Understanding the Microlocal Properties of the Shearlet Transform via the Radon and the Wavelet Transforms

Master Thesis in Applied Mathematics
Delft University of Technology

Dani Rozenbroek

Thesis committee:
Dr. F. Bartolucci
Dr. ir. W.G.M. Groenevelt
Dr. H. N. Kekkonen

Summary

Over the last two decades shearlet analysis has played an important role in the field of microlocal analysis. One of the main advantages of the shearlet transform is its efficiency in capturing anisotropic data structures. This effectiveness can be explained by earlier research: in Bartolucci (2019) it is shown that there is a profound relation between the shearlet transform, the wavelet transform, and the Radon transform. This relation is established through an integral representation.

The first objective of this thesis is to examine the properties of the three transforms as described in the literature and to understand their roles in the resolution of the wavefront set within shearlet analysis. The findings of this literature review are as follows:

(i) The wavelet transform effectively captures isotropic data structures, but falls short in capturing anisotropic features. This limitation motivated the development of the shearlet system.

(ii) The Radon transform in combination with the Fourier transform can be used to determine whether a point is a regular directed point.

(iii) The shearlet transform is able to detect the wavefront set. This result was already shown in Grohs (2010), and has been partially established in a theorem in Bartolucci (2019) using the integral representation of the shearlet transform in terms of the Radon and wavelet transforms.

The result established in Bartolucci (2019), as described in (iii), characterizes *almost* the entire wavefront set, but not the *entire* wavefront set. Therefore, the second goal of this thesis is to generalize this result. Building on these foundations, we formulate and prove a more general theorem. This new theorem represents a new contribution to the existing literature and offers a deeper theoretical understanding of the shearlet transform's role in microlocal analysis through its connections with the Radon and wavelet transforms.

Preface

A mathematician is a machine for turning coffee into theorems.

Paul Erdős

This master's thesis is written to obtain the master's degree in Applied Mathematics. The last six months I explored the field of shearlet analysis and studied some of its marvelous results. I was astonished by the elegance of the deep connection between the shearlet, wavelet, and the Radon transforms and how they can be used in the study of microlocal analysis. It has been a true pleasure to gain these wonderful insights.

I would like to thank Francesca Bartolucci for introducing me this topic and for the guidance she provided during this project. Our weekly meeting where always quite helpful and it always gave me a better understanding. I really appreciate the time that she spent in our meetings and answering my numerous questions. The valuable feedback that she gave on previous versions of the manuscript improved this thesis significantly.

I would like to thank Wolter Groenevelt and Hanne Kekkonen for participating in my thesis committee and for the time that they spent in reading my thesis.

Finally, I would like to thank some friends for the pleasant study sessions and the many coffee breaks, which definitely helped me through this master's project.

Enjoy reading!

Dani Rozenbroek *Amsterdam, August 2025*

Contents

1
Introduction

Shearlets have been extremely popular in the field of applied harmonic analysis and over the last two decades their properties have been studied extensively. In the first section of this chapter, we will provide a motivation for our interest in shearlets through examples from image analysis. We will explain why shearlet improves other methods, such as wavelets, curvelets, and contourlets, and we will provide a short history of the introduction of shearlets in applied harmonic analysis. In Section 1.2 we describe the research goals of this thesis and in Section 1.3 we give an outline of its structure.

1.1 Applications of Shearlets

Applied harmonic analysis is a broad and fast-developing field in mathematics with numerous applications, many of which revolve around the efficient representation and reconstruction of 'signals' and the study of their fundamental properties. Examples of such 'signals' include audio recordings, digital images, and video sequences. Since these signals can be quite large and require extreme storage capacity, the challenge is to develop techniques that reduce the storage requirements in such a way that compresses the information effectively while preserving essential features.

Tools from applied harmonic analysis are extensively used in the field of image analysis, particularly in detecting and characterizing the geometric properties of images. It is often possible to 'reconstruct' the original image from its 'edges'. For example, consider the pictures in Figure 1.1. In Figure 1.1a we see a photo of a cat. By studying only the edges of this image, we can see that the original photo displays a cat. In Figure 1.1b we see an image of a church. Using the simple structure of the window (the gray shape), we see from its shape that it represents a window. In other words, we can reconstruct images from their edges, which represents an efficient approach to data storage and

compression while preserving the essential information. It saves a lot of memory to store only the edges, instead of the full image.

(a) A cat (b) A church window

Figure 1.1 Example of reconstructing an image from its edges. Source: Bartolucci (2019) and Kutyniok and Labate (2012).

In Figure 1.2, we see an image inside a box. This image consists of a smooth region separated by a smooth edge. These kinds of figures are so-called 'cartoon-like images'. In image analysis, these types of images are often modeled as a function $f : \mathbb{R}^2 \to \mathbb{C}$. Based on Figure 1.2, we would expect that this model consists of piecewise regular functions. This motivated the following terminology: the set of cartoon-like images consists of smooth functions, say they are twice continuously differentiable, which are of the form

$$f = f_0 + f_1 \chi_B,$$

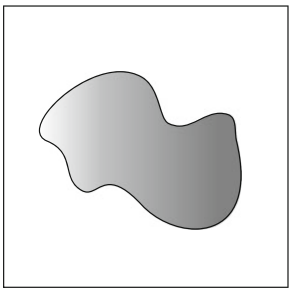


Figure 1.2: Example of a cartoon-like image. Source: Kutyniok and Labate (2012).

where B is a region in $\mathbb{R}^2$ with ∂B a closed curve which is smooth and has bounded curvature. From this observation, we see that we can model a cartoon-like image by a function consisting of a part that describes the interior of the region, and a function that describes the boundary. We see that the important parts of a photo are the edges, since they describe the whole picture. More precisely, the most important parts that describe and characterize a signal are exactly its sharp changes and unusual patterns.

In image science, we are often not only interested in the location of such edges but also in their direction. Microlocal analysis provides a framework for studying the behavior of distributions by examining not only the location of their singularities, but also the directions in which these singularities propagate. A fundamental concept in the theory of microlocal analysis is the wavefront set of a distribution. According to Hörmander

(1983), the wavefront set in $\mathbb{R}^2$ consists of all the points (x, ξ) such that x is an element of the boundary of a region and ξ is perpendicular to the boundary of that region at x. Thus, the wavefront set contains information about both the singular support of a distribution, that is the set of points where the function fails to be smooth, and the specific directions where singularities occur.

In many branches of mathematics, the wavefront set serves as a crucial concept. For example, in the study of partial differential equations, it plays an essential role in analyzing and understanding how singularities propagate through space and time, see Hintz (2025). It is also used in the study of medical imaging and tomography, see Deans (1983) and the references therein. Due to the many applications of the wavefront set, a lot of research has been done and is therefore well understood.

One of the first ideas to analyze singularities of signals was to use the Fourier transform. The Fourier transform has the ability to provide a description of the overall regularity of a signal. However, despite this global description, the Fourier transform has limitations when it comes to identifying the precise locations of singularities or detecting their spatial distribution of singularities.

This limitation of the Fourier transform inspired the development and study of wavelets. The 'wavelet era' began in the mid-late 1980s, and wavelet theory started gaining serious attention, which led to collaboration between mathematicians, physicists, and engineers, such as Yves Meyer, Stéphane Mallat, and Ingrid Daubechies. Consequently, new developments were made and one of the main objectives was to fix the limitation of the Fourier transform. This eventually led to the development of the wavelet transform. This transform is able to simultaneously detect spatial and frequency localization. This is due to the fact that wavelets have the property that they can characterize pointwise smoothness properties of functions. Therefore, the wavelet transform is able to characterize the local regularity of signals. For more details, we refer to the classical paper of Mallat and Hwang (1992).

Wavelets have truly revolutionized image and signal processing, and the theory of wavelets has been implemented in the algorithm of JPEG 2000, which is the current standard for image compression. Furthermore, wavelets have many applications in other fields such as signal and audio processing, time-frequency analysis, inverse problems and the numerical analysis of partial differential equations. See for example Daubechies (1992), and Gröchenig (2001) and the references therein.

Figure 1.3 shows a picture of two parrots (top left) and a version of the same picture with noise added (top right). The wavelet transform is able to capture geometric properties of this noisy image to achieve the most important features, for example, the edges of this picture (bottom left). Finally, a reconstruction is also shown, which is obtained using the shearlet transform (bottom right). Comparing the shearlet-based method and the wavelet-based method, we see that the shearlet-based method gives a better result, since we obtained more information about the original photo.

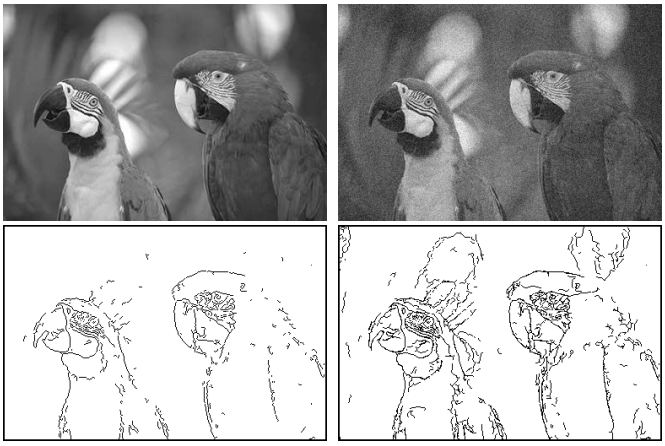
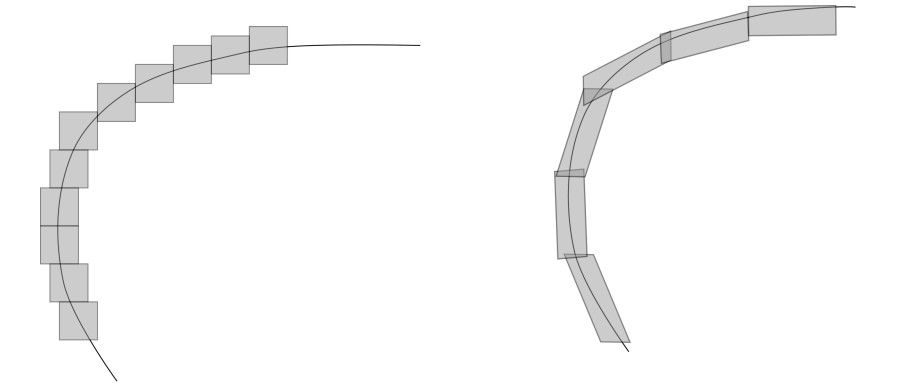


Figure 1.3 Comparison of edge detection using shearlet-based method versus wavelet-based method. From top left, clockwise: Original image, noisy image, shearlet result, and wavelet result. Source: Deng et al. (2009)

The reason why the wavelet-based method performs worse than the shearlet-based method is that the wavelet-based method exhibits limitations when applied to multivariate data structures. Wavelets achieve almost optimal representation efficiency for one-dimensional data containing pointwise singularities, but detecting singularities along curves (or manifolds within higher-dimensional spaces) is a more difficult task. This fundamental limitation arises from the inherently isotropic nature of the wavelet constructions, which employ uniform scaling parameters across all directions and cannot adapt to the geometric features of higher-dimensional data.

When we are dealing with images with edges in a two-dimensional setting, we are working with multivariate functions that are typically governed by anisotropic (i.e. directional) singularities.

In Figure 1.4, we see a curve which is covered with both isotropically shaped squares (Figure 1.4a) and anisotropically shaped and rotated rectangles (Figure 1.4b). From this figure, it is clear that it is suboptimal to describe isotropic data with wavelets to capture singularities along edges. Figure 1.4 indicates that it is more efficient to study function systems with varying types of localization, as well as a method for rotating the elements.



(a) Isotropically shaped squares overlapping a curve

(b) Anisotropically shaped rectangles overlapping a curve

Figure 1.4 Comparison of isotropically and anisotropically shaped regions along a curve.

The main reason why the shearlet-based method performed better than the wavelet-based method is that shearlets have an anisotropic structure. This means that they can handle different directions and shapes more effectively. Therefore, the shearlets are better for capturing geometric structures than wavelets. We hope that Figures 1.3 and 1.4 provide an informal motivation for why shearlets improve wavelets. In the next chapter, we will give a more mathematical argument for why we need shearlets to detect anisotropical data.

Due to the limitations of wavelets for such problems as described above, many researchers were motivated to look for alternative methods. In 1992, the steerable pyramid, and directional filter banks were introduced by Simoncelli et al. (1992) and Bamberger and Smith (1992), respectively. A year later, Antoine et al. (1993) introduced two-dimensional directional wavelets. Curvelets were introduced in Candès and Donoho (2004), and in 2005 Do and Vetterli introduced contourlets. We refer to Do and Vetterli (2005) for more details. Finally, in 2006 the shearlet system was introduced by the famous paper of Guo, Kanghui, Kutyniok, and Labate Guo et al. (2006). Shearlet systems satisfy all the desired properties that are required to handle an anisotropic system. We refer to Kutyniok and Labate (2012) and Labate et al. (2005) for a more detailed history of the emergence of the shearlet system.

In recent years, shearlets have been used in many branches of applied mathematics. For example, they are used to classify handwritten numbers, which is a classic machine learning problem, see Foroozandeh et al. (2020). Shearlets are also valuable in noise detection, and in reconstruction and inverse problems; we discussed these phenomena based on Figure 1.3.

1.2 Goal of this Thesis

In the existing literature a lot of research has been done on how and why shearlets can be helpful in identifying the wavefront set. In Grohs (2010) it is shown that the shearlet transform can successfully detect the wavefront set. The proof in that work uses several ideas from real analysis and it is quite technical.

In Bartolucci (2019) a different approach was taken to study the role of the shearlet transform in detecting the wavefront set. The author established an integral connection between the shearlet transform and the wavelet and Radon transforms. Based on this representation, a theorem was formulated and proved showing that the shearlet transform can determine whether a point belongs to the wavefront set. While this result was already known from earlier research, the integral formula provided new geometric insight into the shearlet transform's ability to resolve the wavefront set of signals.

However, the theorem in Bartolucci (2019) offers valuable new insights; it only characterizes *almost* the entire wavefront set, but not the *entire* wavefront set. This leaves a gap in the existing literature to fully describe the wavefront set via the shearlet transform and the established integral formula. The main goal of this thesis is to generalize this theorem to achieve a complete characterization of the wavefront set using the shearlet transform.

To formulate this theorem and provide a proof, we need to examine the literature and study the Radon and wavelet transforms. We will see that these transforms play an important role in our theoretical framework, since exactly these two transforms explain why the shearlet transform is suitable for detecting the wavefront set. Therefore, the first goal of this thesis is to provide the necessary mathematical background on shearlets, wavelets, and the Radon transform.

With this theoretical background we can formulate and provide a proof of our new theorem. This theorem will represent a new contribution to the existing literature and offers a deeper theoretical understanding of the Shearlet transform's role in microlocal analysis through its connections with the Radon and wavelet transforms.

1.3 Thesis Outline

The structure of this thesis is as follows. We start in Chapter 2 with introducing some notation which will be used throughout this thesis while recalling relevant theory from Fourier analysis. Furthermore, we will devote a few sections to the exploration of the properties of the wavelet, the Radon, and the shearlet transforms, which are all well-known transformations in the literature. Furthermore we will give a mathematical definition of the wavefront set and we will provide an example that the wavelet transform is not always able to describe the wavefront set in higher dimensions.

After laying this foundation, we will study the role of the shearlet transform in microlocal analysis in Chapter 3. In this chapter, we will study the wavefront set and see how and why the shearlet transform can detect singularities. This will be done by providing a motivating example in the first section. Based on this example, we will formulate in the second section our main theorem and explain the new contribution of our work to the literature. The proof of this theorem is outlined in the subsequent two sections.

Finally, in Chapter 4 we give a conclusion of our work and provide an overview of some open problems for follow-up research.

2
Preliminaries

In this chapter we introduce the main theoretical concepts and known results from the literature that will be used throughout this thesis. In Section 2.1, we introduce some notation and define the function spaces that will be used in this thesis. In Section 2.2 we recall some fundamental properties of the Fourier transform. Section 2.3 is devoted to studying the wavefront set and providing some motivating examples. In Section 2.4 we discuss the Radon transform and study some important properties of it. Subsequently, we will introduce the theory of wavelets in Section 2.5 and explain why (higher-dimensional) wavelets are not suitable to describe anisotropic phenomena. This limitation leads to the theory of shearlet systems, which will be discussed in Section 2.6. Finally, in the last section of this chapter, we mention an interesting integral identity that connects the Radon, the wavelet, and the (vertical) shearlet transforms.

2.1 Notation and Function Spaces

In this section we introduce some notations and definitions from real analysis. Furthermore, we introduce the function spaces that we will use throughout this thesis. We refer to classical books on this topic, such as Stein and Shakarchi (2005) and Grafakos (2014), for more details and motivation.

We begin by introducing notation for several variables. While our theory primarily applies to one- and two-dimensional settings, we present the notation in a general d-dimensional context.

A function f is said to be in $L^p(\mathbb{R}^d)$ if

$$\int_{\mathbb{R}^d} |f(x)|^p \mathrm{d}\mu(x) < \infty.$$

Here, the measure $\mathrm{d}\mu$ corresponds to the usual Lebesgue measure on the corresponding σ-algebra $\mathscr{B}(\mathbb{R}^d)$. The L^p-norm is denoted by

$$\|f\|^p_{L^p(\mathbb{R}^d)} := \int_{\mathbb{R}^d} |f(x)|^p \mathrm{d}\mu(x).$$

Hölder's inequality states that for nonzero measurable functions f and g on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ we have

$$\|fg\|_{L^1(\mathbb{R}^d)} \leqslant \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)},$$

where q is the so-called dual exponent of p and satisfies the relation $p^{-1} + q^{-1} = 1$. When we are in the setting of a Hilbert space, we can define an inner product, which is given by

$$\langle f, g \rangle := \int_{\mathbb{R}^d} f(x)\overline{g(x)} \mathrm{d}\mu(x).$$

Let $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. The norm of x is defined by $|x| := \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}$. A multi-index α is an ordered d-tuple of natural numbers, i.e. $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ and we set $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$, which denotes its total size. We write $\partial^\alpha f$ for $\partial_1^{\alpha_1} \partial_2^{\alpha_2} \ldots \partial_d^{\alpha_d} f$. Furthermore, we denote the space of all infinitely differentiable compactly supported functions by $\mathscr{C}_0^\infty(\mathbb{R}^d)$ and the space of all infinitely differentiable functions by $\mathscr{C}^\infty(\mathbb{R}^d)$.

Let $\mathscr{S}(\mathbb{R}^d)$ be the Schwartz space, which roughly consists of rapidly decaying functions, i.e. all the functions in this space are smooth and all of their derives decay faster than any polynomial. More precisely: a function $f \in \mathscr{C}^\infty(\mathbb{R}^d)$ is a Schwartz function if for all multi-indices α en β we have

$$\rho_{\alpha,\beta}(f) = \sup_{x \in \mathbb{R}^d} |x^\alpha \partial^\beta f(x)| < \infty.$$

Note that a function f is in $\mathscr{S}(\mathbb{R}^d)$ if and only of for every multi-index β and every $N \in \mathbb{N}$ there is a constant $C_{N,\beta}$ such that for all $x \in \mathbb{R}^d$ we have

$$|\partial^\beta f(x)| \leqslant \frac{C_{N,\beta}}{(1+|x|)^N}. \tag{2.1}$$

Recall that the dual space is the space of continuous linear functionals on the set of test functions. We introduce the following spaces:

$$(\mathscr{C}_0^\infty(\mathbb{R}^d))' = \mathscr{D}'(\mathbb{R}^d),$$
$$(\mathscr{S}(\mathbb{R}^d))' = \mathscr{S}'(\mathbb{R}^d),$$
$$(\mathscr{C}^\infty(\mathbb{R}^d))' = \mathscr{E}'(\mathbb{R}^d).$$

From these definitions it is clear that the dual spaces are nested as follows:

$$\mathscr{E}'(\mathbb{R}^d) \subseteqq \mathscr{S}'(\mathbb{R}^d) \subseteqq \mathscr{D}'(\mathbb{R}^d).$$

The elements of the space $\mathscr{D}'(\mathbb{R}^d)$ are called *distributions*. Elements of $\mathscr{S}'(\mathbb{R}^d)$ are called *tempered distributions*. Finally, the elements of the space $\mathscr{E}'(\mathbb{R}^d)$ are called *distributions with compact support*.

Since the space $\mathscr{S}'(\mathbb{R}^d)$ is defined to be the topological dual space of $\mathscr{S}(\mathbb{R}^d)$, we have

$$\mathscr{S}'(\mathbb{R}^d) := \{u : \mathscr{S}(\mathbb{R}^d) \to \mathbb{C} : u \text{ is linear and continuous}\}.$$

From this definition, it is clear that a linear functional $u : \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ is a tempered distribution if and only if there exist constants $C > 0$ and $M, K \in \mathbb{N}$ such that for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$ we have

$$|\langle u, \varphi \rangle| \leqslant C \sum_{|a| \leqslant M} \sum_{|\beta| \leqslant K} \rho_{\alpha,\beta}(\varphi). \tag{2.2}$$

We define the space $\mathscr{S}_0(\mathbb{R}^d)$ by all the functions f with vanishing moments, i.e.

$$\mathscr{S}_0(\mathbb{R}^d) := \left\{ \int_{\mathbb{R}^d} x^n f(x) \mathrm{d}x < \infty : \forall n \in \mathbb{N}^d \right\}.$$

Furthermore, we write $f(x) = \mathscr{O}(g(x))$ if and only if

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = C, \quad C \in \mathbb{R},$$

and we write $f(x) \lesssim g(x)$ if there exists a positive constant C such that for every x we have $f(x) \leqslant Cg(x)$. Finally, $\mathbb{R}^+$ denotes the positive real numbers.

2.2 The Fourier Transform

In this section we will give a brief recap of the theory of Fourier analysis and recall some well-known results from this field. We start by introducing the definition of the Fourier transform.

Definition 2.1 Let $f \in \mathscr{S}(\mathbb{R}^d)$. We define the *Fourier transform* of f by

$$\mathscr{F}(f)(\xi) := \widehat{f}(\xi) := \int_{\mathbb{R}^d} f(x) e^{-2\pi i x\cdot\xi} \mathrm{d}x.$$

Definition 2.1 and a change of variables imply that for any $\lambda \in \mathbb{R} \setminus \{0\}$ we have

$$\widehat{f}(\lambda\xi) = \frac{1}{\lambda^d} \widehat{f}\left(\frac{\xi}{\lambda}\right). \tag{2.3}$$

This 'dilation' property will be used in many subsequent sections of this thesis.

We recall some well-known properties of the Fourier transform.

Proposition 2.2 *Let $f, g \in \mathscr{S}(\mathbb{R}^d)$. The following properties hold:*

(i) *The Fourier transform is uniformly continuous on $\mathbb{R}^d$ and*

$$\|\widehat{f}\|_{L^\infty(\mathbb{R}^d)} \leqslant \|f\|_{L^1(\mathbb{R}^d)}.$$

(ii) *Parseval's relation:*

$$\langle f, g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\mathrm{d}\xi = \langle \widehat{f}, \widehat{g}\rangle.$$

(iii) *Plancherel's identity:*

$$\|f\|_{L^2(\mathbb{R}^d)} = \|\widehat{f}\|_{L^2(\mathbb{R}^d)}.$$

Remark 2.3 It is obvious that Definition 2.1 and Formula (2.3) make sense as a convergent integral for functions $f \in L^1(\mathbb{R}^d)$. We can thus extend this definition of the Fourier transform on $L^1(\mathbb{R}^d)$. Furthermore, property (ii) makes also sense whenever f and g are in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. With a slight abuse of notation we denote the Fourier transform of $f \in L^2(\mathbb{R}^d)$ by $\widehat{f}$. Then, we can extend Plancherel's identity to functions belonging to the space $L^2(\mathbb{R}^d)$.

If $u : \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ is a tempered distribution, we can take its (tempered) Fourier transform, which is defined in the following definition.

Definition 2.4 For $u \in \mathscr{S}'(\mathbb{R}^d)$ we define the Fourier transform $\widehat{u}$ of a tempered distribution u by the identity

$$\langle \widehat{u}, \varphi\rangle = \langle u, \widehat{\varphi}\rangle$$

for all functions $\varphi \in \mathscr{S}(\mathbb{R}^d)$.

The following two examples will be useful in this thesis.

Example 2.5 The *Dirac mass* δ_t at a point $t \in \mathbb{R}^d$ is defined by

$$\langle \delta_t, \varphi \rangle = \varphi(t)$$

for every $\varphi \in \mathscr{S}(\mathbb{R}^d)$. We have $\delta_t \in \mathscr{S}'(\mathbb{R}^d)$, since

$$|\langle \delta_t, \varphi \rangle| = |\varphi(t)| \leqslant \|\varphi\|_{L^\infty(\mathbb{R}^d)} = \rho_{0,0}(\varphi). \tag{2.4}$$

Therefore, by (2.2), we conclude that $\delta_t \in \mathscr{S}'(\mathbb{R}^d)$ and hence we can compute its Fourier transform, which results in

$$\langle \widehat{\delta}_t, \varphi \rangle = \langle \delta_t, \widehat{\varphi} \rangle = \widehat{\varphi}(t) = \int_{\mathbb{R}^d} \varphi(x) e^{-2\pi i x \cdot t} \, \mathrm{d}x, \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^d).$$

Hence, $\widehat{\delta}_t$ can be identified with the function $\xi \mapsto e^{-2\pi i \xi \cdot t}$.

Example 2.6 Consider the distribution $\delta_{x_2 = p + q x_1}$. Similarly, we can show that this is a tempered distribution and therefore we can compute its Fourier transform. For any φ in the Schwartz space, we have:

$$\begin{aligned}
\langle \widehat{\delta_{x_2 = p + q x_1}}, \varphi \rangle &= \langle \delta_{x_2 = p + q x_1}, \widehat{\varphi} \rangle \\
&= \int_{-\infty}^{\infty} \widehat{\varphi}(x_1, p + q x_1) \, \mathrm{d}x_1 \\
&= \int_{-\infty}^{\infty} \int_{\mathbb{R}^2} \varphi(y_1, y_2) e^{-2\pi i (x_1 (y_1 + q y_2) + p y_2)} \, \mathrm{d}y_1 \, \mathrm{d}y_2 \, \mathrm{d}x_1 \\
&= \int_{\mathbb{R}^2} \int_{-\infty}^{\infty} \varphi(y_1, y_2) e^{-2\pi i (x_1 (y_1 + q y_2) + p y_2)} \, \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}y_2 \qquad \text{(Fubini)} \\
&= \int_{\mathbb{R}^2} \varphi(y_1, y_2) \delta(y_1 + q y_2) e^{-2\pi i p y_2} \, \mathrm{d}y_1 \, \mathrm{d}y_2.
\end{aligned}$$

Therefore $\widehat{\delta_{x_2 = p + q x_1}}$ can be identified with the function $(\xi_1, \xi_2) \mapsto e^{-2\pi i \xi_2 \cdot p} \delta(\xi_1 + q \xi_2)$.

2.3 The Wavefront Set

In this section, we will give a formal definition of the wavefront set. We start with introducing the definitions, and thereafter we will provide some motivating examples. Our approach is quite influenced by Kutyniok and Labate (2012).

Getting to know the Wavefront Set

By microlocal analysis we mean the mathematical study of singular points of distributions. In microlocal analysis we are not only interested in the location of the singular points, but also in their direction. An important concept in this field is the wavefront set.

The wavefront set is, roughly speaking, the set of singular points together with its direction, i.e. the direction that is perpendicular to the tangent of the curves. So the wavefront set consists of points and directions.

Figure 2.1 shows a discontinuity curve together with some vectors. The black vectors are perpendicular to the discontinuity curve. So the wavefront set consists of all the points on the discontinuity curve together with the vectors that are perpendicular.

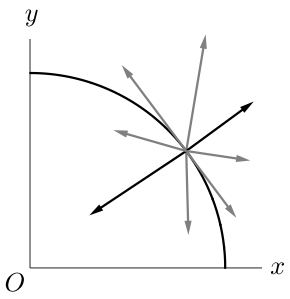


Figure 2.1: Example of the wavefront set.

To motivate the more formal definition of the wavefront set, we recall first the classical Paley-Wiener theorem, see Rudin (1991). This theorem relates the Fourier transform of a compactly supported distribution u to its smoothness and decay properties. Specifically, u is a smooth compactly supported function if and only if its Fourier transform $\mathscr{F}(u)(\xi)$ decays faster than any polynomial as $|\xi| \to \infty$. In other words, the Paley-Wiener theorem characterizes smooth compactly supported functions by a growth condition on their Fourier transform.

Theorem 2.7 (Paley-Wiener) *Assume u is in $\mathscr{E}'(\mathbb{R}^2)$. Then u is in $\mathscr{C}_0^\infty(\mathbb{R}^2)$ if and only if for every $N \in \mathbb{N}$ there exists a constant C_N such that*

$$|\mathscr{F}(u)(\xi)| \leqslant C_N(1+|\xi|)^{-N} \tag{2.5}$$

for every $\xi \in \mathbb{R}^2$.

So the Paley-Wiener theorem implies that if u is in $\mathscr{E}'(\mathbb{R}^2)$, but not in $\mathscr{C}^\infty(\mathbb{R}^2)$, then there exists at least one direction $\xi \neq 0$ such that the Fourier transform of u does not satisfy (2.5) in any neighborhood containing ξ. Heuristically speaking, the direction ξ causes the problem of why u is not smooth.

We introduce the following terminology.

Definition 2.8 Let $f \in \mathscr{D}'(\mathbb{R}^2)$.

(i) A point $x_0 \in \mathbb{R}^2$ is called a *regular point* if there exist a neighborhood U_{x_0} of x_0 and a function $\varphi \in \mathscr{C}_0^\infty(U_{x_0})$, with $\varphi(x_0) \neq 0$, such that $\varphi f \in \mathscr{C}_0^\infty(\mathbb{R}^2)$.

(ii) The *singular support* of f is the complement of the set of regular points. This set is denoted by $\operatorname{sing\,supp}(f)$.

This motivates why we are interested in studying 'regular directed points', since these points enable us to analyze the local behavior of $f \in \mathscr{D}'(\mathbb{R}^2)$ in a neighborhood of a given point $x_0 \in \mathbb{R}^2$ and the decay behavior of its Fourier transform in a certain direction $\xi_0 \in \mathbb{R}^2 \setminus \{0\}$.

Definition 2.9 Assume $f \in \mathscr{D}'(\mathbb{R}^2)$.

(i) A point $(x_0, v_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is called a *regular directed point* of f if there exist a neighborhood U_{x_0} of x_0, a function $\varphi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$, such that $\varphi(x_0) \neq 0$, and a neighborhood V_{v_0} of v_0 such that for every $N \in \mathbb{N}$ we have

$$(\widehat{\varphi f})(\xi) = \mathscr{O}\left((1+|\xi|)^{-N}\right) \tag{2.6}$$

for every $\xi = (\xi_1, \xi_2)$ with $\xi_2/\xi_1 \in V_{v_0}$.

(ii) The complement of the set of regular directed points is called the *wavefront set* and is denoted by $\mathrm{WF}(f)$.

Some Examples

We will give some examples to get a feeling for this terminology.

Example 2.10 Let $\delta_t : \mathscr{S}'(\mathbb{R}) \to \mathbb{C}$ be the Dirac distribution defined as in Example 2.5. We claim that

$$\mathrm{WF}(\delta_t) = \{t\} \times \mathbb{R}.$$

Proof Choose an arbitrary point t' in $\mathrm{WF}(\delta_t)$. If $t' \neq t$ then for any smooth cutoff function φ we have $\varphi\delta_t = 0$ in a neighborhood of t'. Consequently, the Fourier transform of $t \mapsto \varphi\delta_t$ equals zero, i.e. we have $\widehat{(\varphi\delta_t)} = 0$, and thus it decays rapidly in ξ. This means that (t', ξ) cannot belong to the wavefront set for any direction ξ. Consider now the case $t' = t$. We know that δ_t is singular at the point t and from Example 2.5 we know that $\widehat{\delta_t}(\xi) = \exp(2\pi i \xi \cdot t)$. This does not decay fast in any direction $\xi \in \mathbb{R}$. Therefore we obtained $\mathrm{WF}(\delta_t) = \{t\} \times \mathbb{R}$. □

Example 2.11 Define the 'line distribution' $\delta_{x_2=p+qx_1}$ by

$$\langle \delta_{x_2=p+qx_1}, \varphi \rangle := \int_{-\infty}^{\infty} \varphi(x_1, p+qx_1)\mathrm{d}x_1, \quad \varphi \in \mathscr{S}(\mathbb{R}).$$

We claim that

$$\mathrm{WF}(\delta_{x_2=p+qx_1}) = \{(x_1, x_2) : x_2 = p+qx_2\} \times \{-1/q\}.$$

Proof From Example 2.6 we know that

$$\widehat{\delta_{x_2=p+qx_1}}(\xi) = e^{-2\pi i \xi_2 \cdot p}\delta(\xi_1 + q\xi_2)$$

in the sense of tempered distributions. Thus $\widehat{\delta_{x_2=p+qx_1}}$ is of fast decay, except when $\xi_2/\xi_1 = -q$. A similar argument as given in the previous example, yields the desired claim. □

Example 2.12 Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x,y) = \begin{cases} 1 & \text{if } y > 0, \\ 0 & \text{if } y \leqslant 0. \end{cases}$$

A graph of f is shown below in Figure 2.2.

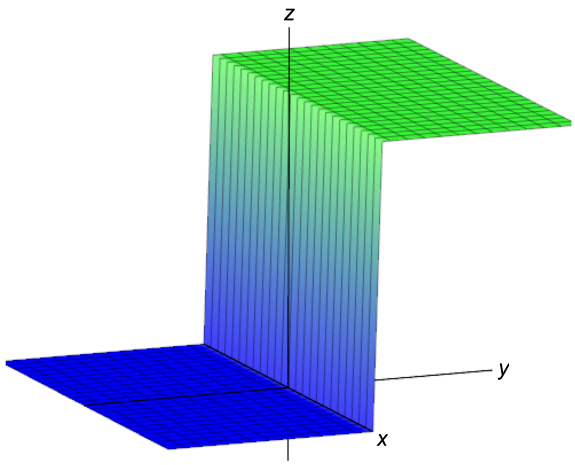


Figure 2.2 The graph of f

From the graph of f, we see clearly that f has a singularity along the axis $(x,0)$, with $x \in \mathbb{R}$, and in the direction perpendicular to the (x,y) plane, i.e. in the direction $(0,\xi_2)$, with $\xi_2 \neq 0$. Therefore, we would expect that

$$\mathrm{WF}(f) = \{(x,0,0,\xi_2) : x \in \mathbb{R}, \xi_2 \neq 0\}.$$

This intuition can be made precise, see Boman (2014) for a proof of this result.

From the three discussed examples above, it was from the beginning 'clear' where the regular directed points are located. If the function f becomes more difficult, then it is more difficult to find and characterize these points. We are therefore interested in some mathematical theory to locate the regular directed points; which will be done in the following sections. In the next section we will introduce the Radon transform and show that the Radon transform in combination with the Fourier transform is able to detect the regular directed points of f.

2.4 The Radon Transform

The Radon transform was introduced by Johann Radon in 1917, and since then this transformation has had significant achievements in applications such as medicine (such as tomography), geophysics, astronomy, and optics. We refer to Deans (1983) for details and for more applications. Our approach in this section is inspired by Deans (1983). Furthermore, we refer to the classical reference Ramm and Katsevich (1996) for more details about the Radon transform.

We introduce the Radon transform in the two-dimensional setting, since the applications of this thesis are in $\mathbb{R}^2$. Moreover, the two-dimensional case provides the most accessible way to motivate the main ideas and develop an intuitive understanding of the definition. It is worth mentioning that our results and definitions can be generalized to the setting of $\mathbb{R}^d$.

We will see that the Radon transform can be obtained from different parametrizations. Therefore we will study the following integral transformations: the 'hyperplane' Radon transform, the polar Radon transform, and the (vertical) affine Radon transform. When we refer to the 'Radon transform', it will be clear from the context which one we mean.

The Radon Transform

Let f be defined on a domain $D \subseteq \mathbb{R}^2$. The line integral of f along all possible lines L defines the Radon transform, denoted by $\mathscr{R}f$. See Figure 2.3.

Let $x = (x,y)$. We can write the line L by $n \cdot x = n_1 x + n_2 y = t$. So in $\mathbb{R}^2$ we write

$$\mathscr{R}f(n,t) = \frac{1}{|n|} \int_{n\cdot x = t} f(x) \mathrm{d}m(x),$$

where $\mathrm{d}m(x)$ is the Euclidean measure on the line $\{x \in \mathbb{R}^2 : n \cdot x = t\}$. In this formula we divide by $|n|$ for convenience. The variables n and t can respectively be thought of as the 'direction' and 'distance' from the origin.

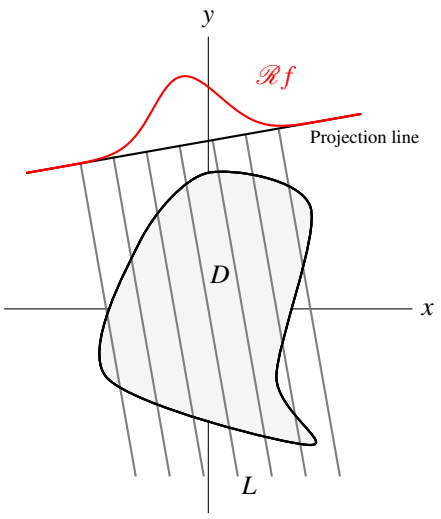


Figure 2.3: Integration along lines

This leads to the following general definition.

Definition 2.13 Let $f \in L^1(\mathbb{R}^2)$. The *Radon transform* $\mathscr{R}f : (\mathbb{R}^2 \setminus \{0\}) \times \mathbb{R} \to \mathbb{C}$ is defined by

$$\mathscr{R}f(n,t) = \frac{1}{|n|} \int_{n \cdot x = t} f(x) \mathrm{d}m(x),$$

where $\mathrm{d}m(x)$ is the Euclidean measure on the hyperline $\{x \in \mathbb{R}^2 : n \cdot x = t\}$.

We make some remarks.

Remark 2.14 (i) First, observe that $\mathscr{R}f(n,t)$ is well-defined. Indeed, fix an $n \in \mathbb{R}^2 \setminus \{0\}$. Then Tonelli's theorem yields:

$$\|\mathscr{R}f\|_{L^1(\mathbb{R}^2)} \leqslant \int_{-\infty}^{\infty} \frac{1}{|n|} \left(\int_{n \cdot x = t} |f(x)| \mathrm{d}m(x) \right) \mathrm{d}t = \int_{\mathbb{R}^2} |f(x)| \mathrm{d}x < \infty,$$

for almost every $t \in \mathbb{R}$.

(ii) The parametrization of a line is not unique. Consider for example the line $x = 1$. Possible parametrization of this line are $((1,0),1)$ and $((2,0),2)$. The latter reduces to $2x = 2$, i.e. $x = 1$.

A very important connection between the Radon transform and the Fourier transform is summarized in the theorem below, which is known in the literature as 'The Fourier Slice Theorem'.

Theorem 2.15 (The Fourier Slice Theorem) *Let $f \in L^1(\mathbb{R}^2)$. Then for all $n \in \mathbb{R}^2 \setminus \{0\}$ and all $\tau \in \mathbb{R}$ we have that*

$$\mathscr{F}(\mathscr{R}f(n,\cdot))(\tau) = \mathscr{F}f(\tau n).$$

Note that the Fourier transform on the right-hand side denotes the two-dimensional Fourier transform, while $\mathscr{F}(\mathscr{R}f(n,\cdot))(\tau)$ denotes the one-dimensional Fourier transform of $\mathscr{R}f(n,t)$ as a function of t, with n fixed.

To give a rigorous proof of the Fourier Slice Theorem, we need Coarea's formula.

Theorem 2.16 (Coarea's formula) *Let $u : \mathbb{R}^2 \to \mathbb{R}$ be Lipschitz function and let $f \in L^1(\mathbb{R}^2)$. Then*

$$\int_{\mathbb{R}^2} f(x) |\nabla u(x)| \, \mathrm{d}x = \int_{-\infty}^{+\infty} \left(\int_{\{u^{-1}(t)\}} f(x) \, \mathrm{d}S(x) \right) \mathrm{d}t.$$

The Coarea Formula can be seen as a kind of 'curvilinear' generalized version of Fubini's Theorem. The proof of this theorem is beyond the scope of this thesis and therefore we refer to Evans and Gariepy (2015).

Coarea's formula can be used to prove the Fourier slice theorem, which will be done below.

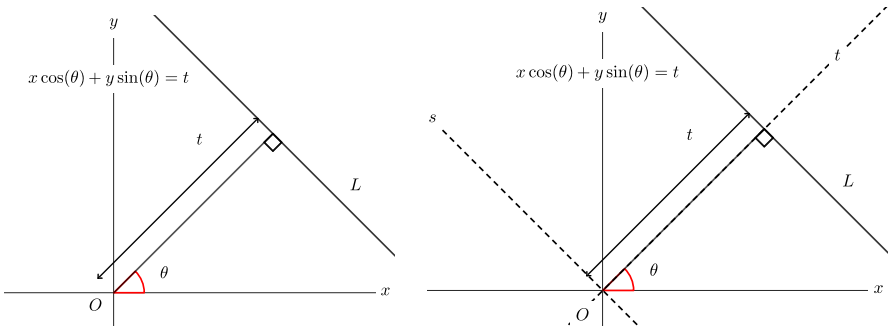
Proof of the Fourier Slice Theorem Let $f \in L^1(\mathbb{R}^2)$. Then for any $n \in \mathbb{R}^2 \setminus \{0\}$ and $\tau \in \mathbb{R}$ fixed we have

$$\begin{aligned}
\mathscr{F}(\mathscr{R}f(n,\cdot))(\tau) &= \int_{-\infty}^{\infty} \mathscr{R}f(n,t)e^{-2\pi it\cdot\tau}\mathrm{d}t \\
&= \int_{-\infty}^{\infty} \frac{1}{|n|}\int_{x\cdot n=t} f(x)\mathrm{d}m(x)\mathrm{e}^{-2\pi it\cdot\tau}\mathrm{d}t \\
&= \int_{\mathbb{R}^2} f(x)\mathrm{e}^{2\pi i(n\tau)\cdot x}\mathrm{d}x \qquad \text{(Coarea's Formula)} \\
&= \mathscr{F}f(n\tau).
\end{aligned}$$

In the third line we used Coarea's formula. Here we used the function $u : \mathbb{R}^2 \to \mathbb{R}$ defined by $u(x) = x\cdot n$, with $u^{-1}(t) = \{x\cdot n = t\}$ and $|\nabla u(x)| = |n|$. □

The Polar Radon Transform

We start with a motivation of our choice of the parameterization of the polar Radon transform. We parametrize the line by the pair (θ,t), where θ is the direction perpendicular to the line L and $|t|$ its distance from the origin. Then a parametrization of L is $x\cos(\theta)+y\sin(\theta)=t$. See Figure 2.4a.



(a) In the (x,y)-plane (b) In the (t,s)-plane

Figure 2.4 A parametrization of a line using polar coordinates.

Now, suppose that we introduce a new coordinate system with axis rotated by the angle θ. If the new axis are labeled by t and s, see Figure 2.4b, then we get

$$x = t\cos(\theta) - s\sin(\theta), \tag{2.7}$$
$$y = t\sin(\theta) + s\cos(\theta). \tag{2.8}$$

This yields the following explicit formula for the polar Radon transform:

$$\mathscr{R}^{\text{pol}} f(\theta,t) = \int_{-\infty}^{\infty} f(t\cos\theta - s\sin\theta, t\sin\theta + s\cos\theta)\mathrm{d}s.$$

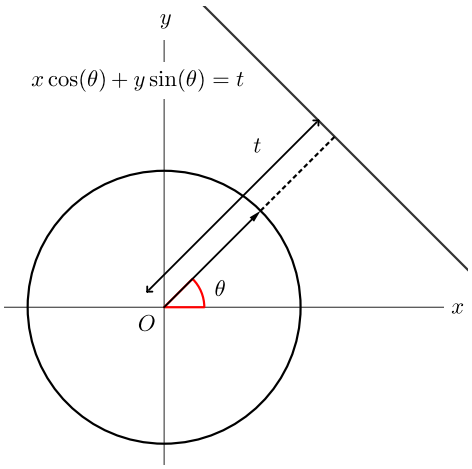


Figure 2.5 Parametrization of the line by the pair $(\theta,t) \in [0,2\pi) \times \mathbb{R}$, where θ is the direction which is perpendicular to the line and t the distance.

Definition 2.17 Let $f \in L^1(\mathbb{R}^2)$. The polar Radon transform of f is the function $\mathscr{R}^{\text{pol}} f : [0,2\pi) \times \mathbb{R} \to \mathbb{C}$ defined by

$$\mathscr{R}^{\text{pol}} f(\theta,t) = \mathscr{R} f(n(\theta),t) = \int_{n(\theta)\cdot x = t} f(x)\,\mathrm{d}m(x). \tag{2.9}$$

The reason why (2.9) is called the polar Radon transform can be seen from Figure 2.5.

Remark 2.18 Since we restrict ourself to the interval $\theta \in [0,2\pi)$, the parametrization of the line is unique.

We will illustrate the theory with an example. This example will be used later on in the thesis.

Example 2.19 Consider the function $f : \mathbb{R}^2 \to [0, \frac{1}{2}]$ defined by

$$f(x,y) := \frac{1}{2}\begin{cases} 1 & \text{if } x^2+y^2 \leqslant 1, \\ 0 & \text{if } x^2+y^2 > 1. \end{cases}$$

Since we are on the unit circle, it is convenient to use polar coordinates. Using the Formulas (2.7) and (2.8), the condition $x^2+y^2 \leqslant 1$ is equivalent with

$$-\sqrt{1-t^2} \leqslant s \leqslant \sqrt{1-t^2}.$$

Since f equals $1/2$ inside the unit disk, the integral formula (2.9) reduces to

$$\mathscr{R}^{\text{pol}} f(\theta,t) = \frac{1}{2}\int_{-\sqrt{1-t^2}}^{\sqrt{1-t^2}} \mathrm{d}s = \sqrt{1-t^2}, \text{ for } |t| \leqslant 1.$$

For $|t| > 1$ the line does not intersect the disk, so the Radon transform is zero. So we conclude:

$$\mathscr{R}^{\text{pol}} f(\theta,t) = \begin{cases} \sqrt{1-t^2}, & |t| \leqslant 1, \\ 0, & |t| > 1. \end{cases}$$

The (Vertical) Affine Radon Transform

The third way to parametrize line by a par (v,t) is by using the vector $n(v) = (1,v)$, which is perpendicular to the line intersecting the x-axis at $(t,0)$. See figure 2.6.

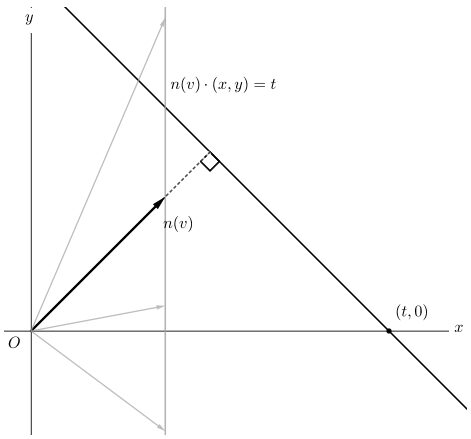


Figure 2.6 Affine Radon Transform

Since the points (x,y) lie on the line $n(v)\cdot(x,y) = x+yv = t$, we have for any $y \in \mathbb{R}$ that x is determined by the relation $x = t - vy$. So if we want to integrate over the line, this is

equivalent with integrating over all possible y-values (see the gray arrows and the gray line in Figure 2.6), which yields

$$\mathscr{R}^{\text{aff}} f(v,t) = \int_{-\infty}^{\infty} f(t-vy,y)\,\mathrm{d}y.$$

These ideas lead to the following definition:

Definition 2.20 Let $f \in L^1(\mathbb{R}^2)$. The *affine Radon transform* of f is the function $\mathscr{R}^{\text{aff}} f : \mathbb{R}\times\mathbb{R} \to \mathbb{C}$ given by

$$\mathscr{R}^{\text{aff}} f(v,t) = \mathscr{R} f(n(v),t)s = \frac{1}{\sqrt{1+|v|^2}} \int_{n(v)\cdot x = t} f(x)\,\mathrm{d}m(x) = \int_{\mathbb{R}} f(t-vy,y)\,\mathrm{d}y.$$

The affine Radon transform $\mathscr{R}^{\text{aff}} f(v,t)$ and the polar Radon transform are related as follows

$$\mathscr{R}^{\text{aff}} f(v,t) = \frac{1}{\sqrt{1+v^2}} \mathscr{R}^{\text{pol}} f\left(\theta_v, \frac{t}{\sqrt{1+v^2}}\right), \quad \theta_v = \arctan(v).$$

By construction, the affine Radon transform is defined by labeling the normal vector $n(v) = (1,v)$ to a line by affine coordinates. This is always possible, except for the horizontal lines. This construction is a limitation, since we cannot 'cover' the whole $\mathbb{R}^2$ plane. This shortage motivates the introduction of the vertical affine Radon transform, which will be done now.

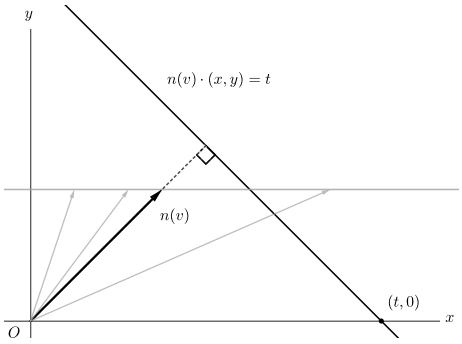


Figure 2.7: Parametrising along horizontal lines, which is obtained by switching the roles of the x-axis and the y-axis in the parametrization of the affine Radon transform.

In this new setting we parametrize lines in $\mathbb{R}^2$ by pairs $(v,t) \in \mathbb{R}^2$ as $\{(x,y) \in \mathbb{R}^2 : vx + y = t\}$. Now we parametrize over all the lines, except the vertical ones. See Figure 2.7 for an illustration. We will denote the unit vector n by $n(v) = (v,1)$, which will be the most convenient way to work with.

This leads to the following definition.

Definition 2.21 The vertical affine Radon transform of any $f \in L^1(\mathbb{R}^2)$ is the function $\mathscr{R}^{\text{v}} f : \mathbb{R}^2 \to \mathbb{C}$ defined by the formula

$$\mathscr{R}^{\text{v}} f(v,t) := \int_{-\infty}^{\infty} f(x,t-vx)\mathrm{d}x.$$

We make two observations.

Remark 2.22

(i) Let $\tilde{f}$ be defined as $\tilde{f}(x,y) := f(y,x)$. Then the following relation holds

$$\mathscr{R}^{\text{aff}} f(v,t) = \mathscr{R}^{\text{v}} \tilde{f}(v,t).$$

(ii) For the vertical affine Radon transform the Fourier slice theorem reads:

$$\mathscr{F}(\mathscr{R}^{\text{v}} f(v,t))(\xi) = \mathscr{F}(f)((\xi v, \xi)). \tag{2.10}$$

We give an example of an explicit calculation of the vertical affine Radon transform of a function. This example will be used in the next chapter.

Example 2.23 Consider the function $f : \mathbb{R}^2 \to [0, \frac{1}{2}]$ defined by

$$f(x,y) := \frac{1}{2} \begin{cases} 1 & \text{if } x^2+y^2 \leqslant 1, \\ 0 & \text{if } x^2+y^2 > 1. \end{cases}$$

We will compute the vertical affine Radon transform. By definition, we have

$$\mathscr{R}^{\text{v}} f(v,t) = \int_{-\infty}^{\infty} f(x, t-vx)\mathrm{d}x = \frac{1}{2} \int_{-\infty}^{\infty} \mathbf{1}_{\{x^2+y^2 \leqslant 1\}}(x, t-vx)\mathrm{d}x.$$

This integral in nonzero whenever $x^2 + (t-vx)^2 \leqslant 1$, i.e. whenever

$$\frac{tv - \sqrt{1-t^2+v^2}}{1+v^2} \leqslant x \leqslant \frac{tv + \sqrt{1-t^2+v^2}}{1+v^2}.$$

This yields

$$\mathscr{R}^{\text{v}} f(v,t) = \begin{cases} \dfrac{\sqrt{1+v^2-t^2}}{1+v^2} & \text{if } t^2 - v^2 \leqslant 1, \\ 0 & \text{if } t^2 - v^2 > 1. \end{cases}$$

Detecting the wavefront set

The Radon transform has deep connections with detecting the wavefront set. The next example shows intuitively how the polar Radon transform is able to detect the wavefront set.

Example 2.24 From Example 2.19 we know that the polar Radon transform of the function $f(x,y) = \frac{1}{2} \cdot \mathbf{1}_{\{x^2+y^2 \leqslant 1\}}$ equals $\mathscr{R}^{\text{pol}} f(\theta,t) = \sqrt{1-t^2} \cdot \mathbf{1}_{\{|t| \leqslant 1\}}$. In Figure 2.8 we see the graph of $f(x)$ together with $\varphi(x) = \mathscr{R}^{\text{pol}} f(\theta, x)$.

From Figure 2.8 it is clear that when we integrate along the vertical lines, the points belonging to the wavefront set of f (i.e. the points where the lines are tangent to the

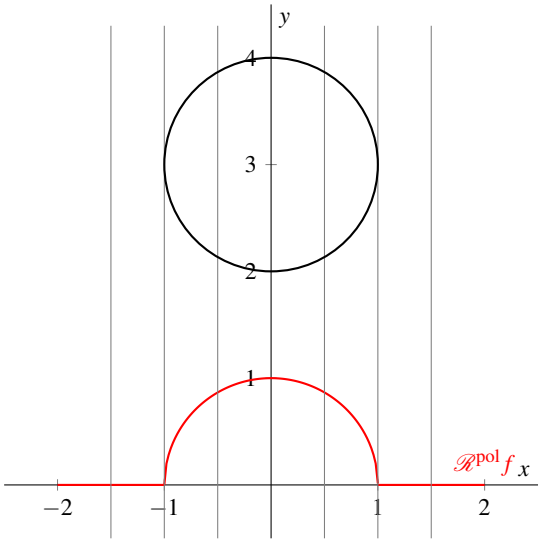


Figure 2.8 The unit disk together with $\varphi(x) = \mathscr{R}^{\mathrm{pol}} f(\theta, x)$.

circle) 'create' a point of singularity of the Radon transform, i.e. the points $x = \pm 1$, since φ is there not differentiable.

The next theorem formalizes the connection between a regular directed point of f and the vertical affine Radon transform. The result for the affine Radon transform is proven in Bartolucci (2019) and the proof for the vertical affine Radon transform is inspired of this work.

Theorem 2.25 *Let $f \in L^p(\mathbb{R}^2)$, where $1 \leqslant p \leqslant \infty$. A point $(x_0, v_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is a regular directed point of f if and only if there exist a neighborhood U_{x_0}, a function $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$ and a neighborhood V_{v_0} of v_0 such that for every $N \in \mathbb{N}$ there exists a constant C_N such that*

$$|\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi)| \leqslant C_n(1+|\xi|)^{-N}$$

for all $v \in V_{v_0}$ and all $\xi > 0$.

So Theorem 2.25 shows how we can use the Fourier transform and the vertical Radon transform to determine whether a point (x_0, v_0) is a regular directed point.

Proof Fix an $f \in L^p(\mathbb{R}^2)$, with $1 \leqslant p \leqslant \infty$. Since $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$, we know it is contained in every L^p space, so in particular we know that for a certain q we have that $\phi \in L^q(\mathbb{R}^2)$, where q satisfies the relation $p^{-1} + q^{-1} = 1$. Hölder's inequality implies that

$$\|\phi f\|_{L^1(\mathbb{R}^2)} \leqslant \|f\|_{L^p(\mathbb{R}^2)} \|\phi\|_{L^q(\mathbb{R}^2)} < \infty.$$

Therefore $\phi f \in L^1(\mathbb{R}^2)$ so the vertical affine Radon transform of ϕf is well-defined.

To prove the theorem, we have to show two directions.
($\Rightarrow$) Suppose that $(x_0, v_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is a regular directed point of f. Then there exist a neighborhood U_{x_0} of x_0, a function $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$ and a neighborhood V_{v_0} of v_0 such that, for every $N \in \mathbb{N}$, there exists by (2.6) a constant C_N with

$$|\mathscr{F}(\phi f)(\xi v, \xi)| \leqslant C_N (1 + |(\xi v, \xi)|)^{-N} \tag{2.11}$$

for all $v \in V_{v_0}$ and $\xi > 0$ We have therefore:

$$\begin{aligned} |\mathscr{F}(\mathscr{R}^{\mathrm{v}} \phi f(v,t))(\xi)| &= |\mathscr{F}(\phi f)(\xi v, \xi)| && \text{(Fourier Slice Theorem)} \\ &\leqslant C_N (1 + |\xi| \sqrt{1+v^2})^{-N} && \text{(By (2.11))} \\ &\leqslant C_N (1 + |\xi|)^{-N} \end{aligned}$$

for all $v \in V_{v_0}$ and $\xi > 0$. This proves the forward implication.

($\Leftarrow$) Suppose now that there exist a neighborhood U_{x_0} of x_0, a function $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$ and a neighborhood V_{v_0} of v_0 such that, for every $N \in \mathbb{N}$, there exists a constant C_N such that

$$|\mathscr{F}(\mathscr{R}^{\mathrm{v}} \phi f(v,t))(\xi)| \leqslant C_N (1 + |\xi|)^{-N} \tag{2.12}$$

for all $v \in V_{v_0}$ and $\xi > 0$. We then have for every $N \in \mathbb{N}$

$$\begin{aligned} |\mathscr{F}(\phi f)(\xi v, \xi)| &= |\mathscr{F}(\mathscr{R}^{\mathrm{v}} \phi f(v,t))(\xi)| && \text{(Fourier Slice Theorem)} \\ &\leqslant C_N (1 + |\xi|)^{-N}. && \text{(By 2.12)} \end{aligned}$$

Since

$$(1 + |\xi|)^{-N} = \left(\frac{\sqrt{1+v^2}}{\sqrt{1+v^2} + |\xi|\sqrt{1+v^2}} \right)^N \leqslant \left(\frac{\sqrt{1+v^2}}{1 + |\xi|\sqrt{1+v^2}} \right)^N = (\sqrt{1+v^2})^N (1 + |(\xi v, v)|)^{-N},$$

we obtain, for every $N \in \mathbb{N}$, that

$$|\mathscr{F}(\phi f)(\xi v, \xi)| \leqslant C_N (\sqrt{1+v^2})^N (1 + |(\xi v, \xi)|)^{-N} \leqslant D_N (1 + |(\xi v, \xi)|)^{-N}$$

for all $v \in V_{v_0}$ and $\xi > 0$. Here is $D_N = C_N \max\{(\sqrt{1+v^2})^N : v \in \overline{V_{v_0}}\}$, which is a constant independent of v. Note that D_n is well-defined, since we take the closure of a bounded set V_{v_0}, which is therefore compact, and hence it has a maximum value. Consequently, the point $(x_0, v_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is a regular directed point of f. □

In the following sections, we explore the theory of wavelet and shearlet systems. This rich theory plays an important role in our analysis of the wavefront set. One big advantage of these systems over the Radon transform is their practical applications, since they can be efficiently implemented in computer programs.

2.5 Wavelets

In this section we describe some general theory about wavelets that we need in this thesis. We start with some general theory about wavelets and thereafter we explain how and why the wavelet transform is capable of detecting (pointwise) singularities. We refer to classical books on wavelet theory such as Daubechies (1992), Walnut (2002), Boggess and Narcowich (2009) or Debnath and Shah (2017), and their references therein.

General Theory

Let $\psi \in L^2(\mathbb{R})$. We call the function ψ a *wavelet* if it satisfies:

$$C_\psi := \int_{\mathbb{R}} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} \mathrm{d}\xi < \infty, \tag{2.13}$$

with $\widehat{\psi}(\xi)$ the Fourier transform of $\psi(t)$. This condition is called the *admissibility condition*, which guarantees that an inverse formula for the wavelet transform exists.

We say that a wavelet has n-vanishing moments if the following condition is satisfied:

$$\int_{-\infty}^{\infty} t^k \psi(t) \mathrm{d}t = 0, \quad k = 1, 2, \ldots, n.$$

We define the *mother wavelet* $W_{b,a}\psi$ by

$$W_{b,a}\psi(x) := |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right), \quad a, b \in \mathbb{R}, \quad a > 0.$$

Here is a the so-called scaling parameter and b is a translation parameter. A straightforward calculation shows that for every $\xi \in \mathbb{R}$ and $a > 0$ we have

$$\mathscr{F}(W_{b,a}\psi)(\xi) = a^{\frac{1}{2}} e^{-2\pi i b\xi} \mathscr{F}(\psi)(a\xi). \tag{2.14}$$

Equation (2.14) will pop up in many steps in the sequel of this thesis.

Define the *wavelet transform* of a function $\psi \in L^2(\mathbb{R})$ by

$$\mathscr{W}_\psi f(b,a) = \langle f, W_{b,a}\psi \rangle = |a|^{-\frac{1}{2}} \int_{\mathbb{R}} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} \mathrm{d}x.$$

For a wavelet that satisfies the admissibility condition, the continuous wavelet transform is an isometry and admits an inversion formula. See the following theorem.

Theorem 2.26 *Let $f \in L^2(\mathbb{R})$. Then*

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathscr{W}_\psi f(a,b)\, \psi_{b,a}(t) \frac{\mathrm{d}b\,\mathrm{d}a}{a^2}, \quad a.e.$$

Furthermore, we have the following isometry relation:

$$\|f\|_{L^2(\mathbb{R})} = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\mathscr{W}_\psi f(a,b)|^2 \frac{\mathrm{d}a\mathrm{d}b}{a^2}. \tag{2.15}$$

A result that we need later on is the following, whose proof can be found in Mallat (2009).

Lemma 2.27 *Let χ be an admissible wavelet with all vanishing moments and of fast decay. Then for every $M \in \mathbb{N}$ there exists a θ with a fast decay such that*

$$\mathscr{F}(\chi)(\xi) = \xi^M \mathscr{F}(\theta\xi).$$

Remark 2.28 Note that Lemma 2.27 implies that $\theta \in L^2(\mathbb{R})$. Indeed, due to the fact that θ is of fast decay, we have for every $m \geqslant 2$ that

$$\int_{-\infty}^{\infty} |\theta(t)|^2 \mathrm{d}t \leqslant \int_{-\infty}^{\infty} \frac{c_m}{1+|t|^m} \mathrm{d}t < \infty,$$

for a certain $c_m \in \mathbb{R}$.

Detecting singularities

One of the key advantages of the wavelet transform is its ability to accurately detect singularities, thanks to its capability to localize information in both the time and frequency domains. The next theorem, whose proof can be found in Holschneider (1995), describes the connection between the decay rate of the wavelet transform at small scales with the smoothness of a function.

Theorem 2.29 *Let $f \in \mathscr{S}'(\mathbb{R})$. Then f is of $\mathscr{C}^\infty$ regularity in an open interval I (in the sense that $\varphi f \in \mathscr{C}^\infty(\mathbb{R})$ for all $\varphi \in \mathscr{C}^\infty(\mathbb{R})$ with compact support inside I) if and only if for some admissible wavelet $\psi \in \mathscr{S}_0(\mathbb{R})$, the wavelet transform satisfies*

$$\mathscr{W}_\psi f(b,a) = \mathscr{O}(a^m), \quad a \to 0$$

for all $m > 0$, uniformly in b on all compact subsets K of I.

Theorem 2.29 says essentially that the wavelet transform detects singularities by analyzing how the function behaves at different scales a. If the wavelet coefficients decay fast as $a \to 0$, this means roughly that the function is smooth and therefore it does not

have a singularity. If the coefficients vanish slowly, then the wavelet transform 'sees' the singularity of the function.

Figure 2.9 shows a signal, which has some singularities, together with the scalogram of the coefficients of the corresponding wavelet transform. Theorem 2.29 implies that whenever the signal is smooth around a point, the wavelet coefficients decay rapidly.

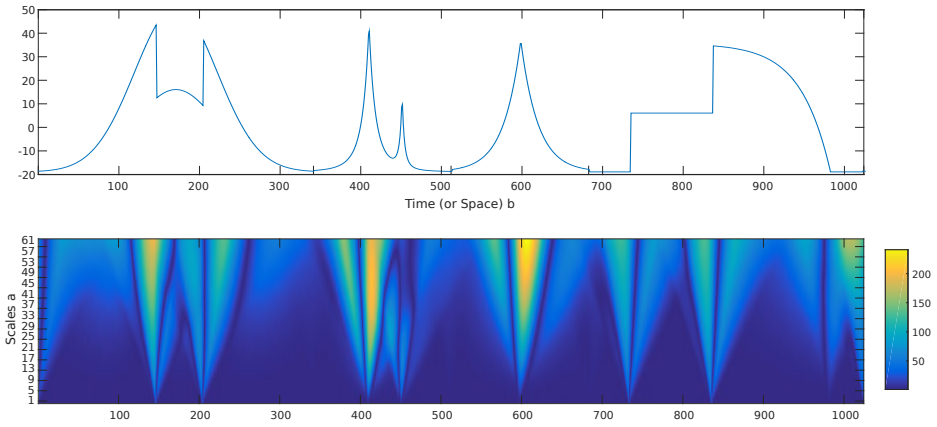


Figure 2.9 Top: A signal with singularities. Bottom: A scalogram of the coefficients of the corresponding wavelet transform.

From the scalogram in Figure 2.9 we see that the signal has some 'jumps'. In the corresponding scalogram we see around these points some cones, with relatively high values. These high values indicated that the coefficients of the wavelet transform go slow to zero around the singularities, due to the high values in the scologram. Around the points where our signal is smooth, the scalogram shows low values of the wavelet transform. In other words, the scalogram gives us information about the local smoothness of our signal and vice versa.

Our next goal is to relate the wavefront set to the cone of influence together and the smoothness properties of the wavelet coefficients.

Let ψ be a wavelet which is compactly supported on $[-M,M]$, with $M>0$. The support of $\psi_{b,a}(t)=a^{-\frac{1}{2}}\psi((t-b)/a)$ is equal to $[b-Ma,b+Ma]$. The cone of influence of b_0 is the set of points (b,a) such that b_0 is in the support of $\psi_{b,a}$. So the cone of influence of b_0 is defined by the set of points (b,a) which satisfies:

$$|b-b_0|\leqslant Ma. \tag{2.16}$$

A visualization is depicted in Figure 2.10.

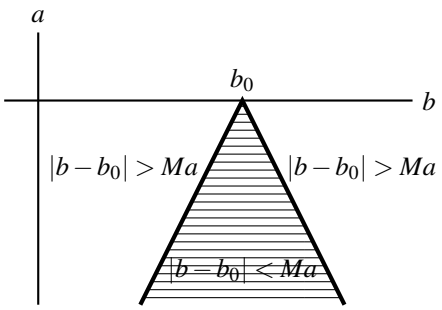


Figure 2.10: Cone of influence

The cone of influence is related to the smoothness of a function and its wavelet coefficients. Since our wavelet has vanishing moments, we know that the wavelet coefficients decay rapidly outside the cone of influence. This means that the singularity's influence is no longer significant. The cone of influence contains the points of singularity, because the wavelets centered near the singularity do not exhibit fast decay. So the cone of influence can be used to identify the region where the wavelet coefficients are influenced by a singularity, i.e, the places where they do not decay rapidly, even as the scale parameter a goes to zero.

Our goal is to relate the coefficients of the wavelet transform with the cone of influence and the wavefront set. To achieve this objective, we need the following terminology and results.

Definition 2.30 A function f is pointwise Lipschitz $\alpha\geqslant 0$ at v, if there exist a $K>0$ and a polynomial p_v of degree $m=\lfloor\alpha\rfloor$ such that

$$\forall t\in\mathbb{R},\quad |f(t)-p_v(t)|\leqslant K|t-v|^{\alpha}. \tag{2.17}$$

A function f is uniformly Lipschitz α over $[a,b]$ if it satisfies (2.17) for all $v\in[a,b]$ with a constant K that is independent of v.

In Mallat (2009) the following result is proven.

Theorem 2.31 *If $f \in L^2(\mathbb{R})$ is Lipschitz $\alpha \leqslant n$ at b_0, then there exists a constant A such that*

$$\forall (b,a) \in \mathbb{R} \times \mathbb{R}^+, \quad |\mathscr{W}_\psi f(b,a)| \leqslant A a^{\alpha+1/2} \left(1 + \left|\frac{b-b_0}{a}\right|\right)^{\alpha}. \tag{2.18}$$

Conversely, if $\alpha < n$ is not an integer and there exist a constant A and an $\alpha' < \alpha$ such that

$$\forall (b,a) \in \mathbb{R} \times \mathbb{R}^+, \quad |\mathscr{W}_\psi f(b,a)| \leqslant A a^{\alpha+1/2} \left(1 + \left|\frac{b-b_0}{a}\right|\right)^{\alpha'}, \tag{2.19}$$

then f is Lipschitz α at b_0.

We will use Theorem 2.31 to prove the following result.

Theorem 2.32 *Let $f \in L^2(\mathbb{R})$ and let (b,a) be a point in the cone of influence of b_0.*

(i) *If f is smooth, i.e. $f \in \mathscr{C}^\infty$, then for every $M \in \mathbb{N}$ and every $(b,a) \in \mathbb{R} \times \mathbb{R}^+$ we have*

$$|\mathscr{W}_\psi f(b,a)| \lesssim a^M, \quad \text{as} \quad a \to 0.$$

(ii) *If f is not smooth in the point b_0, then there is some $M \in \mathbb{N}$ such that*

$$|\mathscr{W}_\psi f(b,a)| \gtrsim a^M, \quad \text{as} \quad a \to 0.$$

Theorem 2.32 states that the decay of the wavelet coefficient of a point that lies in the cone of influence and belongs to the wavefront set goes to zero asymptotically for a fixed polynomial of power M.

Proof Fix $f \in L^2(\mathbb{R})$ and let (b,a) be a point in the cone of influence of b_0. We are interested in the case that $a \to 0$, so we can assume without loss of generality that $a \in (0,1)$.

(i) Suppose that f is a smooth function. Since f is smooth, we know that it is uniformly Lipschitz with constant $\alpha < N$, for every $N \in \mathbb{N}$. So by Theorem 2.31 there exists a constant A such that for every $(b,a) \in \mathbb{R} \times \mathbb{R}^+$ we have

$$|\mathscr{W}_\psi f(b,a)| \leqslant A a^{\alpha+1/2} \left(1 + \left|\frac{b-b_0}{a}\right|\right)^{\alpha}.$$

Since (b,a) lies in the cone of influence of b_0, we know by (2.16) that we have

$$\left|\frac{b-b_0}{a}\right| \leqslant M.$$

Consequently, we obtain:

$$|\mathscr{W}_\psi f(b,a)| \lesssim a^{\alpha+\frac{1}{2}}, \quad \text{as} \quad a \to 0.$$

Since this inequality holds for every $\alpha \leqslant N$, with $N \in \mathbb{N}$, we can always, for every fixed $M \in \mathbb{N}$, find an α with the property $\alpha \geqslant M - \frac{1}{2}$. Since $a \in (0,1)$, we consequently obtain that for every $M \in \mathbb{N}$ we have

$$|\mathscr{W}_\psi f(b,a)| \lesssim a^M, \quad \text{as} \quad a \to 0.$$

(ii) Let b_0 be a point where f is not smooth. Then in particular f is not Lipschitz for every $\alpha \leqslant N$, with $N \in \mathbb{N}$. A similar argument as before and Theorem 2.31 [take the negation of (2.19)] implies that there exists some $M \in \mathbb{N}$ such that

$$|\mathscr{W}_\psi f(b,a)| \gtrsim a^M, \quad \text{as} \quad a \to 0.$$

□

Limitations of the Wavelet Transform

So far, we described how the wavelet transform is able to detect pointwise singularities. For detecting singularities of edges in a two-dimensional setting, one may ask if two-dimensional wavelets, which possess a directional parameter, are able to work with anisotropic data structures. It turns out that the (two-dimensional) wavelet is not always able to capture the anisotropic data. The goal of this subsection is to give a mathematical argument for this. This will be done by considering an explicit example to show the failure of the two-dimensional wavelet to detect the wavefront set of anisotropic data.

We start with introducing some terminology of the two-dimensional wavelet. A two-dimensional function ψ is called a wavelet if it satisfies the admissibility condition

$$\int_{\mathbb{R}^2} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|^2} \mathrm{d}\xi < \infty.$$

Let $(a,\theta,t) \in \mathbb{R}^+ \times \mathbb{T} \times \mathbb{R}^2$, with $\mathbb{T}$ the one dimensional torus. Define the functions $\psi_{a,\theta,t}$ by

$$\psi_{a,\theta,t}(x) = \frac{1}{a}\psi\left(\frac{R_\theta(x-t)}{a}\right). \tag{2.20}$$

In this formula R_θ denotes the rotation by $\theta \in \mathbb{T}$, which can be seen as a matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Given the vector $t = (t_1, t_2)$, and applying the rotation matrix gives:

$$R_\theta(-t) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -t_1 \\ -t_2 \end{pmatrix} = \begin{pmatrix} -t_1\cos\theta + t_2\sin\theta \\ -t_1\sin\theta - t_2\cos\theta \end{pmatrix}.$$

The two-dimensional wavelet transform of a function $f \in \mathscr{S}'(\mathbb{R})$ is defined by

$$\mathscr{W}_\psi^{2D}(f)(a,\theta,t) := \langle f, \psi_{a,\theta,t}\rangle.$$

We will first give an example where the two-dimensional wavelet successfully describes the wavefront set of a distribution.

Example 2.33 Consider the point singularity δ_0, which is in $\mathscr{S}'(\mathbb{R})$ by (2.4). Then for $t = (t_1,t_2)$, where $t \neq 0$, we have that

$$\begin{aligned}\mathscr{W}_\psi^{2D}(\delta_0)(a,\theta,t) &= \langle \delta_0, \psi_{a,\theta,t}\rangle \\ &= a^{-1}\int_{\mathbb{R}^2} \delta_0(x)\psi\left(\frac{R_\theta(x-t)}{a}\right)\mathrm{d}x \\ &= a^{-1}\psi\left(\frac{R_\theta(-t)}{a}\right) = \mathscr{O}(a^{N-1}), \quad \forall N \in \mathbb{N}.\end{aligned}$$

The latter follows from Formula (2.1), since ψ is a Schwartz function. This implies that for $t \neq 0$ we have

$$\mathscr{W}_\psi^{2D}(\delta_0)(a,\theta,t) = \mathscr{O}(a^N) \quad \forall N \in \mathbb{N}.$$

When $t = 0$ we have

$$\mathscr{W}_\psi^{2D}(\delta_0)(a,\theta,0) = \mathscr{O}(a^{-1}).$$

Therefore, we see that the two-dimensional wavelet transform is (in this case) able to describe the wavefront set of δ_0.

Now we will show that the two-dimensional wavelet transform has some limitations and is not able to detect the wavefront set.

Example 2.34 Consider the distribution $\nu = \delta_{x_1=0}$. It is clear that the singularities lie on the line $x_1 = 0$. Example 2.6 implies that

$$\widehat{\nu} = \delta_{\xi_2=0}. \tag{2.21}$$

Therefore we have

$$\begin{aligned}\langle \nu, \psi_{a,\theta,0}\rangle &= \langle \widehat{\nu}, \widehat{\psi}_{a,\theta,0}\rangle && \text{(Plancherel)} \\ &= \int_{\mathbb{R}^2} \delta_{\xi_2=0}(\xi)\widehat{\psi}_{a,\theta,0}(\xi)\mathrm{d}\xi && \text{(By 2.21)} \\ &= a\int_{\mathbb{R}} \psi\left(a\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}\xi_1\\ 0\end{pmatrix}\right)\mathrm{d}\xi_1 && \text{(By Prop 2.2)} \\ &= a\int_{\mathbb{R}} \psi(a\cos(\theta)\xi_1, a\sin(\theta)\xi_1)\mathrm{d}\xi_1\end{aligned}$$

If we let $a \to 0$, then $\mathscr{W}_\psi^{2D}(\nu)(a,\theta,0) = \mathscr{O}(1)$ for every θ. So the two dimensional

wavelet varies smoothly with δ and since $\mathscr{W}_{\psi}^{\mathrm{2D}}(\nu)(a,\theta,0) = \mathscr{O}(1)$ for every θ, it is not possible to distinguish the singularity direction, which corresponds to $\theta = 0$, and the other directions.

In other words, Example 2.34 shows that, when we work with anisotropic phenomena, the two-dimensional wavelet transform is not (always) able to detect the wavefront set of a distribution ν. To overcome this limitation, we will introduce shearlet systems in the next section. It turns out that the shearlet transform is able to detect the wavefront set.

2.6 Shearlet Systems

In this section we introduce the shearlet system. We will use a group theoretic approach, which is strongly inspired by Shah and Tantary (2023). For motivation and completeness we introduce a lot of results, although we don't need them all in the sequel of this thesis. For proofs of most results, we refer to Shah and Tantary (2023) and Dahlke et al. (2009).

Shearlet transform: A Group Theoretic Approach

We start by introducing the main notation and the definition of continuous shearlets. Shearlet systems are composed of three operators, namely: scaling, shearing, and translation. The term 'continuous' indicates that continuous parameter sets are considered.

We start with defining the following four operators. The scaling matrix A_a is defined by

$$A_a = \begin{pmatrix} a & 0 \\ 0 & a^{\frac{1}{2}} \end{pmatrix}, \quad a > 0.$$

The shearing matrix S_s is given by

$$S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \quad s \in \mathbb{R}.$$

Let $\psi \in L^2(\mathbb{R}^2)$. Then the translation operator T_t is defined by:

$$T_t \psi(x) = \psi(x-t), \quad t \in \mathbb{R}^2.$$

Finally, the dilatation operator of a matrix M on $L^2(\mathbb{R}^2)$ is defined by

$$D_M \psi(x) = |\det(M)|^{-\frac{1}{2}} \psi(M^{-1}x), \quad M \in \mathrm{GL}(2,\mathbb{R}).$$

Definition 2.35 For $\psi \in L^2(\mathbb{R}^2)$, the *continuous shearlet system* $\mathrm{SH}(\psi)$ is defined by

$$\mathrm{SH}(\psi) = \left\{ \psi_{a,s,t} = T_t D_{A_a} D_{S_s} \psi : a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2 \right\}.$$

The shearlet group $\mathbb{S}$ is then defined to be the set $\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$ endowed with the group operation $\circ$ defined by

$$(a,s,t) \circ (a',s',t') = (aa', s + s'\sqrt{a}, t + S_s A_a t').$$

It can be shown that this is a locally compact group with a left-invariant Haar measure $\mathrm{d}\mu = a^{-3}\mathrm{d}a\mathrm{d}s\mathrm{d}t$. We let the unitary representation $\sigma : \mathbb{S} \to \mathscr{U}(L^2(\mathbb{R}^2))$ be defined by

$$\sigma_{a,s,t}\psi := \sigma(a,s,t)\psi := T_t D_{A_a} D_{S_s} \psi = |a|^{-\frac{3}{4}} \psi(A_a^{-1} S_s^{-1}(x-t)) \tag{2.22}$$

where $\mathscr{U}(L^2(\mathbb{R}^2))$ denotes the group of unitary operators on $L^2(\mathbb{R}^2)$. It is a tedious, but not a hard job to show that $\sigma_{a,s,t}\psi$ is indeed unitary, i.e. it satisfies

$$\sigma(a,s,t)(\sigma(a',s',t')\psi(x) = \sigma((a,s,t)(a',s',t'))(x)$$

for every $\psi \in L^2(\mathbb{R}^2)$, $x \in \mathbb{R}^2$ and $(a,s,t),(a',s',t') \in \mathbb{S}$.

Furthermore, for any pair of functions $\psi, \phi \in L^2(\mathbb{R}^2)$ we have

$$\begin{aligned}\int_{\mathbb{S}} |\langle \phi, \sigma(a,s,t)\psi \rangle|^2 \,\mathrm{d}\mu = &\int_{\mathbb{R}} \int_0^\infty \left|\widehat{f}(\xi_1,\xi_2)\right|^2 \left\{ \int_0^\infty \int_{\mathbb{R}} \frac{|\widehat{\psi}(\zeta_1,\zeta_2)|^2}{\zeta_1^2} \mathrm{d}\zeta_2 \mathrm{d}\zeta_1 \right\} \mathrm{d}\xi_1 \mathrm{d}\xi_2 \\ &+ \int_{\mathbb{R}} \int_{-\infty}^0 \left|\widehat{f}(\xi_1,\xi_2)\right|^2 \left\{ \int_{-\infty}^0 \int_{\mathbb{R}} \frac{|\widehat{\psi}(\zeta_1,\zeta_2)|^2}{\zeta_1^2} \mathrm{d}\zeta_2 \mathrm{d}\zeta_1 \right\} \mathrm{d}\xi_1 \mathrm{d}\xi_2,\end{aligned}$$

here is $\mathrm{d}\mu$ the left Haar measure on the shearlet group $(\mathbb{S}, \circ)$. This formula gives motivation for the definition of a shearlet.

Definition 2.36 A function $\psi \in L^2(\mathbb{R}^2)$ is called a *shearlet* if

$$\int_{\mathbb{R}^2} \frac{|\widehat{\psi}(\xi)|^2}{|\xi_1|^2} \mathrm{d}\xi < \infty.$$

Based on Definition 2.36, we will now introduce the shearlet transform. The shearlet transform defines a mapping of f in $L^2(\mathbb{R}^2)$ to the components of f associated with the elements of $\mathbb{S}$.

Definition 2.37 Let $\psi \in L^2(\mathbb{R}^2)$ be an admissible shearlet for the square integrable representation $\sigma : \mathbb{S} \to \mathscr{U}(L^2(\mathbb{R}^2))$ of the shearlet group $(\mathbb{S}, \cdot)$ on the Hilbert space $L^2(\mathbb{R}^2)$. Then the *Continuous Shearlet Transform* of $f \in L^2(\mathbb{R}^2)$ is the mapping

$$L^2(\mathbb{R}^2) \ni f \to \mathscr{SH}_\psi f(a,s,t) = \langle f, \sigma(a,s,t)\psi \rangle, \quad (a,s,t) \in \mathbb{S}.$$

Thus, $\mathscr{SH}_\psi$ maps the function f to the coefficients $\mathscr{SH}_\psi f(a,s,t)$ associated with the scale variable $a > 0$, the orientation variable $s \in \mathbb{R}$, and the location variable $t \in \mathbb{R}^2$.

In a similar fashion as in (2.22), we can define the vertical shearlet representation. The vertical shearlet representation is defined by

$$S^{\mathrm{v}}_{b,s,a} f(x) = |a|^{-\frac{3}{4}} f(\tilde{A}_a^{-1} \tilde{S}_s^{-1}(x-b)), \tag{2.23}$$

where

$$\tilde{S}_s = \begin{pmatrix} 1 & 0 \\ -s & 1 \end{pmatrix} \quad \text{and} \quad \tilde{A}_a = a \begin{pmatrix} |a|^{-\frac{1}{2}} & 0 \\ 0 & 1 \end{pmatrix}.$$

This leads to the associated vertical shearlet transform:

$$\mathscr{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b,s,a) := \langle f, S^{\mathrm{v}}_{b,s,a} \psi^{\mathrm{v}} \rangle. \tag{2.24}$$

2.7 The connection between the affine Radon, Wavelet, and Shearlet transforms

In this section we will give a formula with connects the (vertical) affine Radon transform and the wavelet transform with the shearlet transform. This formula will be frequently used in the next chapter. We start by introducing some notions and assumptions.

Let ψ be in $L^2(\mathbb{R}^2)$ and of the form

$$\mathscr{F}\psi(\xi_1,\xi_2) = \mathscr{F}\psi_1(\xi_1)\mathscr{F}\psi_2\left(\frac{\xi_2}{\xi_1}\right), \tag{2.25}$$

with $\psi_1 \in L^2(\mathbb{R})$ satisfying the conditions:

$$0 < \int_{-\infty}^{\infty} \frac{|\mathscr{F}\psi_1(\xi)|^2}{|\xi|} \mathrm{d}\xi < \infty \quad \text{and} \quad \int_{-\infty}^{\infty} |\xi| |\mathscr{F}\psi_1(\xi)|^2 \mathrm{d}\xi < \infty,$$

with $\psi_2 \in L^2(\mathbb{R})$. Then ψ satisfies the admissible condition (2.13). Define the function $\chi_1 \in L^2(\mathbb{R})$ by

$$\mathscr{F}\chi_1(\xi) = |\xi|^{\frac{1}{2}} \mathscr{F}\psi_1(\xi),$$

which is a one-dimensional wavelet.

In Bartolucci (2019) the following theorem has been shown true.

Theorem 2.38 *Let $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$. Then*

$$\mathscr{SH}_{\psi} f(b,s,a) = |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \mathscr{W}_{\chi_1}(\mathscr{R}^{aff} f(v,\cdot))(n(v)\cdot b, a)\, \phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right) \mathrm{d}v.$$

Here $n(v) = (1,v)$.

We can obtain an analog of Theorem 2.38 for the vertical shearlet transform. Suppose that ψ^{v} is an admissible wavelet satisfying:

$$\mathscr{F}\psi^{\mathrm{v}}(\xi_1,\xi_2) = \mathscr{F}\psi(\xi_2,\xi_1).$$

Let χ_1, ϕ_2, ψ_1 and ψ_2 as before. We impose that ψ^{v} satisfies

$$\mathscr{F}\psi^{\mathrm{v}}(\xi_1,\xi_2) = \mathscr{F}\psi_1(\xi_2)\mathscr{F}\psi_2\left(\frac{\xi_1}{\xi_2}\right), \tag{2.26}$$

where $\mathscr{F}\psi_2 = \phi_2$ satisfying the admissible condition:

$$0 < \int_{\mathbb{R}^2} \frac{|\mathscr{F}\psi(\xi)|^2}{|\xi_2|^2} \mathrm{d}\xi < \infty, \quad \xi = (\xi_1,\xi_2).$$

In this setting, we can formulate the following theorem.

Theorem 2.39 *For any $f \in L^1(\mathbb{R}^2)\cap L^2(\mathbb{R}^2)$ we have the formula:*

$$\mathscr{S}\mathscr{H}^{v}_{\psi^{v}} f(b,s,a) = |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \mathscr{W}_{\chi_1}(\mathscr{R}^{v} f(v,\cdot))(n(v)\cdot b, a)\, \phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right) \mathrm{d}v, \tag{2.27}$$

with $n(v) = (v,1)$.

As we have seen before, we know that the wavelet transform is capable of detecting one-dimensional singularities and the Radon transform is well-suited for capturing singularities and edge information of a function in higher dimensions. Theorems 2.38 and 2.39 establish a link between the (vertical) shearlet transform, the wavelet transform, and the (vertical) affine Radon transform. It is therefore reasonable to expect that the shearlet transform is suitable to detect the wavefront set. This intuition is indeed correct, as we will see in the next chapter.

3

Shearlets and the Wavefront Set

In this chapter we show that the shearlet transform can capture the wavefront set. Building on the transforms introduced in the previous chapter (the wavelet, shearlet, and (vertical) Radon transform) along with the result of Theorem 2.39, which establishes the connection between the shearlet transform and the wavelet and affine Radon transforms. We will show how these results play a central theme in our analysis.

The goal of Section 3.1 is to provide an explicit example and show that the shearlet transform is indeed capable of detecting the wavefront set for this specific function, based on properties of the wavelet and Radon transforms. This motivating example will lead to a general theory, which will be the scope of the subsequent three sections.

3.1 A motivating example

Theorem 2.29 implies that a function is smooth around a point if and only if the coefficients of the wavelet transform decay faster than any polynomial. Therefore, the wavelet transform is capable to 'detect' pointwise singularities and we would expect, due to the relationship mentioned in Theorem 2.39, that the shearlet transform is also able to 'detect' these points of singularities. The objective of this section is to provide an example to show that our intuition is indeed correct. We will provide a function for which we can explicitly calculate the vertical Radon transform.

Consider the function $f : \mathbb{R}^2 \to [0, \frac{1}{2}]$, defined by

$$f(x,y) := \frac{1}{2} \begin{cases} 1 & \text{if } x^2 + y^2 \leqslant 1, \\ 0 & \text{if } x^2 + y^2 > 1. \end{cases}$$

In Example 2.23 we computed the corresponding vertical affine Radon transform:

$$\mathscr{R}^{\mathrm{v}} f(v,t) = \begin{cases} \dfrac{\sqrt{1+v^2-t^2}}{1+v^2} & \text{if } t^2 - v^2 \leqslant 1, \\ 0 & \text{if } t^2 - v^2 > 1. \end{cases}$$

Based on this calculation, we define the function

$$\phi(t) = \begin{cases} \sqrt{1-t^2} & \text{if } |t| \leqslant 1, \\ 0 & \text{if } |t| > 1. \end{cases} \tag{3.1}$$

The graph of $\phi(t)$ is shown in Figure 3.1. It is clear from the graph that $\phi(t)$ is smooth except at the points $t = \pm 1$.

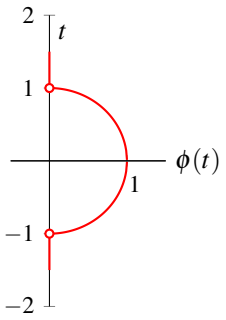


Figure 3.1: Graph of $\phi(t) = \mathscr{R}^{\mathrm{v}} f(0,t)$

The scalogram of the coefficients of the wavelet transform of ϕ is shown in Figure 3.2.

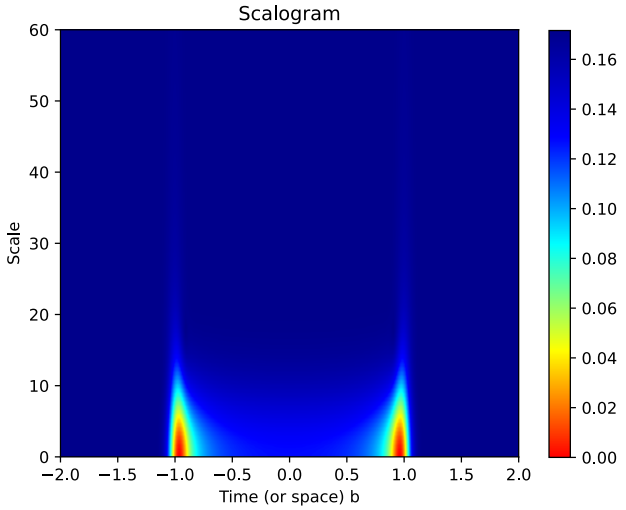


Figure 3.2 Scalogram of the wavelet coefficients of ϕ.

In Figure 3.2 we see two 'cones' around the singular points $t = \pm 1$ of $\phi(t)$. Figure 3.2 suggests that the wavelet coefficients $\mathscr{W}_{\chi_1}\phi(b,a)$ of the function ϕ decrease faster than any polynomial if and only if $a \to 0$ and $b = \pm 1$. To make these observations precise, we need the following auxiliary lemma and make some explicit calculations.

Lemma 3.1 *Let χ_1 be a compactly supported wavelet on the interval $[-1,1]$ and ϕ defined by (3.1) . Then*

$$\mathscr{W}_{\chi_1}(\mathscr{R}^v f(v,t))(b,a) = \frac{1}{\sqrt[4]{1+v^2}} \mathscr{W}_{\chi_1}(\phi)\left(\frac{b}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}}\right).$$

Proof We start with showing that

$$\mathscr{R}^{\mathrm{v}} f(v,t) = \frac{1}{\sqrt[4]{1+v^2}} W_{0,\sqrt{1+v^2}}(\phi)(t). \tag{3.2}$$

For $t^2 - v^2 > 1$, this identity is trivially true, since both sides of (3.2) are equal to zero. So we consider the case $t^2 - v^2 \leqslant 1$. Then we have

$$\begin{aligned}
\frac{1}{\sqrt[4]{1+v^2}} (W_{0,\sqrt{1+v^2}}\phi)(t) &= \frac{1}{\sqrt[4]{1+v^2}} \cdot \frac{1}{\sqrt[4]{1+v^2}} \phi\left(\frac{t}{\sqrt{1+v^2}}\right) \\
&= \frac{1}{\sqrt{1+v^2}} \cdot \sqrt{1 - \frac{t^2}{1+v^2}} = \frac{\sqrt{1+v^2-t^2}}{1+v^2} \\
&= \mathscr{R}^{\mathrm{v}} f(v,t).
\end{aligned}$$

Then for any fixed t we have by (3.2) that

$$\begin{aligned}
\mathscr{W}_{\chi_1}(\mathscr{R}^{\mathrm{v}} f(v,t))(b,a) &= \langle \mathscr{R}^{\mathrm{v}} f, W_{b,a}\chi_1 \rangle \\
&= \frac{1}{\sqrt[4]{1+v^2}} \langle W_{0,\sqrt{1+v^2}}\phi, W_{b,a}\chi_1 \rangle \\
&= \frac{1}{\sqrt{a(1+v^2)}} \int_{-\infty}^{\infty} \phi\left(\frac{t}{\sqrt{1+v^2}}\right) \chi_1\left(\frac{t-b}{a}\right) \mathrm{d}t.
\end{aligned} \tag{3.3}$$

Furthermore, we have:

$$\begin{aligned}
&\frac{1}{\sqrt[4]{1+v^2}} (\mathscr{W}_{\chi_1}\phi)\left(\frac{b}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}}\right) \\
&= \frac{1}{\sqrt[4]{1+v^2}} \langle \phi, W_{\frac{b}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}}} \chi_1 \rangle \\
&= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \phi(t) \chi_1\left(\frac{t - \frac{b}{\sqrt{1+b^2}}}{\frac{a}{\sqrt{1+v^2}}}\right) \mathrm{d}t \\
&= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \phi(t) \chi_1\left(\frac{t\sqrt{1+v^2} - b}{a}\right) \mathrm{d}t \\
&= \frac{1}{\sqrt{a(1+v^2)}} \int_{-\infty}^{\infty} \phi\left(\frac{t}{\sqrt{1+v^2}}\right) \chi_1\left(\frac{t-b}{a}\right) \mathrm{d}t.
\end{aligned} \tag{3.4}$$

Comparing (3.3) and (3.4) we deduce

$$\mathscr{W}_{\chi_1}(\mathscr{R}^{\mathrm{v}} f(v,t))(b,a) = \frac{1}{\sqrt[4]{1+v^2}} \mathscr{W}_{\chi_1}(\phi)\left(\frac{b}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}}\right), \tag{3.5}$$

which proves the desired equality. □

We will integrate $\phi(t)$ along horizontal lines. See Figure 3.3.

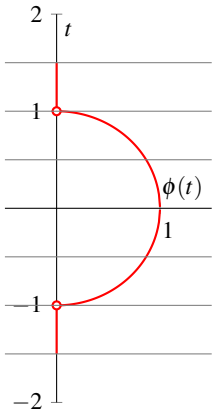


Figure 3.3 Graph of $\phi(t) = \mathscr{R}^{\mathrm{v}} f(0,t)$ together with horizontal lines.

We will now show that the shearlet transform, using the wavelet transform and the vertical Radon transform, is capable of detecting the wavefront set of f.

Consider the point $(0,1)$. This point lies on the boundary of $x^2+y^2 \leqslant 1$. If we integrate along horizontal lines (i.e. with direction '0'), then the horizontal line through the point $(v,1)$, with $v \in \mathbb{R}$, 'touches' this point. Thus, the point $(0,1,0)$ lies in the wavefront set of f. We will show that the decay behavior of the vertical shearlet transform in this point is not faster than any polynomial.

In the setting of the vertical shearlet transform, we have for the point $(0,1)$ that $b = (0,1)$. Consequently, by Figure 3.4, we have $\theta = 0$ and hence $s = \tan(\theta) = 0$. This implies that

$$n(v) \cdot b = (v,1) \cdot (0,1) = 1.$$

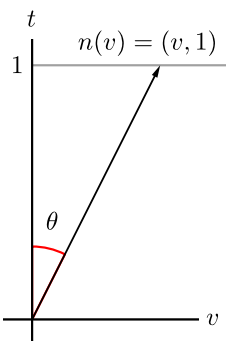


Figure 3.4: Direction of integrating.

In the setting of Theorem 2.39 we take $\phi_2(t) = \mathbf{1}_{[-1,1]}(t)$ and we assume that χ_1 is an admissible wavelet such that $\operatorname{supp}(\chi_1) \subsetneqq [-1,1]$. If we integrate along the horizontal lines, such as depicted in Figure 3.3, we obtain:

$$
\begin{aligned}
|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(0,1,0,a)| &= |a|^{-\frac{3}{4}} \left| \int_{s-|a|^{\frac{1}{2}}}^{s+|a|^{\frac{1}{2}}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{v}} f(v,\cdot))(v,a)\mathrm{d}v \right| && \text{(By Thm. 2.39)} \\
&= |a|^{-\frac{3}{4}} \left| \int_{s-|a|^{\frac{1}{2}}}^{s+|a|^{\frac{1}{2}}} \mathcal{W}_{\chi_1}\phi\left(\frac{1}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}}\right) \frac{\mathrm{d}v}{\sqrt[4]{1+v^2}} \right| && \text{(By (3.5))} \\
&= 2|a|^{-\frac{1}{4}} \left| \frac{1}{\sqrt[4]{1+\bar{v}_a^2}} \mathcal{W}_{\chi_1}\phi\left(\frac{1}{\sqrt{1+\bar{v}_a^2}}, \frac{a}{\sqrt{1+\bar{v}_a^2}}\right) \right| && \text{(MVT)} \\
&\geqslant \frac{1}{\sqrt[4]{1+\bar{v}_a^2}} \left| \mathcal{W}_{\chi_1}\phi\left(\frac{1}{\sqrt{1+\bar{v}_a^2}}, \frac{a}{\sqrt{1+\bar{v}_a^2}}\right) \right| && (|a| \leqslant 1)
\end{aligned}
$$

In the third line we used the mean value theorem for integrals, with $\bar{v}_a \in [-|a|^{\frac{1}{2}}, |a|^{\frac{1}{2}}]$. In the last line we used that $|a| \leqslant 1$. This condition is not restrictive, since we are interested in the case $a \to 0$. As $a \to 0$, then $\bar{v}_a \to 0$, so $1/\sqrt{1+\bar{v}_a^2} \to 1$.

We claim that the point $(1/\sqrt{1+\bar{v}_a^2}, a/\sqrt{1+\bar{v}_a^2})$ belongs to the cone of influence created by the singularity $t = 1$, which is a singular point of ϕ. To prove this claim, we need to show that for every $a \in \mathbb{R}$ we have

$$
\left| 1 - \frac{1}{\sqrt{1+\bar{v}_a^2}} \right| \leqslant \frac{|a|}{\sqrt{1+\bar{v}_a^2}}.
$$

This is equivalent with

$$
-\sqrt{2|a|+a^2} \leqslant \bar{v} \leqslant \sqrt{2|a|+a^2}.
$$

But we always have $|a|^{\frac{1}{2}} \leqslant \sqrt{2|a|+a^2}$. This proves the claim.

Since ϕ is not smooth in the point $t = 1$, by Theorem 2.32 part (ii) we know that there is an $N \in \mathbb{N}$ such that $|\mathcal{W}_{\chi_1}(\phi)(1/\sqrt{1+\bar{v}_a^2}, a/\sqrt{1+\bar{v}_a^2})| \gtrsim a^N$. From this we observe that the wavelet coefficients do not have a fast decay whenever $a \to 0$. Consequently we have

$$
|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(0,1,0,a)| \gtrsim a^N
$$

for some $N \in \mathbb{N}$. Therefore, based on this example, we see that the shearlet transform does not decay too fast around a point belonging to the wavefront set. An intuitive reasoning is that this is caused due to the fact that the wavelet coefficients do not decay quick as $a \to 0$. If they decayed too quickly, then by Theorem 2.29 they indicate smoothness rather than the presence of a singularity.

Now consider the point $(1,0)$. From Figure 3.1 it is clear that ϕ is smooth around the point $(1,0)$, so the point $(1,0,0)$ does not lie in the wavefront set. The following calculation shows that the coefficients of the shearlet transform equal $\mathcal{O}(a^N)$. Since $n(v) = (v,1)$, we get $n(v)\cdot(1,0) = v$ and therefore:

$$\begin{aligned}
|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(1,0,0,a)| &= |a|^{-\frac{3}{4}} \left| \int_{s-|a|^{\frac{1}{2}}}^{s+|a|^{\frac{1}{2}}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{v}}(v,\cdot))(v,a)\mathrm{d}v \right| && \text{(By Thm. 2.39)} \\
&= |a|^{-\frac{3}{4}} \left| \int_{s-|a|^{\frac{1}{2}}}^{s+|a|^{\frac{1}{2}}} \mathcal{W}_{\chi_1}\phi\left(\frac{v}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}}\right) \frac{\mathrm{d}v}{\sqrt[4]{1+v^2}} \right| && \text{(By (3.5))} \\
&= 2|a|^{-\frac{1}{4}} \left| \frac{1}{\sqrt[4]{1+\bar{v}_a^2}} \mathcal{W}_{\chi_1}\phi\left(\frac{\bar{v}_a}{\sqrt{1+\bar{v}_a^2}}, \frac{a}{\sqrt{1+\bar{v}_a^2}}\right) \right| && \text{(MVT)} \\
&= \mathcal{O}(a^{N-\frac{1}{4}}).
\end{aligned}$$

The latter follows by Theorem 2.29, which can be applied since ϕ is smooth around $(0,1)$. So we conclude that for every $N \in \mathbb{N}$ we have

$$|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(1,0,0,a)| = \mathcal{O}(a^{N-\frac{1}{4}}), \quad \text{as } a \to 0.$$

Since this holds for every $N \in \mathbb{N}$, we can therefore conclude that $|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(0,1,0,a)| = \mathcal{O}(a^N)$ for every $N \in \mathbb{N}$ as $a \to 0$. From this we see that if a point does not belong to the wavefront set, then the shearlet coefficients show fast decay.

Based on this example, we formulate the following conjecture.

Conjecture 3.2 *Let f a function. A point (b,s) does not belong to the wavefront set of f if and only if for every $N \in \mathbb{N}$ we have*

$$\mathcal{SH}^{v}_{\psi^{v}} f(b,s,a) = \mathcal{O}(a^N), \quad \text{as} \quad a \to 0.$$

This conjecture turns out to be true under some conditions of the function f and the wavelet ψ. We will make our intuition precise in the next section and provide a proof.

We close this section with two final remarks.

Remark 3.3 (i) The example presented in this section shows that the shearlet transform can successfully detect the wavefront set. This capability arises from the presence of singularities in the function f, which induce singularities in $\phi = \mathcal{R}^{\mathrm{v}}$. Moreover, the result of Theorem 2.39 plays a key role in our analysis, as it establishes that the wavelet transform effectively characterizes the smoothness of univariate functions.

(ii) Although we derived an explicit expression for the Radon transform of f, this is in fact not really necessary. What really matters is the location of its singularities.

This point becomes evident when we have a second look at the formula given in Theorem 2.39. From this expression, we see that the decay behavior of the wavelet coefficients within the integral, which is determined by the location of the singularities of the vertical affine Radon transform, influences the decay behavior of the shearlet.

Remark 3.4 Based on the motivation example given in this paragraph, we will now give a sketch of the connection between the shearlet transform and a point in the wavefront set of an arbitrarily function. This sketch is incomplete and can perhaps be made precise in further research.

Suppose $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ and assume $(b_0, s_0) \in \mathrm{WF}(f)$. Assume without loss of generality that $\mathrm{supp}(\chi_2) \subsetneqq [-1,1]$ and ϕ_2 is the characteristic function of the interval $[-1,1]$. Let $n(v) = (1,v)$ and $b_0 = (b_{0,1}, b_{0,2})$. Keeping Theorem 2.39 in mind, it follows that

$$|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b_0, s_0, a)| = |a|^{-\frac{3}{4}} \left| \int_{s_0 - |a|^{\frac{1}{2}}}^{s_0 + |a|^{\frac{1}{2}}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{aff}} f(v, \cdot))(n(v) \cdot b_0, a) \mathrm{d}v \right|.$$

By the mean value theorem, we know that there is a $\bar{v}_a$ in the interval

$$(s_0 - |a|^{\frac{1}{2}}, s_0 + |a|^{\frac{1}{2}}) \tag{3.6}$$

such that this integral equals

$$\begin{aligned} |\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b_0, s_0, a)| &= |a|^{-\frac{3}{4}} [s_0 + |a|^{\frac{1}{2}} - (s_0 - |a|^{\frac{1}{2}})] |\mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{v}}(\bar{v})a, \cdot))(n(\bar{v}_a) \cdot b_0, a)| \\ &= 2|a|^{-\frac{1}{4}} |\mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{v}}(\bar{v}_a, \cdot))(n(\bar{v}_a) \cdot b_0, a)|. \end{aligned}$$

We are interested in the behavior whenever a approaches zero, so the interval (3.6) becomes small. Therefore, write $\bar{v}_a = s_0 + \varepsilon_a$, with ε_0 represents a very small constant. In this case we have

$$n(\bar{v}_a) \cdot b = (1, s_0 + \varepsilon_a) \cdot (b_{0,1}, b_{0,2}) = b_{0,1} + s_0 b_{0,2} + \varepsilon_a b_{0,2}.$$

Consequently

$$|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b_0, s_0, a)| = 2|a|^{-\frac{1}{4}} |\mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{v}}(\bar{v}_a, \cdot))(b_{0,1} + s_0 b_{0,2} + \varepsilon_a b_{0,2}, a)|.$$

We would like to make the following argument: observe that $\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b_0, s_0, a)$ is the only shearlet coefficient that contains $\mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{v}}(f(s_0, \cdot))(n(s_0) b_0, a)$ as $a \to 0$ and since $(b_0, s_0) \in \mathrm{WF}(f)$, it follows that the wavelet coefficients contain information about this singularity. We therefore obtain:

$$\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b_0, s_0, a)| \gtrsim a^N.$$

3.2 Detecting the Wavefront set using Shearlets

In the previous section, we saw an explicit example where we showed that the shearlet transform is capable to detect a point in the wavefront set. This example shows the role of the shearlet transform in the study of wavefront set resolution in shearlet analysis. Based on this example, we will formulate a general theorem in the following subsection. This theorem is one of the main results of this thesis and constitutes a novel contribution to the existing literature, as we will discuss in more details in the subsequent subsection.

The Main Theorem

In the setting of our main theorem of this thesis, we will assume that ψ is of the form (2.25) and ψ^{v} is of the form (2.26). Furthermore, from now on we assume that we have $\chi_1 \in \mathscr{S}_0(\mathbb{R})$ and $\phi_2 \in \mathscr{C}_0^\infty(\mathbb{R})$. Finally, we impose a growth condition on ψ and ψ^{v}; they are a rapidly decreasing functions, i.e. for every $x \in \mathbb{R}^2$ and every $N \in \mathbb{N}$ we have $\psi(x) = \mathscr{O}((1+|x|)^{-N}$ and $\psi^{\mathrm{v}}(x) = \mathscr{O}((1+|x|)^{-N}$.

Theorem 3.5 *Let f be a function in $L^2(\mathbb{R}^2)$ and let (b_0,s_0) be a point in $\mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$. Define the following two spaces:*

- $\mathscr{D}_1 := \{(b_0,s_0) \in \mathbb{R}^2 \times [-1,1]$: *there exist neighborhoods V_{b_0} of b_0 and V_{s_0} of s_0 such that for every $N > 0$ there exists a $C_N > 0$ satisfying $|\mathscr{SH}_\psi f(b,s,a)| \leqslant C_N|a|^N$ as $a \to 0$ for every $b \in V_{b_0}$ and every $s \in V_{s_0}\}$.*
- $\mathscr{D}_2 := \{(b_0,s_0) \in \mathbb{R}^2 \times ((-\infty,-1) \cup (1,\infty))$: *there exist neighborhoods V_{b_0} of b_0 and V_{s_0} of s_0 such that for every $N > 0$ there exists a $C_N > 0$ satisfying $|\mathscr{SH}^v_{\psi^v} f(b,s,a)| \leqslant C_N|a|^N$ as $a \to 0$ for every $b \in V_{b_0}$ and every $1/s \in V_{s_0}\}$.*

Then $\mathscr{D}_1 \cup \mathscr{D}_2 = WF(f)^c$.

Since we work with the sets $\mathscr{D}_1$ and $\mathscr{D}_2$, we split $\mathbb{R}^2$ in two different regions. In each cone we use the shearlet transform or the vertical shearlet transform. In this way we work with 'cone-adapted' shearlets, see Figure 3.5.

First, note that $\mathscr{D}_1 \cap \mathscr{D}_2 = \varnothing$, so if $(b,s) \in \mathscr{D}_1 \cup \mathscr{D}_2$, then we have either $(b,s) \in \mathscr{D}_1$ or $(b,s) \in \mathscr{D}_2$. If $(b,s) \in \mathscr{D}_1$, then we are in the green region and we are integrating along vertical lines. In this case, we work with the affine shearlet transform. Note that the tangent of the corresponding parametrization is between -1 and 1. If $(b,s) \in \mathscr{D}_2$, then we are integrating over the horizontal lines. In this case, we work with the vertical shearlet transform.

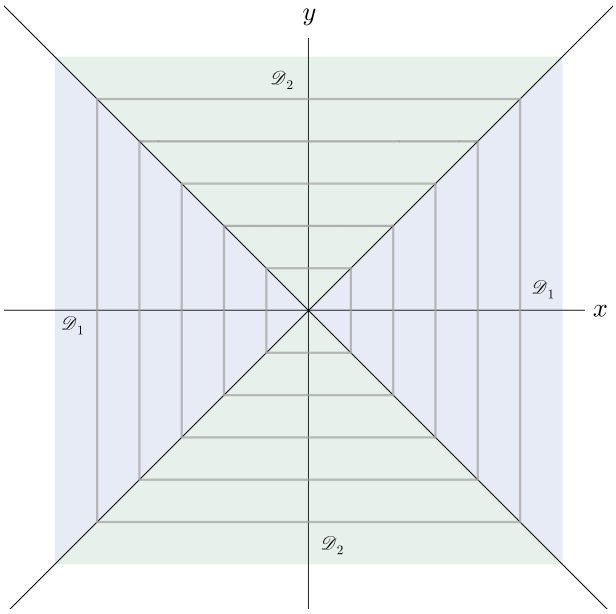


Figure 3.5 Cone-adapted shearlets

To prove Theorem 3.5, we need to show two inclusions, i.e. we need to show the following two cases:

- $(b,s) \in \mathrm{WF}(f)^c$ only if $(b,s) \in \mathscr{D}_1 \cup \mathscr{D}_2$;
- If $(b,s) \in \mathscr{D}_1 \cup \mathscr{D}_2$, then $(b,s) \in \mathrm{WF}(f)^c$.

The proof of Theorem 3.5 is quite long and therefore the following two sections are devoted to providing the proof. We split it into the 'Only if' part and the 'If' part, which correspond respectively with the two bullets above.

Contribution to the Literature

Theorem 3.5 is one of the most important theorems of this thesis. While the result has been established for more than a decade, Grohs's original 2010 proof was quite technical, see Grohs (2010). Some years later, in 2019, Bartolucci obtained an integral representation of the shearlet transform in terms of the wavelet and the Radon transforms. This integral representation was used to formulate and prove a theorem for detecting the wavefront set using the shearlet transform. In Bartolucci (2019), only the shearlet transform was used, but the vertical shearlet transform was not used. The problem with

this framework is that we are not able to cover the whole wavefront set, because only (a version of) the region $\mathscr{D}_1$ is taken into account.

The new contribution to the literature is that we consider not only the shearlet transform, but also the vertical shearlet. Working with cone-adapted shearlets, i.e. by considering the sets $\mathscr{D}_1$ and $\mathscr{D}_2$, we make a distinction between the use of the shearlet transform and the vertical shearlet transform. This approach allows us to generalize the results from Bartolucci (2019) to cover all possible directions in the wavefront set.

A chronological visualization of the key results in the study of the wavefront set, along with the results of this thesis, is shown in Figure 3.6.

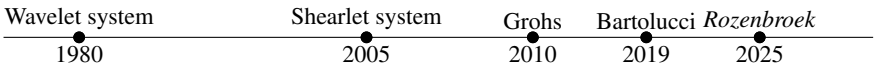


Figure 3.6 Timeline of the study of the wavefront set

The proof of the vertical shearlet transform, presented in the next two sections, follows the approach outlined for the shearlet transform in Bartolucci (2019). In addition, we offer much motivation and include the intermediate steps in detail.

3.3 'Only if' part of the proof of Theorem 3.5

Suppose that $(b,s) \in \mathrm{WF}(f)^c$. If $(b,s) \notin \mathscr{D}_2$, then in Bartolucci (2019) it is shown that $(b,s) \in \mathscr{D}_1$. So our task is to show that if $(b,s) \notin \mathscr{D}_1$, then $(b,s) \in \mathscr{D}_2$. This will be achieved in this section.

From Theorem 2.25 we know that the wavefront set of an L^p function f can be described using the decay properties of the Fourier transform of the function $t \mapsto \mathscr{R}^{\mathrm{v}}(\phi f)(v,t)$. In Section 2.5, we saw that wavelets are able to detect whether a function is smooth at a point. So it makes sense to expect that the decay properties of the wavelet coefficients of $t \mapsto \mathscr{R}^{\mathrm{v}}(\phi f)(v,t)$ can tell us something about the microlocal features of f.

We now formalize and prove this intuition in the lemma below.

Lemma 3.6 *Let $f \in L^2(\mathbb{R}^2)$ and let $(x_0,v_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ be a regular directed point of f. Let $\chi \in L^2(\mathbb{R})$ be an admissible wavelet with all vanishing moments such that $\mathscr{F}(\chi) \in L^1(\mathbb{R})$. Then, there exist a neighborhood U_{x_0} of x_0, a neighborhood V_{v_0} of v_0 and a function $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$ such that for every $N > 0$ there exists a constant C_N such that*

$$|\mathscr{W}_\chi(\mathscr{R}^{\mathrm{v}}(\phi f))(v,t)| \leqslant C_N |t|^N \tag{3.7}$$

for every $v \in V_{v_0}$ and $t \in \mathbb{R}$.

Proof Let $(x_0, v_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ be a regular directed point of $f \in L^2(\mathbb{R}^2)$. Theorem 2.25 implies that there exist a neighborhood U_{x_0} of x_0, a neighborhood V_{v_0} of v_0 and a function $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$ such that, for every $N \in \mathbb{N}$, there exists a constant C_N with

$$|\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi)| \leqslant C_N(1+|\xi|)^{-N} \tag{3.8}$$

for all $v \in V_{v_0}$ and $\xi > 0$. By the definition of the wavelet transform and Parseval's identity, we have

$$\mathscr{W}_\chi(\mathscr{R}^v(\phi f)(v,t))(b,a) = \langle \mathscr{R}^v(\phi f)(v,t), W_{b,a}\chi \rangle = \langle \mathscr{F}(\mathscr{R}^v(\phi f)(v,t)), \mathscr{F}(W_{b,a}\chi) \rangle.$$

Parseval's identity is justified by the fact that both $W_{b,a}\chi$ and $\mathscr{R}^{\mathrm{v}}(\phi f)$ are in L^2, the latter is a consequence of Hölder's inequality and the fact that ϕ is in $\mathscr{C}_0^\infty$. Combining this result with Formula (2.14), we obtain

$$|\mathscr{W}_\chi(\mathscr{R}^v(\phi f)(v,t))(b,a)| \leqslant |a|^{\frac{1}{2}} \int_{-\infty}^{\infty} |\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi))| \mathscr{F}\chi(a\xi)| \mathrm{d}\xi. \tag{3.9}$$

For every $M > 0$ we can find an α such that $0 < \alpha < \frac{2M}{2M+1}$. We split the integral in (3.9) as the sum of A and B, where

$$A := |a|^{\frac{1}{2}} \int_{|\xi| \leqslant |a|^{-\alpha}} |\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi))| \mathscr{F}\chi(a\xi)| \mathrm{d}\xi,$$
$$B := |a|^{\frac{1}{2}} \int_{|\xi| > |a|^{-\alpha}} |\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi))| \mathscr{F}\chi(a\xi)| \mathrm{d}\xi.$$

Estimating A yields:

$$\begin{aligned}
A &= |a|^{\frac{1}{2}+M} \int_{|\xi| \leqslant |a|^{-\alpha}} |\xi|^M |\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi))| \mathscr{F}(\theta)(a\xi)| \mathrm{d}\xi && \text{(Lemma 2.27)} \\
&\leqslant |a|^{\frac{1}{2}+M(1-\alpha)} \int_{|\xi| \leqslant |a|^{-\alpha}} |\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(\xi))| \mathscr{F}(\theta)(a\xi)| \mathrm{d}\xi && \text{(Since } |\xi| \leqslant |a|^{-\alpha}\text{)} \\
&\leqslant |a|^{\frac{1}{2}+M(1-\alpha)} \int_{|\xi| \leqslant |a|^{-\alpha}} |\mathscr{F}((\phi f)(v\xi,\xi))| \mathscr{F}(\theta)(a\xi)| \mathrm{d}\xi && \text{(Fourier slice theorem)} \\
&\leqslant |a|^{\frac{1}{2}+M(1-\alpha)} \|\mathscr{F}(\phi f)\|_{L^\infty} \int_{|\xi| \leqslant |a|^{-\alpha}} |\mathscr{F}(\theta)(a\xi)| \mathrm{d}\xi && \text{(Hölder)} \\
&\leqslant |a|^{M(1-\alpha)} \|\phi f\|_{L^1} \int_{|\xi| \leqslant |a|^{-\alpha}} |\mathscr{F}(W_{x+vy,a}\theta)(\xi)| \mathrm{d}\xi && \text{(Prop. Fourier transform)} \\
&\leqslant \sqrt{2}|a|^{M(1-\alpha)-\frac{\alpha}{2}} \|\phi f\|_{L^1} \|\mathscr{F}(W_{x+vy,a}\theta)\|_{L^2} && \text{(Cauchy-Schwarz)} \\
&= \sqrt{2}|a|^{M(1-\alpha)-\frac{\alpha}{2}} \|\phi f\|_{L^1} \|\theta\|_{L^2}. && \text{(Plancherel)}
\end{aligned}$$

We can also estimate B by

$$\begin{aligned}
B &\leqslant C_N|a|^{\frac{1}{2}}\int_{|\xi|>|a|^{-\alpha}}(1+|\xi|)^{-N}|\mathscr{F}\chi(a\xi)|\mathrm{d}\xi && \text{(By (3.8))}\\
&\leqslant C_N|a|^{\frac{1}{2}}\int_{|\xi|>|a|^{-\alpha}}|\xi|^{-N}|\mathscr{F}\chi(a\xi)|\mathrm{d}\xi && \text{(Since } |\xi|\leqslant|1+\xi|\text{)}\\
&\leqslant C_N|a|^{\alpha N+\frac{1}{2}}\int_{|\xi|>|a|^{-\alpha}}|\mathscr{F}\chi(a\xi)|\mathrm{d}\xi && \text{(Since } |\xi|>|a|^{-\alpha}\text{)}\\
&= C_N|a|^{\alpha N-\frac{1}{2}}\int_{|\xi|>|a|^{-\alpha}}|\mathscr{F}\chi(\xi)|\mathrm{d}\xi && \text{(Dilation)}\\
&\leqslant C_N|a|^{\alpha N-\frac{1}{2}}\|\mathscr{F}\chi\|_{L^1}. && \text{(Since } \{|\xi|>|a|^{-\alpha}\}\subseteqq\mathbb{R}^2\text{)}
\end{aligned}$$

The latter is finite, since we imposed that $\mathscr{F}\chi\in L^1$. The constant C_N is independent of v and since our previous two estimates work for every positive M and N, we obtain

$$|\mathscr{W}_\chi(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(b,a)| = \mathscr{O}(|a|^{M(1-\alpha)-\frac{\alpha}{2}}+|a|^{\alpha N-\frac{1}{2}}),$$

with $v\in V_{v_0}$. Consequently, since χ has all vanishing moments, we deduce for every $N>0$ that

$$|\mathscr{W}_\chi(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(b,a)| = \mathscr{O}(|a|^N).$$

So we obtain (3.7) for every $v\in V_{v_0}$ and every $t\in\mathbb{R}$. □

With this lemma in mind, we can now prove the 'Only if' part of the proof of Theorem 3.5.

Proof 'Only if' part of Theorem 3.5 Let $(b_0,s_0)\in\mathbb{R}^2\times\mathbb{R}\setminus\{0\}$ be a regular directed point of $f\in L^2(\mathbb{R}^2)$. Lemma 3.6 implies that there exist a neighborhood U_{b_0} of b_0, a neighborhood V_{s_0} of s_0 and a function $\phi\in\mathscr{C}_0^\infty(\mathbb{R}^2)$, with $\phi(x_0)\neq 0$, such that for every $N>0$ we have

$$|\mathscr{W}_\chi(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(b,a)|\leqslant C_N|a|^N, \tag{3.10}$$

for a certain constant C_N and for all $v\in V_{s_0}$ and $b\in\mathbb{R}$. Furthermore, we can assume without loss of generality that $\phi\equiv 1$ on U_{b_0}. So for every $b\in U_{b_0}$ we can assume that $|x-b|>\delta$ for some positive δ.

We are interested in the behavior of $|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}f(b,s,a)|$ as $a\to 0$. By (3.10) we know how the wavelet transform of the vertical Radon transform of ϕf behaves. Therefore, we write $f=(1-\phi)f+\phi f$ and we get the following inequality

$$|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}f(b,s,a)|\leqslant|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(1-\phi)f(b,s,a)|+|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)|.$$

So to study the behavior of the shearlet coefficient of f, it suffices to study the behavior of the shearlet coefficients of $(1-\phi)f$ and ϕf.

We start with giving an estimate of the shearlet coefficients $|\mathcal{SH}^{\mathrm{v}}_{\psi}((1-\phi)f)(b,s,a)|$. We have:

$$\begin{aligned} S^{\mathrm{v}}_{b,s,a}\psi &= |a|^{-\frac{3}{4}}|\psi(\tilde{A}_a^{-1}\tilde{S}_s^{-1}(x-b)), && \text{(By (2.23))} \\ &\leqslant C_N|a|^{-\frac{3}{4}}(1+|\tilde{A}_a^{-1}\tilde{S}_s^{-1}(x-b)|)^{-N} && \text{(Since } \psi \text{ is rapidly decreasing)} \end{aligned}$$

For every $x \in \mathbb{R}^2$ and $M \in \mathrm{GL}(2,\mathbb{R})$ we have $|x| = |MM^{-1}x| \leqslant \|M\||M^{-1}x|$, where $\|\cdot\|$ denotes the spectral norm[1] of a matrix. So we obtain.

$$\|M\|^{-1}|x| \leqslant |M^{-1}x|.$$

The eigenvalues of the matrix $\tilde{A}_a^{\dagger}\tilde{A}_a$ are a^2 and $|a|^{\frac{1}{2}}$. Since we are interested in the behavior of the shearlet coefficients as $a \to 0$, we may consider the case that $|a| < 1$. Consequently, we deduce $\|\tilde{A}_a\| = |a|^{\frac{1}{2}}$. Furthermore, we have for all $s \in \mathbb{R}$ that

$$\|\tilde{S}_s\|^2 = \frac{1}{2}s^2 + 1 + \frac{1}{2}\sqrt{s^4+4s^2}. \tag{3.11}$$

From these observations, it follows that:

$$\begin{aligned} |S^{\mathrm{v}}_{b,s,a}\psi(x)| &\leqslant C_N|a|^{-\frac{3}{4}}(1+|\tilde{A}_a^{-1}\tilde{S}_s^{-1}(x-b)|)^{-N} \\ &\leqslant C_N|a|^{-\frac{3}{4}}(1+\|\tilde{A}_a\|^{-1}\|\tilde{S}_s\|^{-1}|x-b|)^{-N} && (3.12) \\ &= C_N|a|^{-\frac{3}{4}}(1+|a|^{-\frac{1}{2}}\|\tilde{S}_s\|^{-1}|x-b|)^{-N} && (3.13) \\ &= C_N|a|^{-\frac{3}{4}+\frac{N}{2}}\|\tilde{S}_s\|^N(\|\tilde{S}_s\||a|^{\frac{1}{2}}+|x-b|)^{-N}. \end{aligned}$$

Combining the above estimate with the definition of the vertical shearlet transform, see (2.24), gives that for every $b \in U_{b_0}$ and all $s \in V_{s_0}$ we have

$$\begin{aligned} |\mathcal{SH}^{\mathrm{v}}_{\psi}((1-\phi)f)(b,s,a)| &= |\langle (1-\phi)f, S^{\mathrm{v}}_{b,s,a}\psi\rangle| \\ &\leqslant \int_{\mathbb{R}^2} |1-\phi(x)||f(x)||S^{\mathrm{v}}_{b,s,a}\psi(x)|\mathrm{d}x \\ &\leqslant C_N|a|^{-\frac{3}{4}+\frac{1}{2}N}\|\tilde{S}_s\|^N\int_{\mathbb{R}^2}|1-\phi(x)||f(x)|(\|\tilde{S}_s\||a|^{\frac{1}{2}}+|x-b|)^{-N}\mathrm{d}x \\ &\leqslant C_N|a|^{-\frac{3}{4}+\frac{1}{2}N}\|\tilde{S}_s\|^N\int_{\mathbb{R}^2}|1-\phi(x)||f(x)||x-b|^{-N}\mathrm{d}x \\ &= C_N|a|^{-\frac{3}{4}+\frac{1}{2}N}\|\tilde{S}_s\|^N\int_{|x-b|>\delta}|1-\phi(x)||f(x)||x-b|^{-N}\mathrm{d}x \\ &\leqslant C_N|a|^{-\frac{3}{4}+\frac{1}{2}N}\|\tilde{S}_s\|^N\|1-\phi\|_{L^\infty}\|f\|_{L^2}\left(\int_{|x-b|>\delta}|x-b|^{-2N}\mathrm{d}x\right)^{\frac{1}{2}} \\ &\leqslant C|a|^{-\frac{3}{4}+\frac{1}{2}N}, \end{aligned}$$

[1] The spectral norm of a matrix M is defined as the square root of the maximum eigenvalue of $M^{\dagger}M$, where $M^{\dagger}$ is the Hermitian transpose.

for some C which is independent of b and v. Note that we assumed that $s \in V_{s_0}$ and therefore there is some $\varepsilon > 0$ such that $s \leqslant \varepsilon + s_0$ and thus $\|\tilde{S}_s\|^2 \leqslant C_{s_0,\varepsilon}$. This gives an estimate of the shearlet transform of the function $(1-\phi)f$.

We will now estimate the shearlet coefficients of the function ϕf. In our analysis, Theorem 2.39 plays a vital role. In the setting of this theorem, we may assume without loss of generality that $\operatorname{supp}(\phi_2) \subsetneqq [-1,1]$. Then $\phi_2((v-s)|a|^{-1/2})$ equals zero outside the interval $(s-|a|^{\frac{1}{2}}, s+|a|^{\frac{1}{2}})$. Since we are interested in the case that $a \to 0$, we have for sufficiently small a that $(s-|a|^{\frac{1}{2}}, s+|a|^{\frac{1}{2}}) \subsetneqq V_{s_0}$. We will work with this a. Note that $\phi f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, so Theorem 2.39 is applicable and consequently for every $b \in U_{b_0}$ and $s \in V_{s_0}$ we obtain

$$
\begin{aligned}
&|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f)(b,s,a)| \\
&\leqslant |a|^{-\frac{3}{4}} \int_{\mathbb{R}^2} |\mathscr{W}_{\tilde{\chi}}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(n(v)\cdot b, a)| |\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}v && \text{(By Theorem 2.39)} \\
&= |a|^{-\frac{3}{4}} \int_{V_{s_0}} |\mathscr{W}_{\tilde{\chi}}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,t))(n(v)\cdot b, a)| |\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}v && \text{(Porperty of } \phi_2\text{)} \\
&\leqslant C_N |a|^{-\frac{3}{4}+N} \int_{V_{s_0}} |\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}v && \text{(By (3.10))} \\
&\leqslant C_N \sqrt{2} |a|^N \|\phi_2\|_{L^2} && \text{(Cauchy-Schwarz)} \\
&\leqslant D|a|^N && \text{(Since } \phi \in \mathscr{C}_0^\infty\text{)}
\end{aligned}
$$

Here is D a constant independent of b and v. Combining both results, we obtain

$$
\begin{aligned}
|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}} f(b,s,a)| &\leqslant |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(1-\phi)f(b,s,a)| + |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)| \\
&\leqslant C|a|^{-\frac{3}{4}+\frac{1}{2}N} + D|a|^N.
\end{aligned}
$$

Since this equality holds for every N we deduce that $(b,s) \in \mathscr{D}_2$ and hence we have shown the 'Only if' part of Theorem 3.5. □

3.4 'If' part of the proof of Theorem 3.5

Suppose that $(b,s) \in \mathscr{D}_1 \cup \mathscr{D}_2$. If $(b,s) \in \mathscr{D}_1$, then we are done, since this case was already shown in Bartolucci (2019). So assume $(b,s) \in \mathscr{D}_2$. We will show that $(b,s) \in \mathrm{WF}(f)^c$. To obtain this result, we start with introducing the definition of a cone.

Definition 3.7 A set $\Gamma \subsetneq \mathbb{R}^2$ is called a *cone* if for every $x \in \Gamma$ and for every $\lambda \in \mathbb{R} \setminus \{0\}$ we have $\lambda x \in \Gamma$. The set $\Gamma_{\xi_0} \cap \overline{B(0,1)}^c$ will be denoted by $\Gamma^*_{\xi_0}$.

Figure 3.7 shows an illustration of the cone $\Gamma^*_{\xi_0}$.

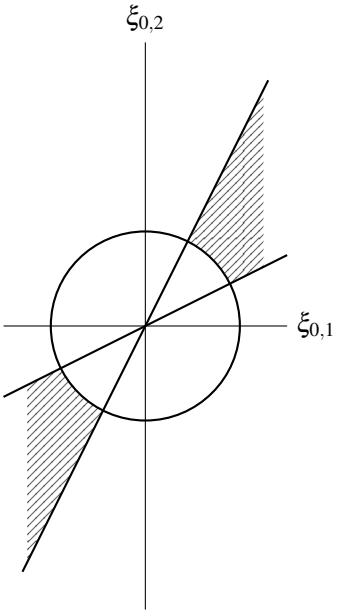


Figure 3.7: The cone $\Gamma^*_{\xi_0}$

We make the following remark, which will be used later on in the thesis.

Remark 3.8 For every $\xi_0 \in \Gamma_{\xi_0}$, where $\xi_0 = (\xi_{0,1}, \xi_{0,2})$, we have that the ratio of $\xi_{0,1}$ and $\xi_{0,2}$, is constant. In other words, we have that $\xi_{0,1}/\xi_{0,2} \leqslant M$ for some $M > 0$.

To prove the 'If' part of Theorem 3.5, we will do some preparations. First, in Lemma 3.9 we show that a point (x_0, ξ_0) is a regular directed point of f if for every $\varphi \in L^1(\Gamma^*_{\xi_0})$ we have

$$\left| \int_{\Gamma^*_{\xi_0}} |\xi|^N \mathscr{F}(\phi f)(\xi)\, \varphi(\xi)\, d\xi \right| \leqslant C_N \|\varphi\|_{L^1(\Gamma^*_{\xi_0})}. \tag{3.14}$$

After proving this particular estimate, our proof is in reach.

Lemma 3.9 *Let $f \in \mathscr{D}'(\mathbb{R}^2)$. The following statements are equivalent:*

(i) *A point $(x_0, \xi_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is a regular directed point of f if there exist a neighborhood U_{x_0} of x_0, a function $\phi \in \mathscr{C}^\infty_0(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$, and a conic neighborhood Γ_{ξ_0} of ξ_0 such that, for every $N \in \mathbb{N}$, there exists a constant C_N such that*

$$|\mathscr{F}(\phi f)(\xi)| \leqslant C_N (1 + |\xi|)^{-N}$$

for all $\xi \in \Gamma_{\xi_0}$.

(ii) *A point $(x_0, \xi_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is a regular directed point of f if there exist a neighborhood U_{x_0} of x_0, a function $\phi \in \mathscr{C}^\infty_0(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$, and a conic*

neighborhood Γ_{ξ_0} *of* ξ_0 *such that, for every* $N \in \mathbb{N}$*, there exists a constant* $C_N > 0$ *such that*

$$|\mathscr{F}(\phi f)(\xi)| \leqslant C_N |\xi|^{-N}$$

for all $\xi \in \Gamma^*_{\xi_0} = \Gamma_{\xi_0} \cap \overline{B(0,1)}^c$*, with* $B(0,1)$ *the ball of center* 0 *and radius* 1.

(iii) *A point* $(x_0, \xi_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ *is a regular directed point of* f *if there exist a neighborhood* U_{x_0} *of* x_0*, a function* $\phi \in \mathscr{C}^\infty_0(\mathbb{R}^2)$ *satisfying* $\phi(x_0) \neq 0$ *and a conic neighborhood* Γ_{ξ_0} *of* ξ_0 *such that, for every* $N \in \mathbb{N}$*, the functional*

$$\varphi \mapsto \int_{\Gamma^*_{\xi_0}} |\xi|^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\,\mathrm{d}\xi$$

is continuous on $L^1(\Gamma^*_{\xi_0})$*, i.e.,*

$$\left| \int_{\Gamma^*_{\xi_0}} |\xi|^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\,\mathrm{d}\xi \right| \leqslant C_N \|\varphi\|_{L^1(\Gamma^*_{\xi_0})},$$

for every $\varphi \in L^1(\Gamma^*_{\xi_0})$.

Proof We start with proving (i) $\Rightarrow$ (ii). Take $f \in \mathscr{D}'(\mathbb{R}^2)$ and assume that the point $(x_0, \xi_0) \in \mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ is a regular directed point of f. Then, by our assumption, there exists a neighborhood U_{x_0} of x_0, a function $\phi \in \mathscr{C}^\infty_0(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$, and a conic neighborhood Γ_{ξ_0} of ξ_0 such that, for every $N \in \mathbb{N}$, there exists a constant $C_N > 0$ such that

$$|\mathscr{F}(\phi f)(\xi)| \leqslant C_N (1 + |\xi|)^{-N}$$

for all $\xi \in \Gamma_{\xi_0}$. So in particular it holds for every $\xi \in \Gamma^*_{\xi_0}$. Since $\xi \in \Gamma^*_{\xi_0}$, we have $|\xi| > 1$. Consequently, for every $N \in \mathbb{N}$ we have

$$|\mathscr{F}(\phi f)(\xi)| \leqslant C_N (1 + |\xi|)^{-N} \leqslant C_N |\xi|^{-N},$$

which proves the first implication.

Now we prove (ii) $\Rightarrow$ (i). Let $f \in \mathscr{D}'(\mathbb{R}^2)$ and let (x_0, ξ_0) in $\mathbb{R}^2 \times \mathbb{R} \setminus \{0\}$ be a regular directed point of f. Suppose that there is a neighborhood U_{x_0} of x_0, a function $\phi \in \mathscr{C}^\infty_0(\mathbb{R}^2)$ satisfying $\phi(x_0) \neq 0$, and a conic neighborhood Γ_{ξ_0} of ξ_0 such that, for every $N \in \mathbb{N}$, there exists a constant $C_N > 0$ such that

$$|\mathscr{F}(\phi f)(\xi)| \lesssim |\xi|^{-N}$$

for all $\xi \in \Gamma^*_{\xi_0} = \Gamma_{\xi_0} \cap \overline{B(0,1)}^c$. Then, for $|\xi| > 1$ we have $(1 + |\xi|) \leqslant 2|\xi|$ and consequently we have

$$|\mathscr{F}(\phi f)(\xi)| \lesssim |\xi|^{-N} \lesssim (1 + |\xi|)^{-N}.$$

Now consider the case $|\xi| \leqslant 1$. Note that the Fourier transform of ϕf is essentially bounded and consequently there exists an $M > 0$ such that $|\mathscr{F}(\phi f)(\xi)| \leqslant M$. So for every $N \in \mathbb{N}$ we have

$$|\mathscr{F}(\phi f)(\xi)| \leqslant M \cdot \frac{(1+|\xi|)^N}{(1+|\xi|)^N} \leqslant \frac{2M}{(1+|\xi|)^N},$$

in the last step we used our assumption that $|\xi| \leqslant 1$.

Finally, we show that (ii) $\Leftrightarrow$ (iii). Note that

$$|\mathscr{F}(\phi f)(\xi)| \leqslant C_N |\xi|^{-N}$$

holds for every $\xi \in \Gamma^*_{\xi_0}$ if and only if the functional $\xi \mapsto |\xi|^N \mathscr{F}(\phi f)(\xi)$ is bounded. Since $\xi \mapsto |\xi|^N \mathscr{F}(\phi f)(\xi)$ belongs to $L^\infty(\Gamma^*_{\xi_0})$, this is equivalent with requiring that the functional

$$\varphi \mapsto \int_{\Gamma_{\xi_0}} |\xi|^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\mathrm{d}\xi$$

is continuous on $L^1(\Gamma_{\xi_0})$, in other words, we require that for every $\varphi \in L^1(\Gamma^*_{\xi_0})$ we have

$$\left| \int_{\Gamma^*_{\xi_0}} |\xi|^N \mathscr{F}(\phi f)(\xi)\, \varphi(\xi)\, d\xi \right| \leqslant C_N \|\varphi\|_{L^1(\Gamma^*_{\xi_0})}. \tag{3.15}$$

□

Lemma 3.10 *Let $f \in L^2(\mathbb{R}^2)$ and $\varphi \in \mathscr{C}_0(\Gamma^*_{\xi_0})$ we have*

$$\left| \int_{\Gamma^*_{\xi_0}} \xi^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\mathrm{d}\xi \right| \leqslant \int_{\mathbb{R}^2} \int_{\mathbb{R}} \int_{\mathbb{R}^+} |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f)(b,s,a)||a|^{-\frac{1}{4}-N} \tag{3.16}$$
$$\times \int_{\Gamma^*_{\xi_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2 \frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3}.$$

Here is $\tilde{\chi}_1$ a wavelet that satisfies the relationship $\mathscr{F}(\chi_1^{2N})(\tau) = |\tau|^N \mathscr{F}(\tilde{\chi}_1)(\tau)$.

Proof Suppose that $f \in L^2(\mathbb{R}^2)$ and $\varphi \in \mathscr{C}_0(\Gamma^*_{\xi_0})$. Observe that for any ψ^{v} we have

$$\begin{aligned}
\int_{\Gamma^*_{\xi_0}} \xi^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\mathrm{d}\xi &= \int_{\mathbb{R}^2} \xi^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\mathrm{d}\xi && \text{(Since } \varphi \in \mathscr{C}_0(\Gamma^*_{\xi_0})\text{)} \\
&= \langle \mathscr{F}(\phi f), |\cdot|^N \overline{\varphi}\rangle && \text{(Definition inner product)} \\
&= \langle \phi f, \mathscr{F}^{-1}|\cdot|^N \overline{\varphi}\rangle && \text{(Fourier transform is unitary)} \\
&= \langle \mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f), \mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\mathscr{F}^{-1}|\cdot|^N \overline{\varphi})\rangle && \text{(Shearlet transform is an isometry)}
\end{aligned}$$

and consequently, we obtain

$$\begin{aligned}&\int_{\Gamma^*_{\xi_0}} \xi^N \mathscr{F}(\phi f)(\xi)\varphi(\xi)\mathrm{d}\xi\\&\quad= \int_{\mathbb{R}^2}\int_{\mathbb{R}}\int_{\mathbb{R}^+} \mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)\overline{\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\mathscr{F}^{-1}|\cdot|^N\overline{\phi})(b,s,a)}\frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3}.\end{aligned} \tag{3.17}$$

We will estimate $\overline{\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\mathscr{F}^{-1}|\cdot|^N\overline{\phi})(b,s,a)}$ to obtain the estimate displayed in (3.16).

By Formula (2.27) we have

$$\begin{aligned}&\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\mathscr{F}^{-1}(\cdot)^N\overline{\varphi})(b,s,a)\\&\quad= |a|^{-\frac{3}{4}}\int_{\mathbb{R}} \mathscr{W}_{\chi_1}(\mathscr{R}^{\mathrm{v}}(\mathscr{F}^{-1}(\cdot)^N\overline{\varphi})(v,\cdot))(n(v)\cdot b,a)\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)\mathrm{d}v.\end{aligned} \tag{3.18}$$

Note that

$$\begin{aligned}&\mathscr{W}_{\chi_1}(\mathscr{R}^{\mathrm{v}}(\mathscr{F}^{-1}|\cdot|^N\overline{\varphi})(v,\cdot))(n(v)\cdot b,a)\\&\quad= \langle\mathscr{R}^{\mathrm{v}}(\mathscr{F}^{-1}|\cdot|^N\overline{\varphi})(v,\cdot), W_{n(v)\cdot b,a}\chi_1\rangle && \text{(Def. wavelet transform)}\\&\quad= \langle\mathscr{F}(\mathscr{R}^{\mathrm{v}}(\mathscr{F}^{-1}|\cdot|^N\overline{\varphi})(v,\cdot)), \mathscr{F}(W_{n(v)\cdot b,a}\chi_1)\rangle && \text{(Parseval)}\\&\quad= \langle|n(v)|^N\overline{\varphi}(\cdot n(v)), \mathscr{F}(W_{n(v)\cdot b,a}\chi_1)\rangle && \text{(Fourier slice theorem)}\\&\quad= (1+v^2)^{\frac{N}{2}}\langle\overline{\varphi}(\cdot n(v)), \mathscr{F}(W_{n(v)\cdot b,a}\chi_1)\rangle. && \text{(Rewriting)}\end{aligned}$$

We have

$$\begin{aligned}|\langle\overline{\varphi}(\cdot n(v)), \mathscr{F}(W_{n(v)\cdot b,a}\chi_1)\rangle| &\leqslant \int_{\mathbb{R}} |\overline{\varphi}(\xi n(v))||\overline{\mathscr{F}(W_{n(v)\cdot b,a}\chi_1(\xi))}|\mathrm{d}\xi\\&= |a|^{\frac{1}{2}}\int_{\mathbb{R}} |\overline{\varphi}(\xi n(v))||e^{2\pi i\xi n(v)\cdot b}||\overline{\mathscr{F}\chi_1(a\xi)}|\mathrm{d}\xi && \text{(By (2.14))}\\&= |a|^{\frac{1}{2}-N}\int_{\mathbb{R}} |\overline{\varphi(\xi n(v))}||a\xi|^N|\overline{\mathscr{F}\chi_1(a\xi)}|\mathrm{d}\xi && \text{(Rewriting)}\\&= |a|^{\frac{1}{2}-N}\int_{\mathbb{R}} |\varphi(\xi n(v))||\overline{\mathscr{F}\tilde{\chi}_1(a\xi)}|\mathrm{d}\xi, && \text{(Defenition } \tilde{\chi}_1\text{)}\end{aligned}$$

where in the last line we defined $\tilde{\chi}_1$ such that $\mathscr{F}\tilde{\chi}_1(\xi) = |\xi|^N\mathscr{F}\chi_1(\xi)$ holds. Combining the latter estimate together with (3.18) yields:

$$\begin{aligned}&|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\mathscr{F}^{-1}(\cdot)^N\overline{\varphi})(b,s,a)|\\&\quad\leqslant |a|^{-\frac{1}{4}-N}\int_{\mathbb{R}}\int_{\mathbb{R}} (1+v^2)^{\frac{N}{2}}|\varphi(\xi n(v))||\mathscr{F}(\tilde{\chi}_1(a\xi)||\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi\mathrm{d}v.\end{aligned}$$

Using the substitutions $v \mapsto \xi_1/\xi_2$ and $\xi \mapsto \xi_2$ we obtain

$$\varphi(\xi n(v)) = \varphi(\xi(v,1)) = \varphi(\xi_2(\xi_1/\xi_2,1)) = \varphi(\xi_1,\xi_2) \tag{3.19}$$

and consequently

$$\begin{aligned}
&|\mathcal{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\mathcal{F}^{-1}(\cdot)^N\overline{\varphi})(b,s,a)| \\
&\leqslant |a|^{-\frac{1}{4}-N}\int_{\mathbb{R}}\int_{\mathbb{R}}\left(1+\frac{\xi_1^2}{\xi_2^2}\right)^{\frac{N}{2}}|\varphi(\xi_1,\xi_2)||\mathcal{F}(\tilde{\chi}_1(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)||\xi_2|^{-1}\mathrm{d}\xi_1\mathrm{d}\xi_2 \\
&\leqslant C|a|^{-\frac{1}{4}-N}\int_{\Gamma^*_{\xi_0}}|\varphi(\xi_1,\xi_2)||\mathcal{F}(\tilde{\chi}_1(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2.
\end{aligned}$$

In the second inequality, we used the fact that φ is compactly supported in $\Gamma^*_{\xi_0}$ and therefore we know that $|\xi_2| \geqslant 1$. By Remark 3.8 we know that $(\xi_1,\xi_2) \mapsto (1+\xi_1^2/\xi_2^2)^{\frac{N}{2}} \leqslant C$ holds for some positive C. Combining the last inequality and (3.17) gives the desired formula (3.16). □

Finally, we need the following result, whose proof can be found in Kutyniok and Labate (2009).

Lemma 3.11 *Let $\psi^{\mathrm{v}} \in L^2(\mathbb{R}^2)$, which is in the form as described in Section 2.7, where $\psi_1 \in L^2(\mathbb{R})$ is a wavelet, $\mathcal{F}\psi_1 \in \mathcal{C}_0^\infty(\mathbb{R})$ such that* $\mathrm{supp}(\mathcal{F}\psi_1) \subsetneq [-2,-1/2]\cup[1/2,2]$. *Furthermore, on ψ_2 we impose the following restrictions: $\|\psi_2\|_{L^2}=1$, $\mathcal{F}\psi_2 \in \mathcal{C}_0^\infty(\mathbb{R})$ such that* $\mathrm{supp}(\mathcal{F}\psi_2) \subsetneq [-1,1]$ *and $\mathcal{F}\psi_2 > 0$ on the interval $(-1,1)$.*
Then for $f \in L^2(\mathbb{R}^2)$ and $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^2)$ we have for every $N \in \mathbb{N}$ that there exists a constant C_N such that

$$|\mathcal{SH}^{v}_{\psi^{\mathrm{v}}}\phi f(b,s,a)| \leqslant C_N|a|^N. \tag{3.20}$$

Remark 3.12 From Lemma 3.11 we see that the imposed conditions for the wavelets are quite strict. In our proof we don't need these specific conditions; the only thing that we need in our proof is the estimate (3.20). It is not known if (3.20) also holds for our (less restrictive) choice of ψ^{v}. This is a nice question for follow-up research.

Now we have set up our preparations, we are ready to prove the 'If' part of the proof of Theorem 3.5.

Proof 'If' part of Theorem 3.5 By Lemma 3.9 part (iii), it suffices to show that there exist a neighborhood U'_{b_0} of b_0, a function $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^2)$ such that $\phi(x_0) \neq 0$ and a neighborhood V'_{s_0} of s_0 such that for every $L \in \mathbb{N}$ the functional

$$\varphi \mapsto \int_{\Gamma^*_{s_0}}|\xi|^L\mathcal{F}(\phi f)(\xi)\varphi(\xi)\mathrm{d}\xi$$

is continuous on $L^1(\Gamma^*_{s_0})$. Here is $\Gamma^*_{s_0} = \Gamma_{s_0} \cap \overline{B(0,1)}^c$, where Γ_{s_0} is the cone parametrized by the interval V'_{s_0}. Choose $\varepsilon'_0 > 0$ such that $V'_{s_0} := (s_0 - \varepsilon', s_0 + \varepsilon')$ is contained in $V_{s_0} = (s_0 - \varepsilon, s_0 + \varepsilon)$ in such a way that there is always a positive distance between the boundary's of V'_{s_0} and V_{s_0}. Finally, let $\phi \in \mathscr{C}_0^\infty(\mathbb{R}^2)$ such that $\operatorname{supp}(\phi) \subsetneqq U'_{b_0} \subsetneqq U_{b_0}$.

By Formula (3.15) and Lemma 3.10 it suffices to show that for every $L > 0$ there exists a constant $C_L > 0$ such that

$$\int_{\mathbb{R}^2} \int_{\mathbb{R}} \int_{\mathbb{R}^+} |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}} \phi f(b,s,a)||a|^{-\frac{1}{4}-N} \tag{3.21}$$
$$\times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right) |\mathrm{d}\xi_1 \mathrm{d}\xi_2 \frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \leqslant C_L \|\varphi\|_{L^1(\Gamma^*_{s_0})},$$

where φ is in $\mathscr{C}_0(\Gamma^*_{s_0})$. We split the above integral in eight integrals, call them $I_1, I_2, \ldots, I_8$, and analyze these in eight different regions. See Table 3.1 for an overview. In the following estimates, we consider C as a general positive constant which may vary from expression to expression.

Integral	a	s	b
I_1	$\lvert a\rvert < 1$	$s \in V_{s_0}$	$b \in U'_{b_0}$
I_2	$\lvert a\rvert \geqslant 1$	$s \in V_{s_0}$	$b \in U'_{b_0}$
I_3	$\lvert a\rvert < 1$	$s \in V^c_{s_0}$	$b \in U'_{b_0}$
I_4	$\lvert a\rvert \geqslant 1$	$s \in V^c_{s_0}$	$b \in U'_{b_0}$
I_5	$\lvert a\rvert < 1$	$s \in V_{s_0}$	$b \in U'^c_{b_0}$
I_6	$\lvert a\rvert \geqslant 1$	$s \in V_{s_0}$	$b \in U'^c_{b_0}$
I_7	$\lvert a\rvert < 1$	$s \in V^c_{s_0}$	$b \in U'^c_{b_0}$
I_8	$\lvert a\rvert \geqslant 1$	$s \in V^c_{s_0}$	$b \in U'^c_{b_0}$

Table 3.1 *Overview of the split integrals*

We start with estimating the first integral, I_1. Since Lemma 3.11 holds for every $N > 0$, we may choose N such that $N > 2\frac{1}{4} + L$. In this case, we have

$$\begin{aligned} I_1 &= \int_{U'_{b_0}} \int_{s_0-\varepsilon}^{s_0+\varepsilon} \int_{|a|<1} |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f)(b,s,a)||a|^{-\frac{1}{4}-L} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right) |\mathrm{d}\xi_1 \mathrm{d}\xi_2 \frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \\ &\leqslant C \int_{U'_{b_0}} \int_{s_0-\varepsilon}^{s_0+\varepsilon} \int_{|a|<1} |a|^{N-3\frac{1}{4}-L} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right) |\mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}a\mathrm{d}s\mathrm{d}b \quad \text{(By Lemma 3.11)} \end{aligned}$$

$$\leqslant C \int_{U'_{b_0}} \int_{s_0-\varepsilon}^{s_0+\varepsilon} \int_{|a|<1} |a|^{N-3\frac{1}{4}-L} \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)| \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}a \mathrm{d}s \mathrm{d}b$$
$$\leqslant C \|\varphi\|_{L^1(\Gamma^*_{s_0})}.$$

In the preceding inequality we made use of the fact that ϕ_2 is compactly supported in $\Gamma^*_{s_0}$, and hence bounded, and the fact that $\mathscr{F}(\tilde{\chi}_1)$ is bounded. The latter is a consequence of the fact that we imposed that χ_1 is in $\mathscr{S}_0(\mathbb{R})$ and consequently it follows that $\tilde{\chi}_1$ is in $\mathscr{S}_0(\mathbb{R})$ too.

For I_2 we select N such that $N < 2\frac{1}{4} + L$, which is again always possible due to the fact that Lemma 3.11 holds for every N. A similar calculation as above shows that $I_2 \leqslant C\|\varphi\|_{L^1(\Gamma^*_{s_0})}$.

Now we will estimate I_3. We first note that by Lemma 2.27 we have

$$|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)| \leqslant |a|^{-\frac{3}{4}} \int_{\mathbb{R}} |\mathscr{W}_{\chi_1}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,\cdot))(n(v)\cdot b,a)| |\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}v. \tag{3.22}$$

We have

$$\begin{aligned}
&|\mathscr{W}_{\chi_1}(\mathscr{R}^{\mathrm{v}}(\phi f)(v,\cdot))(n(v)\cdot b,a)| && (3.23)\\
&= |\langle \mathscr{R}^{\mathrm{v}}(\phi f)(v,\cdot)), W_{n(v)\cdot b,a}\chi_1\rangle| && \text{(Def. Wavelet transform)}\\
&= |\langle \mathscr{F}\mathscr{R}^{\mathrm{v}}(\phi f)(v,\cdot)), \mathscr{F}(W_{n(v)\cdot b,a}\chi_1)\rangle| && \text{(Parsevel)}\\
&\leqslant \int_{\mathbb{R}} |\mathscr{F}(\phi f)(v\xi)| \overline{|\mathscr{F}(W_{n(v)\cdot b,a}\chi_1)(\xi)|} \mathrm{d}\xi && \text{(Fourier slice theorem)}\\
&\leqslant |a|^{\frac{1}{2}} \|\mathscr{F}(\phi f)\|_{L^\infty(\mathbb{R})} \int_{\mathbb{R}} |\mathscr{F}(\chi_1)(a\xi)| \mathrm{d}\xi && \text{(By (2.14))}\\
&= |a|^{-\frac{1}{2}} \|\mathscr{F}(\phi f)\|_{L^\infty(\mathbb{R})} \|\mathscr{F}\chi_1\|_{L^1(\mathbb{R})} && (3.24)
\end{aligned}$$

Combining (3.22) and (3.24) yields:

$$|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)| \leqslant |a|^{-1\frac{1}{4}} |\mathscr{F}(\phi f)\|_{L^\infty(\mathbb{R})} \|\mathscr{F}\chi_1\|_{L^1(\mathbb{R})} \int_{\mathbb{R}} |\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}v.$$

Since we assumed that $\mathrm{supp}(\phi_2) \subseteqq [-1,1]$, we obtain

$$|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)| \leqslant |a|^{-1\frac{1}{4}} |\mathscr{F}(\phi f)\|_{L^\infty(\mathbb{R})} \|\mathscr{F}\chi_1\|_{L^1(\mathbb{R})} \int_{s-|a|^{\frac{1}{2}}}^{s+|a|^{\frac{1}{2}}} |\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}v.$$

By the Cauchy-Schwarz inequality we finally obtain

$$|\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}\phi f(b,s,a)| \leqslant \sqrt{2}|a|^{-\frac{3}{4}} |\mathscr{F}(\phi f)\|_{L^\infty(\mathbb{R})} \|\mathscr{F}\chi_1\|_{L^1(\mathbb{R})} \|\phi_2\|_{L^2(\mathbb{R})}.$$

This inequality implies that

$$\begin{aligned} I_3 &= \int_{U'_{b_0}} \int_{|s-s_0|>\varepsilon} \int_{|a|<1} |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f)(b,s,a)||a|^{-\frac{1}{4}-L} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}\xi_1 \mathrm{d}\xi_2 \frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \\ &\leqslant C \int_{U'_{b_0}} \int_{|s-s_0|>\varepsilon} \int_{|a|<1} |a|^{-4-L} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}a\mathrm{d}s\mathrm{d}b \end{aligned}$$

We know that $\phi_2 \in \mathscr{C}^\infty_0(\mathbb{R})$, and hence it is also in $\mathscr{S}(\mathbb{R})$ and thus, by (2.1), we have

$$|\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)| \leqslant |a|^{-\frac{N}{2}} \left|\frac{\xi_1}{\xi_2} - s\right|^{-N},$$

where we choose $N > 1$.

Furthermore, Lemma 2.27 implies that for every $M \in \mathbb{N}$ there exists a $\theta \in L^2(\mathbb{R})$ such that for every $\xi_2 \in \mathbb{R}$ we have

$$\mathscr{F}(\chi)(\xi_2) = \xi_2^M \mathscr{F}(\theta\xi_2),$$

here we choose $M > 3 + L + \frac{N}{2}$.

Furthermore, from these observations, we get the following estimate for I_3:

$$\begin{aligned} I_3 &\leqslant C \int_{U'_{b_0}} \int_{|s-s_0|>\varepsilon} \int_{|a|<1} |a|^{M-4-L-\frac{N}{2}} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\theta)(a\xi_2)||\xi_2|^M \left|\frac{\xi_1}{\xi_2} - s\right|^{-N} \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}a\mathrm{d}s\mathrm{d}b \end{aligned}$$

Since we have $v = \xi_1/\xi_2 \in V_{s_0} = (s_0 - \varepsilon', s_0 + \varepsilon')$ we have $|\xi_1/\xi_2 - s_0| < \varepsilon'$. Since we are integrating in the region $|s - s_0| > \varepsilon > \varepsilon' > 0$, we know that

$$0 < |s - s_0| - \varepsilon' < |s - s_0| - |\xi_1/\xi_2 - s_0|,$$

from which it follows:

$$|\xi_1/\xi_2 - s_0|^{-N} \leqslant ||s - s_0| - |\xi_1/\xi_2 - s_0||^{-N} < ||s - s_0| - \varepsilon'|^{-N}. \tag{3.25}$$

The first inequality is an application of the reverse triangle inequality. Hence we obtain

$$\begin{aligned} I_3 &\leqslant C \int_{U'_{b_0}} \int_{|s-s_0|>\varepsilon} \int_{|a|<1} |a|^{M-4-L-\frac{N}{2}} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\theta)(a\xi_2)||\xi_2|^M \left||s-s_0|-\varepsilon'\right|^{-N} \mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}a\mathrm{d}s\mathrm{d}b \\ &\leqslant C\|\varphi\|_{L^1(\Gamma^*_{s_0})}. \end{aligned}$$

A similar calculation shows that $I_4 \leqslant C\|\varphi\|_{L^1(\Gamma^*_{s_0})}$, but then we choose M such that $M < 3+L+\frac{N}{2}$.

Now will estimate I_5.

$$\begin{aligned} I_5 &= \int_{U'^c_{b_0}} \int_{|s-s_0|\leqslant\varepsilon} \int_{|a|<1} |\mathscr{S}\mathscr{H}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f)(b,s,a)||a|^{-\frac{1}{4}-L} \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2\frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \\ &\leqslant \int_{U'^c_{b_0}} \int_{|s-s_0|\leqslant\varepsilon} \int_{|a|<1} |a|^{-\frac{1}{4}-L} \int_{\mathrm{supp}(\phi)} |\phi f(x)||S^{\mathrm{v}}_{b,s,a}\psi^{\mathrm{v}}(x)|\mathrm{d}x \\ &\quad \times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2\frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \end{aligned}$$

Note that we are integrating in the region $|a|<1$, so we are in the setting of Equation (3.13), so we get, for $N > 6-2L$, that

$$|S^{\mathrm{v}}_{b,s,a}\psi^{\mathrm{v}}(x)| \leqslant C_N|a|^{\frac{N}{2}-\frac{3}{4}}(|a|^{\frac{1}{2}}+\|\tilde{S}_s\|^{-1}|x-b|)^{-N} \leqslant C_N|a|^{\frac{N}{2}-\frac{3}{4}}\|\tilde{S}_s\|^N|x-b|^{-N}.$$

Equation (3.11) implies that

$$\|\tilde{S}_s\|^2 = \frac{1}{2}s^2+1+\frac{1}{2}\sqrt{s^4+4s^2}.$$

Since $s \leqslant \varepsilon+s_0$, we deduce that $\|\tilde{S}_s\|^2 \leqslant C_{s_0,\varepsilon}$. Hence

$$|S^{\mathrm{v}}_{b,s,a}\psi^{\mathrm{v}}(x)| \leqslant C|a|^{\frac{N}{2}-\frac{3}{4}}|x-b|^{-N}.$$

Since $b \in U'^c_{b_0}$ and $x \in \mathrm{supp}(\phi) \in U'_{b_0}$, we can always find a $C>0$ such that $|x-b| \geqslant C|b_0-b|$. Consequently, we obtain

$$|S^{\mathrm{v}}_{b,s,a}\psi^{\mathrm{v}}(x)| \leqslant C|a|^{\frac{N}{2}-\frac{3}{4}}|b_0-b|^{-N}$$

and thus

$$\begin{aligned}
I_5 &\leqslant \int_{U_{b_0}^{tc}} \int_{|s-s_0|\leqslant\varepsilon} \int_{|a|<1} |a|^{\frac{N}{2}-4-L} \int_{\mathrm{supp}(\phi)} |\phi f(x)||b_0-b|^{-N}\mathrm{d}x \\
&\quad \times \int_{\Gamma_{s_0}^*} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}a\mathrm{d}s\mathrm{d}b \\
&\leqslant \int_{U_{b_0}^{tc}} \int_{|s-s_0|\leqslant\varepsilon} \int_{|a|<1} |a|^{\frac{N}{2}-4-L}\|\phi f\|_{L^1(\mathbb{R}^2)}|b_0-b|^{-N} \\
&\quad \times \int_{\Gamma_{s_0}^*} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}a\mathrm{d}s\mathrm{d}b \\
&\leqslant C\|\varphi\|_{L^1(\Gamma_{s_0}^*)}.
\end{aligned}$$

For I_6 we have $|a| \geqslant 1$ and hence Equation (3.12) implies that for $N > \frac{L}{2}+\frac{3}{2}$ we have

$$\begin{aligned}
|S_{b,s,a}^{\mathrm{v}}\psi^{\mathrm{v}}(x)| &\leqslant C_N|a|^{-\frac{3}{4}}(1+|a|^{-2}\|\tilde{S}_s\|^{-1}|x-b|)^{-N} \\
&= C_N|a|^{2N-\frac{3}{4}}\|\tilde{S}_s\|^{-1}|x-b|^{-N}.
\end{aligned}$$

A similar argument as before, yields:

$$|S_{b,s,a}^{\mathrm{v}}\psi^{\mathrm{v}}(x)| \leqslant C|a|^{2N-\frac{3}{4}}|x-b|^{-N}.$$

In the same spirit as the estimates for I_5, we can show that $I_6 \leqslant C\|\varphi\|_{L^1(\Gamma_{s_0}^*)}$. The latter is a consequence of our choice for $N > \frac{L}{2}+\frac{3}{2}$.

Now we will estimate the last two integrals, I_7 and I_8. We start with estimating I_7. Equation (3.12) implies that

$$|S_{b,s,a}^{\mathrm{v}}\psi^{\mathrm{v}}(x)| \leqslant C_N|a|^{-\frac{3}{4}}(1+\|\tilde{A}_a\|^{-1}\|\tilde{S}_s\|^{-1}|x-b|)^{-N}.$$

Since we integrate in the region $|a|<1$ we know that $\|\tilde{A}_a\|$ equals $|a|^{\frac{1}{2}}$. Consequently:

$$|S_{b,s,a}^{\mathrm{v}}\psi^{\mathrm{v}}(x)| \leqslant C_N|a|^{\frac{N}{2}-\frac{3}{4}}\|\tilde{S}_s\|^{N}|x-b|^{-N}. \tag{3.26}$$

It therefore follows that

$$\begin{aligned}
I_7 &= \int_{U_{b_0}^{tc}} \int_{|s-s_0|\geqslant\varepsilon} \int_{|a|<1} |\mathscr{S}\mathscr{H}_{\psi^{\mathrm{v}}}^{\mathrm{v}}(\phi f)(b,s,a)||a|^{-\frac{1}{4}-L} \\
&\quad \times \int_{\Gamma_{s_0}^*} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2-s}{|a|^{\frac{1}{2}}}\right)|\mathrm{d}\xi_1\mathrm{d}\xi_2\frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \\
&\leqslant \int_{U_{b_0}^{tc}} \int_{|s-s_0|\geqslant\varepsilon} \int_{|a|<1} |a|^{-\frac{1}{4}-L} \int_{\mathrm{supp}(\phi)} |\phi f(x)||S_{b,s,a}^{\mathrm{v}}\phi^{\mathrm{v}}(x)|\mathrm{d}x
\end{aligned}$$

$$\times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}\xi_1\mathrm{d}\xi_2 \frac{\mathrm{d}a\mathrm{d}s\mathrm{d}b}{|a|^3} \quad \text{(Def. } \mathscr{SH}^{\mathrm{v}}_{\psi^{\mathrm{v}}}(\phi f)\text{)}$$

$$\leqslant \int_{U^{\prime c}_{b_0}} \int_{|s-s_0|\geqslant\varepsilon} \int_{|a|<1} |a|^{-4-L+\frac{N}{2}} \|\tilde{S}_s\|^N \int_{\mathrm{supp}(\phi)} |\phi f(x)||x-b|^{-N}\mathrm{d}x$$

$$\times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\tilde{\chi}_1)(a\xi_2)||\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)| \mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}a\mathrm{d}s\mathrm{d}b. \quad \text{(By (3.26)}$$

We have $b \in U^{\prime c}_{b_0}$ and $x \in \mathrm{supp}(\phi) \subsetneqq U^{\prime}_{b_0}$ we can find a C such that $|x-b| \geqslant C|b_0-b|$. Furthermore, we know that ϕ_2 has fast decay, so for every $\tilde{N}$, we can find a constant C such that

$$|\phi_2\left(\frac{\xi_1/\xi_2 - s}{|a|^{\frac{1}{2}}}\right)| \leqslant C|a|^{-\frac{\tilde{N}}{2}} \left|\frac{\xi_1}{\xi_2} - s\right|^{-\tilde{N}}.$$

Here we choose $\tilde{N} > 1+N$. We can do this, since the inequality above holds for every $\tilde{N}$. Finally, since $\tilde{\chi}_1 \in \mathscr{S}_0$, we know that there exists a $\theta \in \mathscr{S}_0$ such that for every $\xi_2 \in \mathbb{R}$ and every M we have $\mathscr{F}(\tilde{\chi}_1)(a\xi_2) = a^M \xi_2^M \mathscr{F}(\theta\xi_2)$. Here we choose M such that $M > 3 - L - \frac{N}{2} + \frac{\tilde{N}}{2}$.

Combining these three observations with our earlier estimates, yields:

$$I_7 \leqslant \int_{U^{\prime c}_{b_0}} \int_{|s-s_0|\geqslant\varepsilon} \int_{|a|<1} |a|^{M-4-L+\frac{N}{2}-\frac{\tilde{N}}{2}} |b_0-b|^{-N} \int_{\mathrm{supp}(\phi)} |\phi f(x)|\mathrm{d}x$$

$$\times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\theta)(a\xi_2)||\xi_2|^M \|\tilde{S}_s\|^N \left|\frac{\xi_1}{\xi_2} - s\right|^{-\tilde{N}} \mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}a\mathrm{d}s\mathrm{d}b$$

$$\leqslant \int_{U^{\prime c}_{b_0}} \int_{|s-s_0|\geqslant\varepsilon} \int_{|a|<1} |a|^{M-4-L+\frac{N}{2}-\frac{\tilde{N}}{2}} |b_0-b|^{-N} \|\phi f\|_{L^1}$$

$$\times \int_{\Gamma^*_{s_0}} |\varphi(\xi_1,\xi_2)||\mathscr{F}(\theta)(a\xi_2)||\xi_2|^M \|\tilde{S}_s\|^N |||s-s_0| - \varepsilon'|^{-\tilde{N}} \mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}a\mathrm{d}s\mathrm{d}b \quad \text{(By 3.25)}$$

$$\leqslant C\|\varphi\|_{L^1(\Gamma^*_{s_0})}.$$

The integrals above converge, due to our imposed choices of $\tilde{N}$ and M.

The final case I_8 can be treated similar. The only difference is that we choose M such that $M < 3 - L - \frac{N}{2} + \frac{\tilde{N}}{2}$.

Combining all eight cases leads to the proof of (3.21). Therefore, our theorem has been proven. □

4
Conclusions and discussion

In this final chapter, we give an overview of the results of this thesis. In Section 4.1 we state our conclusions from our literature review and summarize our contribution to the existing literature in the study of the roles of the Radon transform and the wavelet transform in the resolution of the wavefront set in shearlet analysis. Finally, in Section 4.2 we outline potential directions for future research based on our work.

4.1 Conclusion

The goal of this thesis was to study the microlocal properties of the shearlet transform via the Radon and wavelet transforms. This was done by first studying some well-known results from the literature. This literature review gave the following insights.

(i) The wavelet transform is capable of detecting pointwise singularities, see Theorem 2.29, but it generally falls short in capturing the wavefront set, as shown in Example 2.34. The main reason for this limitation lies in the isotropic nature of the wavelet system.

(ii) Theorem 2.25 shows that the Radon transform is able to determine whether a point belongs to the wavefront set.

(iii) The shearlet transform can characterize points in the wavefront set. This result was already established in Grohs (2010). The integral formula connecting the (vertical) shearlet transform with the wavelet and the (vertical) affine Radon transforms, as presented in Theorems 2.38 and 2.39, offers a new proof approach for this result. However, the theorem in Bartolucci (2019) characterizes almost the entire wavefront set, but not completely.

In our study, the results in Theorems 2.38 and 2.39 played a vital role, as they highlight

the connection between the (vertical) shearlet transform and both the wavelet and (vertical) Radon transforms. This relationship is fundamental to understanding how shearlets can be used effectively in microlocal analysis.

The novel contribution of this thesis to the literature is the formulation (and proof) of a general theorem that extends the results from Bartolucci (2019) to cover all possible directions in the wavefront set. In Theorem 3.5 we work with so-called 'cone-adapted' shearlets, i.e. we use the shearlet transform in a set $\mathscr{D}_1$ and we use the vertical shearlet transform in the set $\mathscr{D}_2$. This approach enables us to characterize the whole wavefront set. The proof was inspired by previous work of Bartolucci (2019), but we provided more motivation and intermediate steps.

Figure 4.1 presents a timeline illustrating the study of the wavefront set using the tools discussed in this thesis.

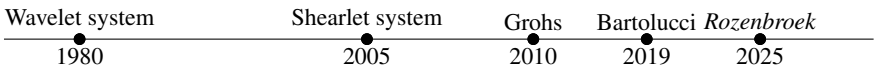


Figure 4.1 Timeline of the study of the wavefront set

4.2 Discussion and Suggestions for Follow-Up Research

In this thesis, we have established a deeper understanding of the connection between the shearlet transform and the wavefront set. However, several interesting questions remain unanswered. We provide an overview of these open questions:

(i) In Remark 3.4, we sketched an approach to prove directly, using the decay properties of the vertical affine Radon transform, that if a point (b_0, s_0) does not lie in the wavefront set of f, then the coefficients of the vertical shearlet of the function f at (b,s,a), with $b \in U_{b_0}$ and $s \in V_{s_0}$, decay rapidly as $a \to 0$. We gave an outline of an idea in Remark 3.4, but making it mathematically precise remains a topic for further research.

(ii) The estimate (3.13) in Lemma 3.11 is used in our work, but the conditions in which it is formulated in Lemma 3.11 are quite strict, and we do not use all these conditions. Future research could aim to prove (3.13) under weaker conditions on the wavelets, such that the estimate still holds.

(iii) In this thesis, we developed most of the results for functions in the space $L^2(\mathbb{R}^2)$. Follow-up research could extend our framework to the space of distributions $\mathscr{D}'(\mathbb{R}^2)$. This extension would provide deeper insights into the theory.

Answering these open problems could yield new insights into the microlocal properties of the shearlet transform, using the Radon and wavelet transform. These insights can be relevant for theoretical understanding, but also for more practical and applied contexts.

References

Antoine, Jean-Pierre, Demanet, Laurent, Jacques, Laurent, and Vandergheynst, Pierre. 1993. Wavelets on the 2-sphere: A group-theoretical approach. *Applied and Computational Harmonic Analysis*, **10**(3), 183–200. 5

Bamberger, Robert H, and Smith, Michael J. 1992. A filter bank for the directional decomposition of images: Theory and design. *IEEE transactions on signal processing*, **40**(4), 882–893. 5

Bartolucci, Francesca. 2019. *Radon transforms: Unitarization, Inversion and Wavefront sets*. Ph.D. thesis, Università degli Studi di Genova, Genova. PhD Programme in Mathematics and Applications, Supervised by Professor Filippo De Mari and Professor Ernesto De Vito. iii, 2, 6, 23, 34, 44, 45, 50, 61, 62

Boggess, Albert, and Narcowich, Francis J. 2009. *A First Course in Wavelets with Fourier Analysis*. Second edn. Hoboken, NJ: Wiley. 25

Boman, Jan. 2014. *Lecture Notes*. https://www.icts.res.in/sites/default/files/Jan_Boman_Lecture_Notes_0.pdf. 15

Candès, Emmanuel J, and Donoho, David L. 2004. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Communications on pure and applied mathematics*, **57**(2), 219–266. 5

Dahlke, Stephan, Kutyniok, Gitta, Steidl, Gabriele, and Teschke, Gerd. 2009. Shearlet coorbit spaces and associated Banach frames. *Journal of Fourier Analysis and Applications*, **15**(1), 1–27. 32

Daubechies, Ingrid. 1992. *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61. Society for Industrial and Applied Mathematics (SIAM). 3, 25

Deans, Stanley R. 1983. *The Radon Transform and Some of Its Applications*. New York: John Wiley & Sons. 3, 16

Debnath, Lokenath, and Shah, Firdous A. 2017. *Lecture Notes on Wavelet Transforms*. Birkhäuser. 25

Deng, Dejun, Geller, David, Walnut, David F., and Weiss, Guido (eds). 2009. *Four Short Courses on Harmonic Analysis: Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis*. Applied and Numerical Harmonic Analysis. Boston, MA: Birkhäuser. 4

Do, Minh N, and Vetterli, Martin. 2005. The contourlet transform: An efficient directional multiresolution image representation. *IEEE transactions on image processing*, **14**(12), 2091–2106. 5

Evans, Lawrence C., and Gariepy, Ronald F. 2015. *Measure Theory and Fine Properties of Functions*. 2nd edn. Texts in Mathematics. Boca Raton, FL: CRC Press. 17

Foroozandeh, Atefeh, Hemmat, Ataollah Askari, and Rabbani, Hossein. 2020. Use of the Shearlet Transform and Transfer Learning in Offline Handwritten Signature Verification and Recognition. *Sahand Communications in Mathematical Analysis*, **17**(3), 1–31. 6

Grafakos, Loukas. 2014. *Classical Fourier Analysis*. 3rd edn. Graduate Texts in Mathematics, vol. 249. New York: Springer. 8

Grohs, Philipp. 2010. Continuous shearlet frames and resolution of the wavefront set. *Monatshefte für Mathematik*, **164**(4), 393–426. Received 16 March 2010, Accepted 10 November 2010; Published online 4 December 2010. iii, 6, 44, 61

Gröchenig, Karlheinz. 2001. *Foundations of Time-Frequency Analysis*. Boston: Birkhäuser. 3

Guo, Kun, Labate, Demetrio, and Lim, Wang-Q. 2006. The theory of wavelets revisited by the framework of shearlets. *IMA Journal of Applied Mathematics*, **71**(4), 421–448. 5

Hintz, Peter. 2025. *An Introduction to Microlocal Analysis*. Springer. 3

Holschneider, M. 1995. *Wavelets: An Analysis Tool*. New York: The Clarendon Press, Oxford University Press. 26

Hörmander, Lars. 1983. *The Analysis of Linear Partial Differential Operators I*. Grundlehren der Mathematischen Wissenschaften, vol. 256. Berlin: Springer-Verlag. Distribution theory and Fourier analysis. 2

Kutyniok, Gitta, and Labate, Demetrio. 2009. Resolution of the wavefront set using continuous shearlets. *Transactions of the American Mathematical Society*, **361**(5), 2719–2754. 54

Kutyniok, Gitta, and Labate, Demetrio (eds). 2012. *Shearlets: Multiscale Analysis for Multivariate Data*. Boston: Birkhäuser. 2, 5, 12

Labate, Demetrio, Lim, Wang-Q, Kutyniok, Gitta, and Weiss, Guido. 2005. Sparse multidimensional representation using shearlets. *Wavelets XI*, **5914**, 59140U. 5

Mallat, S., and Hwang, W. L. 1992. Singularity Detection and Processing with Wavelets. *IEEE Transactions on Information Theory*, **38**(2), 617–643. 3

Mallat, Stéphane. 2009. *A Wavelet Tour of Signal Processing: The Sparse Way*. 3rd edn. Academic Press. 26, 29

Ramm, A. G., and Katsevich, A. I. 1996. *The Radon Transform and Local Tomography*. Research Notes in Mathematics, vol. 40. CRC Press. 16

Rudin, Walter. 1991. *Functional Analysis*. 2nd edn. International Series in Pure and Applied Mathematics. New York: McGraw-Hill, Inc. 13

Shah, Firdous A., and Tantary, Azhar Y. 2023. *Wavelet Transforms: Kith and Kin*. 1st edn. Boca Raton, FL: Chapman and Hall/CRC. 32

Simoncelli, Edward P, Freeman, William T, Adelson, Edward H, and Heeger, David J. 1992. Steerable pyramid: A flexible architecture for multi-scale derivative computation. *Proceedings of the IEEE International Conference on Image Processing*. 5

Stein, Elias M., and Shakarchi, Rami. 2005. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton Lectures in Analysis, vol. 3. Princeton, NJ: Princeton University Press. 8

Walnut, David F. 2002. *An Introduction to Wavelet Analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser. 25