552, pp. 1016–1022, 1973.

[48] K. Jin, M. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging", *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[49] P. Jonas and A. Louis, "A Sobolev space analysis of linear regularization methods for ill-posed problems", *Journal of Inverse and Ill-posed Problems*, vol. 9, no. 1, pp. 59–74, 2001.

[50] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, ser. Applied Mathematical Sciences. New York: Springer, 2005, vol. 160.

[51] A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, ser. Classics in Applied Mathematics 33. Philadelphia: SIAM, 2001.

[52] I. Kazantsev, "Tomographic reconstruction from arbitrary directions using ridge functions", *Inverse Problems*, vol. 14, no. 3, pp. 635–645, 1998.

[53] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*. New York: Springer, 1996.

[54] V. Kolehmainen, S. Siltanen, S. Järvenpää, J. Kaipio, P. Koistinen, M. Lassas, J. Pirttilä, and E. Somersalo, "Statistical inversion for medical X-ray tomography with few radiographs: II. Application to dental radiology", *Physics in Medicine and Biology*, vol. 48, no. 10, pp. 1465–1490, 2003.

[55] K. Königsberger, *Analysis 2*, 5th ed. Berlin: Springer, 2004.

[56] R. Kress, *Numerical Analysis*, ser. Graduate Texts in Mathematics 181. New York: Springer, 1998.

[57] H. Kruse, "Resolution of reconstruction methods in computerized tomography", *SIAM Journal on Scientific and Statistical Computing*, vol. 10, no. 3, pp. 447–474, 1989.

[58] B. Logan and L. Shepp, "Optimal reconstruction of a function from its projections", *Duke Mathematical Journal*, vol. 42, no. 4, pp. 645–659, 1975.

[59] A. Louis, *Inverse und schlecht gestellte Probleme*. Stuttgart: Teubner, 1989.

[60] ——, "Approximate inverse for linear and some nonlinear problems", *Inverse Problems*, vol. 12, no. 2, pp. 175–190, 1996.

[61] ——, "A unified approach to regularization methods for linear ill-posed problems", *Inverse Problems*, vol. 15, no. 2, pp. 489–498, 1999.

[62] ——, "Filter design in three-dimensional cone beam tomography: Circular scanning geometry", *Inverse Problems*, vol. 19, no. 6, S31–S40, 2003.

[63] ——, "Combining image reconstruction and image analysis with an application to two-dimensional tomography", *SIAM Journal on Imaging Sciences*, vol. 1, no. 2, pp. 188–208, 2008.

[64] ——, "Optimal reconstruction kernels in medical imaging", in *Optimization in Medicine*, C. Alves, P. Pardalos, and L. Vincente, Eds., ser. Springer Optimization and Its Applications. Berlin: Springer, 2008, pp. 153–168.

[65] A. Louis and P. Maass, "A mollifier method for linear operator equations of the first kind", *Inverse Problems*, vol. 6, no. 3, pp. 427–440, 1990.

[66] ——, "Contour reconstruction in 3-D X-ray CT", *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 764–769, 1993.

[67] A. Louis and F. Natterer, "Mathematical problems of computerized tomography", *Proceedings of the IEEE*, vol. 71, no. 3, pp. 379–389, 1983.

[68] A. Louis and T. Schuster, "A novel filter design technique in 2D computerized tomography", *Inverse Problems*, vol. 12, no. 5, pp. 685–696, 1996.

[69] A. Louis and T. Weber, "Image reconstruction and image analysis in tomography: Fan beam and 3D cone beam", in *Interdisciplinary Workshop on Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, 2007, pp. 211–229.

[70] A. Louis, T. Weber, and D. Theis, "Computing reconstruction kernels for circular 3D cone beam tomography", *IEEE Transactions on Medical Imaging*, vol. 27, no. 7, pp. 880–886, 2008.

[71] W. Madych, "Summability and approximate reconstruction from Radon transform data", in *Integral Geometry and Tomography*, E. Grinberg and T. Quinto, Eds., ser. Contemporary Mathematics. Providence: American Mathematical Society, 1990, vol. 113, pp. 189–219.

[72] ——, "Tomography, approximate reconstruction, and continuous wavelet transforms", *Applied and Computational Harmonic Analysis*, vol. 7, no. 1, pp. 54–100, 1999.

[73] ——, "Radon's inversion formulas", *Transactions of the American Mathematical Society*, vol. 356, no. 11, pp. 4475–4491, 2004.

[74] M. McCann and M. Unser, "High-quality parallel-ray X-ray CT back projection using optimized interpolation", *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4639–4647, 2017.

[75] E. Miqueles, N. Koshev, and E. Helou, "A backprojection slice theorem for tomographic reconstruction", *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 894–906, 2018.

[76] P. Munshi, "Error analysis of tomographic filters. I: Theory", *NDT & E International*, vol. 25, no. 4-5, pp. 191–194, 1992.

[77] P. Munshi, M. Maisl, and H. Reiter, "Experimental aspects of the approximate error formula for tomographic reconstruction", *Materials Evaluation*, vol. 55, no. 2, pp. 188–191, 1997.

[78] P. Munshi, R. Rathore, K. Ram, and M. Kalra, "Error estimates for tomographic inversion", *Inverse Problems*, vol. 7, no. 3, pp. 399–408, 1991.

[79] ——, "A numerical investigation of tomographic filters", *Transactions of the American Nuclear Society*, vol. 68, pp. 200–202, 1993.

[80] ——, "Error analysis of tomographic filters II: Results", *NDT & E International*, vol. 26, no. 5, pp. 235–240, 1993.

[81] F. Natterer, "A Sobolev space analysis of picture reconstruction", *SIAM Journal on Applied Mathematics*, vol. 39, no. 3, pp. 402–411, 1980.

[82] ——, *The Mathematics of Computerized Tomography*, ser. Classics in Applied Mathematics 32. Philadelphia: SIAM, 2001.

[83] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, ser. SIAM Monographs on Mathematical Modeling and Computation. Philadelphia: SIAM, 2001.

[84] T. Nielsen, S. Hitziger, M. Grass, and A. Iske, "Filter calculation for X-ray tomosynthesis reconstruction", *Physics in Medicine and Biology*, vol. 57, no. 12, pp. 3915–3930, 2012.

[85] X. Pan, E. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?", *Inverse Problems*, vol. 25, no. 12, p. 123 009, 2009.

[86] D. Pelt and K. Batenburg, "Fast tomographic reconstruction from limited data using artificial neural networks", *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5238–5251, 2013.

[87] ——, "Improving filtered backprojection reconstruction by data-dependent filtering", *IEEE Transactions on Image Processing*, vol. 23, no. 11, pp. 4750–4762, 2014.

[88] ——, "Accurately approximating algebraic tomographic reconstruction by filtered backprojection", in *13th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2015, pp. 158–161.

[89] L. Plantagie and K. Batenburg, "Approximating algebraic tomography methods by filtered backprojection: A local filter approach", *Fundamenta Informaticae*, vol. 135, no. 1-2, pp. 1–19, 2014.

[90] D. Popov, "On convergence of a class of algorithms for the inversion of the numerical Radon transform", in *Mathematical Problems of Tomography*, I. Gel'fand and S. Gindikin, Eds., ser. Translations of Mathematical Monographs 81. Providence: American Mathematical Society, 1990, pp. 7–65.

[91] F. Qi, D. Niu, J. Cao, and S. Chen, "Four logarithmically completely monotonic functions involving gamma function", *Journal of the Korean Mathematical Society*, vol. 45, no. 2, pp. 559–573, 2008.

[92] G. Qu, "Convergence of FBP algorithm for tomography", *Acta Mathematicae Applicatae Sinica*, vol. 32, no. 4, pp. 963–968, 2016.

[93] E. Quinto, "An introduction to X-ray tomography and Radon transforms", in *The Radon Transform, Inverse Problems, and Tomography*, G. Olafsson and E. Quinto, Eds., ser. Proceedings of Symposia in Applied Mathematics. Providence: American Mathematical Society, 2006, vol. 63, pp. 1–24.

[94] J. Radon, "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten", *Berichte Sächsische Akademie der Wissenschaften*, vol. 69, pp. 262–277, 1917.

[95] ——, "On the determination of functions from their integral value along certain manifolds", *IEEE Transactions on Medical Imaging*, vol. 5, no. 4, pp. 176–176, 1986, Translated by P. C. Parks from the original German text.

[96] G. Ramachandran and A. Lakshminarayanan, "Three-dimensional reconstruction from radiographs and electron micrographs: Application of convolutions instead of Fourier transforms", *Proceedings of the National Academy of Science USA*, vol. 68, no. 9, pp. 2236–2240, 1971.

[97] P. Raviart, "An analysis of particle methods", in *Numerical Methods in Fluid Dynamics*, F. Brezzi, Ed., ser. Lecture Notes in Mathematics. Como: Springer, 1983, vol. 1127, pp. 243–324.

[98] A. Rieder, "On filter design principles in 2D computerized tomography", in *Radon Transforms and Tomography*, T. Quinto, L. Ehrenpreis, A. Faridani, F. Gonzalez, and E. Grinberg, Eds., ser. Contemporary Mathematics. Providence: American Mathematical Society, 2001, vol. 278, pp. 207–226.

[99] ——, *Keine Probleme mit Inversen Problemen*. Wiesbaden: Vieweg, 2003.

[100] A. Rieder and A. Faridani, "The semidiscrete filtered backprojection algorithm is optimal for tomographic inversion", *SIAM Journal on Numerical Analysis*, vol. 41, no. 3, pp. 869–892, 2003.

[101] A. Rieder and A. Schneck, "Optimality of the fully discrete filtered backprojection algorithm for tomographic inversion", *Numerische Mathematik*, vol. 108, no. 1, pp. 151–175, 2007.

[102] A. Rieder and T. Schuster, "The approximate inverse in action with an application to computerized tomography", *SIAM Journal on Numerical Analysis*, vol. 37, no. 6, pp. 1909–1929, 2000.

[103] ——, "The approximate inverse in action II: Convergence and stability", *Mathematics of Computation*, vol. 72, no. 243, pp. 1399–1415, 2003.

[104] S. Rit and R. Clackdoyle, "2D filtered backprojection for fan-beam CT with independent rotations of the source and the detector", in *Third International Conference on Image Formation in X-Ray Computed Tomography*, 2014, pp. 249–252.

[105] S. Rit, R. Clackdoyle, P. Keuschnigg, and P. Steininger, "Filtered-backprojection reconstruction for a cone-beam computed tomography scanner with independent source and detector rotations", *Medical Physics*, vol. 43, no. 5, pp. 2344–2352, 2016.

[106] S. Rowland, "Computer implementation of image reconstruction formulas", in *Image Reconstruction from Projections: Implementation and Applications*, T. Gabor, Ed. Berlin: Springer, 1979, pp. 9–79.

[107] H. Royden, *Real Analysis*, 3rd ed. Englewood Cliffs: Prentice Hall, 1988.

[108] W. Rudin, *Functional Analysis*, 2nd ed., ser. International Series in Pure and Applied Mathematics. New York: McGraw-Hill, 1991.

[109] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*, ser. Applied Mathematical Sciences. New York: Springer, 2009, vol. 167.

[110] B. Schomburg, "On the approximation of the delta distribution in Sobolev spaces of negative order", *Applicable Analysis*, vol. 36, no. 1-2, pp. 89–93, 1990.

[111] T. Schuster, *The Method of Approximate Inverse: Theory and Applications*, ser. Lecture Notes in Mathematics 1906. Berlin: Springer, 2007.

[112] L. Shepp and J. Kruskal, "Computerized tomography: The new medical X-ray technology", *The American Mathematical Monthly*, vol. 85, no. 6, pp. 420–439, 1978.

[113] L. Shepp and B. Logan, "The Fourier reconstruction of a head section", *IEEE Transactions on Nuclear Science*, vol. NS-21, no. 3, pp. 21–43, 1974.

[114] S. Siltanen, V. Kolehmainen, S. Järvenpää, J. Kaipio, P. Koistinen, M. Lassas, J. Pirttilä, and E. Somersalo, "Statistical inversion for medical X-ray tomography with few radiographs: I. General theory", *Physics in Medicine and Biology*, vol. 48, no. 10, pp. 1437–1463, 2003.

[115] L. Slater, *Generalized Hypergeometric Functions*. Cambridge: Cambridge University Press, 1966.

[116] K. Smith and F. Keinert, "Mathematical foundations of computed tomography", *Applied Optics*, vol. 24, no. 23, pp. 3950–3957, 1985.

[117] K. Smith, D. Salmon, and S. Wagner, "Practical and mathematical aspects of the problem of reconstructing objects from radiographs", *Bulletin of the American Mathematical Society*, vol. 83, no. 6, pp. 1227–1270, 1977.

[118] E. Stein, *Singular Integrals and Differentiability Properties of Functions*. Princeton: Princeton University Press, 1970.

[119] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton: Princeton University Press, 1971.

[120] E. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd ed. Oxford: Clarendon Press, 1948.

[121] Y. Wei, G. Wang, and J. Hsieh, "An intuitive discussion on the ideal ramp filter in computed tomography (I)", *Computers and Mathematics with Applications*, vol. 49, no. 5-6, pp. 731–740, 2005.

[122] P. Wells and P. Munshi, "An investigation of the theoretical error in tomographic images", *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 93, no. 1, pp. 87–92, 1994.

[123] H. Wendland, *Scattered Data Approximation*, ser. Cambridge Monographs on Applied and Computational Mathematics 17. Cambridge: Cambridge University Press, 2005.

[124] D. Werner, *Funktionalanalysis*, 7th ed. Berlin: Springer, 2011.

[125] D. Widder, *The Laplace Transform*. Princeton: Princeton University Press, 1941.

[126] Y. Xu, "A new approach to the reconstruction of images from Radon projections", *Advances in Applied Mathematics*, vol. 36, no. 4, pp. 388–420, 2006.

Zusammenfassung

Die Methode der gefilterten Rückwärtsprojektion (filtered back projection), kurz FBP-Methode, stellt eine weit verbreitete Rekonstruktionstechnik in der Computertomographie dar, bei der eine unbekannte bivariate Funktion durch Kenntnis ihrer Radon-Daten wiederhergestellt wird. Die Rekonstruktion basiert auf der klassischen FBP-Formel, die eine analytische Inversion der Radon-Transformation aus vollständigen Radon-Daten liefert. Allerdings ist die FBP-Formel sensitiv gegenüber Störungen in den Radon-Daten und somit numerisch instabil. Ein Standardansatz zur Stabilisierung ist die Verwendung eines Tiefpass-Filters von beschränkter Bandbreite und mit einer Fenster-Funktion mit kompaktem Träger. Dies reduziert die Störungssensibilität der Rekonstruktionsformel, liefert jedoch nur eine inexakte Approximation der Zielfunktion.

Das Hauptziel dieser Arbeit ist die Analyse des inhärenten FBP-Rekonstruktionsfehlers, der durch die Einführung des Tiefpass-Filters hervorgerufen wird. Zu diesem Zweck entwickeln wir Fehlerabschätzungen in Sobolev-Räumen mit gebrochener Ordnung und stellen quantitative Kriterien bereit, mit denen die Leistungsfähigkeit des verwendeten Tiefpass-Filters anhand der zugehörigen Fenster-Funktion a priori evaluiert werden kann. Die gewonnenen Fehlerschranken hängen ab von der Bandbreite des Tiefpass-Filters, der Flachheit der zugehörigen Fenster-Funktion im Ursprung, der Glattheit der Zielfunktion und der Ordnung der verwendeten Sobolev-Norm, in der der Rekonstruktionsfehler gemessen wird.

Des Weiteren beweisen wir Konvergenz der approximativen FBP-Rekonstruktion gegen die Zielfunktion in den betrachteten Sobolev-Normen, wenn die Bandbreite des Tiefpass-Filters gegen Unendlich strebt. Dabei ermitteln wir asymptotische Konvergenzraten in der Bandbreite und beobachten insbesondere Saturation der Konvergenzordnung bei fraktionalen Raten in Abhängigkeit von Glattheitseigenschaften der zum Filter gehörigen Fenster-Funktion.

Schließlich entwickeln wir Konvergenzraten auch für den Fall von Störungen in den Radon-Daten, wenn das Störungslevel gegen Null strebt. Dazu beweisen wir Fehlerabschätzungen für den Datenfehler und kombinieren diese mit unseren Resultaten für den Approximationsfehler. Weiterhin wird die Bandbreite des Tiefpass-Filters an das Störungslevel der Radon-Daten gekoppelt, um die gewünschte Konvergenz des Rekonstruktionsfehlers zu erzielen.

Die theoretischen Resultate werden gestützt durch numerische Experimente.

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den 31. Januar 2018