

# Shearlets and microlocal analysis\*

Ph. Grohs

Research Report No. 2011-27  
May 2011

Seminar für Angewandte Mathematik  
Eidgenössische Technische Hochschule  
CH-8092 Zürich  
Switzerland

---

\*This work was supported by the European Research Council under grant ERC AdG 247277

# Shearlets and Microlocal Analysis\*

Philipp Grohs

**Abstract** Although wavelets are optimal for describing pointwise smoothness properties of univariate functions, they fail to efficiently characterize the subtle geometric phenomena of multidimensional singularities in high-dimensional functions. Mathematically these phenomena can be captured by the notion of the wavefront set which describes point- and direction-wise smoothness properties of tempered distributions. After familiarizing ourselves with the definition and basic properties of the wavefront set we show that the shearlet transform offers a simple and convenient way to characterize the wavefront set in terms of the decay properties of the shearlet coefficients.

## 1 Introduction

One of the main reasons for the popularity of the wavelet transform is its ability to characterize pointwise smoothness properties of functions. This property has proven to be extremely useful in both pure and applied mathematics. To give a random example we mention the beautiful work [15], where the pointwise smoothness of the Riemann function is studied with a precision that had not been achievable before with other methods.

For multidimensional functions, however, pointwise smoothness does not fully capture the geometric features of the singularity set: it is also of interest in which direction the function is singular. A useful notion to capture this additional information is the wavefront set which has been defined in the introduction. It has its origins in the work of Lars Hörmander on the propagation of singularities of pseudodifferential operators [16, 20].

---

Philipp Grohs  
ETH Zürich, Rämistrasse 101, 8001 Zürich, Switzerland e-mail: pgrohs@math.ethz.ch

\* This work was supported by the European Research Council under grant ERC AdG 247277

It turns out that the wavelet transform is unable to describe the Wavefront Set of a tempered distribution: even though in general the multidimensional wavelet transform *does* possess a directional parameter<sup>2</sup> [1], the fact that the degree of anisotropy of the wavelet elements does not change throughout different scales implies that microlocal phenomena occurring in frequency cones with small opening angles cannot be detected, compare also the discussions in [3].

The purpose of this chapter is to show that shearlets actually can describe directional smoothness properties of tempered distributions: it turns out that the wavefront set can be characterized as the point-direction pairs for which the shearlet coefficients are not of fast decay as the scale parameter tends to zero. Such a result is of great interest in both theory, since it provides a simple and elementary analysis tool to study refined notions of smoothness, and practice, where it is used for the detection and classification of edges in images, compare Chapter 3 in this volume.

The first proof of this result has been given in [17] for “classical”, bandlimited shearlets. In [9], an extension to general shearlet generators has been obtained.

## 1.1 Getting to know the Wavefront Set

Recall from the introduction the definition of the wavefront set of a bivariate tempered distribution:

**Definition 1.** Let  $f$  be a tempered distribution on  $\mathbb{R}^2$ . We say that  $t_0 \in \mathbb{R}^2$  is a *regular point* if there exists a neighborhood  $U_{t_0}$  of  $t_0$  such that  $\varphi\psi \in C^\infty$ , where  $\varphi$  is a smooth cutoff function with  $\varphi \equiv 1$  on  $U_{t_0}$ . The complement of the (open) set of regular points is called *singular support* of  $f$  and denoted

$$\text{sing supp}(f).$$

Furthermore, we call  $(t_0, s_0)$  a *regular directed point* if there exists a neighborhood  $U_{t_0}$  of  $t_0$ , a smooth cutoff function  $\varphi$  with  $\varphi \equiv 1$  on  $U_{t_0}$  and a neighborhood  $V_{s_0}$  of  $s_0$  such that

$$(\varphi f)^\wedge(\eta) = O((1 + |\eta|)^{-N}) \quad \text{for all } \eta = (\eta_1, \eta_2) \text{ such that } \frac{\eta_2}{\eta_1} \in V_{s_0} \text{ and } N \in \mathbb{N}. \quad (1)$$

The *wavefront set*  $\text{WF}(f)$  is the complement of the set of regular directed points.

The wavefront set is usually defined in the Fourier domain. An intuitive reason for this definition is as follows: let us assume that we are given a function with a singularity (think of a jump) in some direction. Then, if we zoom in on the singularity, all that remains are oscillations in the direction orthogonal to the singularity which corresponds to slow Fourier decay.

At first sight, this definition might not feel too natural, especially for readers with not much experience in Fourier analysis. Therefore, in order to get some feeling for

<sup>2</sup> We thank J.-P. Antoine for pointing this out.

this notion we first consider some examples for which we can immediately compute the wavefront sets.

*Example 1.* The Dirac distribution  $\delta_t$ , defined by  $\langle \delta_t, \varphi \rangle := \varphi(t)$  has singular support  $\{t\}$ . Clearly, at  $x = t$  this distribution is non regular in any direction. This is reflected by the non-decay of  $\hat{\delta}_t := \exp(2\pi i \langle \cdot, t \rangle)$ . It follows that  $\text{WF}(\delta_t) \subset \{t\} \times \mathbb{R}$ . On the other hand we have  $\text{WF}(\delta_t) \supset \{t\} \times \mathbb{R}$  since  $\delta_t$  is regular locally around any point  $t' \neq t$ . In summary, we obtain

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

*Example 2.* The line distribution  $\delta_{x_2=p+qx_1}$ , defined by  $\langle \delta_{x_2=p+qx_1}, \varphi \rangle := \int_{\mathbb{R}} \varphi(x_1, p+qx_1)$  has singular support  $\{(x_1, x_2) : x_2 = p+qx_1\}$ . To describe the Wavefront Set of  $\delta_{x_2=p+qx_1}$  we compute

$$\begin{aligned} \hat{\delta}_{x_2=p+qx_1}(\omega) &= \langle \delta_{x_2=p+qx_1}, \exp(2\pi i \langle \omega, x \rangle) \rangle \\ &= \int_{\mathbb{R}} \exp(2\pi i(\omega_1 x_1 + \omega_2(p+qx_1))) dx_2 \\ &= e^{2\pi i p \omega_2} \int_{\mathbb{R}} \exp(2\pi i(\omega_1 + q\omega_2)x_1) dx_2 \\ &= e^{2\pi i p \omega_1} \delta_{\omega_1 + q\omega_2 = 0}. \end{aligned}$$

We remark that, despite the fact that the operations above do not seem to be well-defined at first sight, it is possible to make them rigorous by noting that the equalities above are “in the sense of oscillatory integrals”, compare [20]. It follows that  $\hat{\delta}_{x_2=p+qx_1}(\omega)$  is of fast decay, except when  $\omega_2/\omega_1 = -1/q$ , and therefore

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

Before we go to the next example we pause to introduce the Radon transform [8]. As we shall see later it will serve us as a valuable tool in the proofs of the later sections.

**Definition 2.** The Radon transform of a function  $f$  is defined by

$$\mathcal{R}f(u, s) := \int_{x_2 \in \mathbb{R}} f(u - sx_2, x_2) dx_2, \quad u, s \in \mathbb{R}, \quad (2)$$

whenever this expression makes sense.

Observe that our definition of the Radon transform differs from the most common one which parameterizes the directions in terms of the angle and not the slope as we do. It turns out that our definition is particularly well-adapted to the mathematical structure of the shearlet transform. The next theorem already indicates that the Radon transform provides a useful tool in studying microlocal phenomena.

**Theorem 1 (Projection Slice Theorem).** *With  $\omega \in \mathbb{R}$  and  $\hat{\cdot}$  denoting the univariate resp. bivariate Fourier transform on the left resp. right-hand side we have the*

equality

$$(\mathcal{R}f(u, s))^\wedge(\omega) = \hat{f}(\omega(1, s)). \quad (3)$$

*Proof.*

$$\begin{aligned} (\mathcal{R}f(u, s))^\wedge(\omega) &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(u - sx_2, x_2) e^{-2\pi i u \omega} dx_2 du \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(\tilde{u}, x_2) e^{-2\pi i (\tilde{u} + sx_2) \omega} dx_2 d\tilde{u} = \hat{f}(\omega(1, s)). \end{aligned}$$

By the Projection Slice Theorem, another way of stating that  $(t_0, s_0)$  is a regular directed point is that

$$(\mathcal{R}\Phi f(u, s))^\wedge(\omega) = O(|\omega|^{-N}) \quad \text{and } s \in V_{s_0}, \text{ for all } N \in \mathbb{N}.$$

or in other words, that  $\mathcal{R}\Phi f(u, s)$  is  $C^\infty$  in  $u$  around  $s = s_0$ . We can now consider the next example, the indicator function of the unit ball.

*Example 3.* We let  $f = \chi_B$  with  $B = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\}$ . Clearly we have

$$\text{sing supp}(f) = \partial B = \{(x_1, x_2) : x_1^2 + x_2^2 = 1\}.$$

In order to describe the Wavefront Set of  $f$  we pick a bump function  $\varphi$  around a point  $t \in \partial B$  with  $t_2/t_1 = s_0$  and look at the Radon transform

$$\mathcal{R}\varphi f(u, s) = \int_{\frac{us - \sqrt{1+s^2-u^2}}{1+s^2}}^{\frac{us + \sqrt{1+s^2-u^2}}{1+s^2}} \varphi(u - sx_2, x_2) dx_2. \quad (4)$$

This expression will always be zero unless

$$u \in [t_1 + st_2 - \varepsilon, t_1 + st_2 + \varepsilon]$$

with an arbitrarily small  $\varepsilon > 0$  depending on the diameter of  $\varphi$  around  $t$ . By the definition of  $t$  we have

$$t_1 = \frac{1}{\sqrt{1+s_0^2}}, \quad t_2 = \frac{s_0}{\sqrt{1+s_0^2}},$$

and therefore  $u^2$  will be close to

$$(t_1 + st_2)^2 = \frac{(1 + ss_0)^2}{1 + s_0^2}.$$

It follows that  $u^2 - 1 - s^2$  is arbitrarily close to

$$\frac{(1 - ss_0)^2 - (1 + s^2)(1 + s_0^2)}{1 + s_0^2},$$

which is  $\neq 0$  whenever  $s \neq s_0$ . But if  $u^2 - 1 - s^2$  stays away from zero, by (4), the function  $\mathcal{R}\Phi f$  is  $C^\infty$  and therefore  $(t, s)$  is a regular directed point for  $s \neq s_0$ . The same argument implies that  $\mathcal{R}\Phi f$  is not smooth for  $s = s_0$  and we arrive at

$$\text{WF}(f) = \{(t, s) : t_1^2 + t_2^2 = 1, t_2 = st_1\}.$$

We hope that this last example convinced the reader that indeed the Radon transform is a useful tool for our purposes (compare [3, 17] where similar statements are shown using much less elementary tools such as Bessel functions and the method of stationary phase). It also gives a geometrical interpretation of the wavefront set: take a family of translated lines with a prescribed slope  $s$  and compute the integrals of  $f$  restricted to these lines. If these integrals do not vary smoothly with the translation parameter, then we have a point in the wavefront set.

## 1.2 Contributions

The main result that we would like to present is the fact that the wavefront set can be characterized by the magnitude of the shearlet coefficients as follows:

**Theorem 2.** *Let  $\psi$  be a Schwartz function with infinitely many vanishing moments in  $x_1$ -direction. Let  $f$  be a tempered distribution and  $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ , where  $\mathcal{D}_1 = \{(t_0, s_0) \in \mathbb{R}^2 \times [-1, 1] : \text{for } (s, t) \text{ in a neighborhood } U \text{ of } (s_0, t_0), |\mathcal{S}\mathcal{H}_\psi f(a, s, t)| = O(a^k) \text{ for all } k \in \mathbb{N}, \text{ with the implied constant uniform over } U\}$  and  $\mathcal{D}_2 = \{(t_0, s_0) \in \mathbb{R}^2 \times (1, \infty] : \text{for } (1/s, t) \text{ in a neighborhood } U \text{ of } (s_0, t_0), |\mathcal{S}\mathcal{H}_\psi f(a, s, t)| = O(a^k) \text{ for all } k \in \mathbb{N}, \text{ with the implied constant uniform over } U\}$ . Then*

$$\text{WF}(f)^c = \mathcal{D}.$$

The proof of this result will require some preparations. In particular we need to study continuous reconstruction formulas which allow to reconstruct an arbitrary function from its shearlet coefficients. For classical shearlet generators such a formula is given in (??) in the introduction. In Section 2 we develop analogous formulas for arbitrary shearlet generators. Then, using these representations, in Section 3 we prove our main result, Theorem 2.

## 1.3 Other Ways to Characterize the Wavefront Set

The shearlet transform is not the only decomposition that is capable of characterizing the Wavefront Set. As an example we mention the so-called FBI transform which is defined by

$$f \mapsto Tf(x, \xi, h) := \alpha_h \langle f, \exp(-2\pi i(x - \cdot)^2/2h) \exp(2\pi i(x - \cdot) \cdot \xi/h) \rangle,$$

where  $x, y \in \mathbb{R}^2$  and  $h$  is a semiclassical parameter (see [20] for more information on semiclassical analysis) and  $\alpha_h$  is some parameter. This transform can be interpreted as a semiclassical version of the Gabor transform [11] where the semiclassical Fourier transform is defined by

$$f \mapsto \hat{f}^h(\xi) := \int_{\mathbb{R}^2} f(x) \exp(2\pi i x \cdot \xi / h).$$

Heisenberg's uncertainty principle asserts that a time-frequency window must have area at least  $h$ . Therefore, by letting  $h \rightarrow 0$  the time-frequency localization gets arbitrarily good which makes the FBI transform a useful tool in microlocal analysis. An important result is that the decay rate of  $Tf(x, \xi, h)$  for  $h \rightarrow 0$  determines whether the pair  $(x, \xi_2/\xi_1)$  lies in the wavefront set of  $f$  [20].

Another transform which – being also based on parabolic scaling – is much closer to the shearlet transform is the curvelet transform [3]. The curvelet transform is also capable of characterizing the wavefront set. Another transform based on parabolic scaling with analogous properties is the transform introduced by Hart Smith in [22].

## 2 Reproduction Formulas

A crucial role in the proof of Theorem 2 will be played by so-called reproduction formulas which allow to reconstruct an arbitrary function from its shearlet coefficients. The first such formula is given in [17] for classical shearlet generators and further studies can be found in [10]. We will follow this latter work in our exposition.

*Example 4.* To give some motivation we mention the continuous wavelet transform which is defined by mapping a function  $f$  to its transform coefficients

$$\mathcal{W} \mathcal{T}_\psi f(a, b) := \langle f, \psi_{a,b} \rangle,$$

where

$$\psi_{a,b}(\cdot) := a^{-1/2} \psi\left(\frac{\cdot - b}{a}\right), \quad a, b \in \mathbb{R}.$$

It is well known that whenever the Calderon condition

$$C_\psi^{\text{wav}} := \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty$$

holds, we have the reconstruction formula

$$f = \frac{1}{C_\psi^{\text{wav}}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{W} \mathcal{T}_\psi(a, b) \overline{\psi_{a,b}} \frac{da}{a} db.$$

The measure  $\frac{da}{a} db$  comes from the fact that the wavelet transform carries the structure of a group representation of the affine group for which this measure is the left

Haar measure [14]. Another way to see why this measure is natural in the wavelet context is that the operations of dilation by  $a$  and translation by  $b$  map a unit square in  $(a, b)$ -space to a rectangle with volume  $a^{-1}$ . We also want to mention that it is not necessary to consider the wavelet transform over all frequencies  $a$ . Under some assumptions on  $\psi$  one can show that there exists a smooth function  $\Phi$  such that

$$f = \frac{1}{C_{\psi}^{\text{wav}}} \left( \int_{\mathbb{R}} \int_0^1 \mathcal{W} \mathcal{T}_{\psi}(a, b) \overline{\psi_{a,b}} \frac{da}{a} db + \int_{\mathbb{R}} \langle f, \Phi(\cdot - b) \rangle \overline{\Phi(\cdot - b)} db \right). \quad (5)$$

See [7] for more information on wavelets.

We would like to find conditions for a formula similar to (5) to hold for the shearlet transform. In the case of the full shearlet transform where we have a group structure at hand, such a formula follows from standard arguments, see e.g. [6].

*Remark 1.* The group structure provides us with the natural invariant measure for the shearlet transform: it is given by  $\frac{da}{a^3} ds dt$ . A heuristic explanation for the power of  $-3$  in the density is the fact that this measure divides the parameter space into unit cells of side  $a$  by  $\sqrt{a}$  in space (hence a factor  $a^{-3/2}$ ), unit intervals of length  $\sqrt{a}$  on the space of directions (hence a factor  $\sqrt{a}$ ) and finally a factor of  $a^{-1}$  since  $a$  is a scale parameter, see also [4].

In Example 4 we have seen an integral formula which is a  $C_{\psi}^{\text{wav}}$ -multiple of the identity. In the shearlet setting the corresponding constant arises in the following admissibility condition, compare also [5]. In the following we will assume that  $\psi$  satisfies this condition. All the results regarding the resolution of the Wavefront Set also hold without this assumption, but in that case we would have to split the frequency domain into four half-cones depending on the signs of the coordinates  $\xi_1, \xi_2$ .

**Definition 3.** A function  $\psi$  is called *admissible* if

$$C_{\psi} = \int_{\mathbb{R}} \int_{-\infty}^0 \frac{|\hat{\psi}(\xi)|^2}{|\xi_1|^2} d\xi_1 d\xi_2 = \int_{\mathbb{R}} \int_0^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi_1|^2} d\xi_1 d\xi_2 < \infty. \quad (6)$$

For our purposes it is necessary that the directional parameter varies only in a compact set, otherwise the implicit constants in Theorem 2 would deteriorate. Therefore we would like to find representations similar to (5) for the cone-adapted shearlet transform.

The main result is as follows:

**Theorem 3.** *We have the representation formula*

$$C_{\psi} f = \int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 \mathcal{S} \mathcal{H}_{\psi} f(a, s, t) \psi_{a,s,t} a^{-3} da ds dt + \int_{\mathbb{R}^2} \langle f, \Phi(\cdot - t) \rangle \Phi(\cdot - t) dt \quad (7)$$

which is valid for all  $f \in L^2(C)^{\vee}$  with a smooth function  $\Phi$  and  $C_{\psi}$  being the constant from the shearlet admissibility condition, see Definition 3. An analogous statement is true for the vertical cone  $\tilde{C}$ .



An important role in the proof of this theorem will be played by the function

$$\Delta_\psi(\xi) := \int_{-2}^2 \int_0^1 \left| \hat{\psi} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right|^2 a^{-3/2} da ds. \quad (8)$$

The reason for this fact is given in the next lemma:

**Lemma 1.** *The representation (7) holds if and only if*

$$\Delta_\psi(\xi) + |\hat{\Phi}(\xi)|^2 = C_\psi \quad \text{for all } \xi \in \mathbb{C}. \quad (9)$$

*Proof.* First we note that (7) is equivalent to

$$\begin{aligned} C_\psi^2 \|f\|_2^2 &= \int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 |\langle f, \psi_{a,s,t} \rangle|^2 a^{-3} da ds dt \\ &\quad + \int_{\mathbb{R}^2} |\langle f, \Phi(\cdot - t) \rangle|^2 dt. \end{aligned} \quad (10)$$

This follows from polarization. Taking the Fourier transform of both sides in (10) yields

$$\begin{aligned} C_\psi \|\hat{f}\|_2^2 &= \int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 |\langle \hat{f}, \hat{\psi}_{a,s,t} \rangle|^2 a^{-3} da ds dt \\ &\quad + \int_{\mathbb{R}^2} |\langle \hat{f}, (\Phi(\cdot - t))^\wedge \rangle|^2 dt \\ &= \int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 \langle \hat{f}, \hat{\psi}_{a,s,t} \rangle \overline{\langle \hat{f}, \hat{\psi}_{a,s,t} \rangle} a^{-3} da ds dt \\ &\quad + \int_{\mathbb{R}^2} \langle \hat{f}, (\Phi(\cdot - t))^\wedge \rangle \overline{\langle \hat{f}, (\Phi(\cdot - t))^\wedge \rangle} dt \end{aligned}$$

Plugging in the explicit formula for the Fourier transform lets us rewrite the above equation as follows:

$$\begin{aligned} &= \int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 \int_{\mathbb{R}^2} \hat{f}(\xi) \overline{a^{3/4} e^{-2\pi i t \xi} \hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1))} d\xi \times \\ &\quad \int_{\mathbb{R}^2} \overline{\hat{f}(\eta)} a^{3/4} e^{-2\pi i t \eta} \hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1)) d\eta a^{-3} da ds dt \\ &\quad + \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \hat{f}(\xi) \overline{\exp(-2\pi i t \xi) \hat{\Phi}(\xi)} d\xi \times \\ &\quad \int_{\mathbb{R}^2} \overline{\hat{f}(\eta)} \exp(-2\pi i t \eta) \hat{\Phi}(\eta) d\eta dt \\ &= \int_{-2}^2 \int_0^1 \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \exp(-2\pi i(\eta - \xi)t) \hat{f}(\xi) \overline{a^{3/4} \hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1))} \times \\ &\quad \overline{\hat{f}(\eta)} a^{3/4} \hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1)) d\eta d\xi dt a^{-3} da ds \\ &\quad + \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \exp(-2\pi i(\eta - \xi)t) \hat{f}(\xi) \overline{\hat{\Phi}(\xi)} \overline{\hat{f}(\eta)} \hat{\Phi}(\eta) d\xi d\eta dt. \end{aligned}$$

An application of Parseval's formula yields

$$C_\psi \|\hat{f}\|_2^2 = \|\hat{f}\|_2^2 \left( \int_{-2}^2 \int_0^1 \left| \hat{\psi} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right|^2 a^{-3/2} dad s + |\hat{\Phi}(\xi)|^2 \right).$$

This implies the statement.

Due to the previous lemma the goal in proving Theorem 3 is to show that the (more precisely: any) function  $\hat{\Phi}$  defined by the equation (9) is smooth. To this end it suffices to show that

$$|\hat{\Phi}(\xi)|^2 = O(|\xi|^{-N}) \quad \text{for } \xi \in C, \xi \rightarrow \infty.$$

Before we do this we would like to understand the function  $\Delta_\psi$  better. It turns out that if we allow to integrate over  $\mathbb{R} \times \mathbb{R}_+$  instead of  $[-2, 2] \times [0, 1]$ , the integral is equal to the admissibility constant  $C_\psi$ .

**Lemma 2.** *We have*

$$C_\psi = \int_{\mathbb{R}} \int_{\mathbb{R}_+} \left| \hat{\psi} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right|^2 a^{-3/2} dad s. \quad (11)$$

*Proof.* We make the substitution  $\eta_1(a, s) = -a\xi_1$ ,  $\eta_2(a, s) = a^{1/2}(\xi_2 - s\xi_1)$ . The Jacobian of this substitution equals  $a^{1/2}\xi_1^2 = a^{1/2}(\eta_1/a)^2 = a^{-3/2}\eta_1^2$  which shows the desired result.

Now we can prove the Fourier decay of  $\hat{\Phi}$ .

**Lemma 3.** *We have*

$$|\hat{\Phi}(\xi)|^2 = O(|\xi|^{-N}), \quad \text{for all } N \in \mathbb{N} \text{ and } |\xi_2|/|\xi_1| \leq 3/2. \quad (12)$$

*Proof.* By Lemma 2 we have that

$$\begin{aligned} |\hat{\Phi}(\xi)|^2 &= \left( \int_{a \in \mathbb{R}_+, |s| > 2} \left| \hat{\psi} \left( a\xi_1, \sqrt{a}(\xi_2 - s\xi_1) \right) \right|^2 a^{-3/2} dad s \right. \\ &\quad \left. + \int_{a > 1, |s| < 2} \left| \hat{\psi} \left( a\xi_1, \sqrt{a}(\xi_2 - s\xi_1) \right) \right|^2 a^{-3/2} dad s \right). \end{aligned}$$

We will analyze these two integrals separately, starting with the second one: Due to the smoothness of  $\psi$  and the fact that  $s$  only varies in a compact set we can estimate

$$\begin{aligned} \int_{a > 1, |s| > 2} \left| \hat{\psi} \left( a\xi_1, \sqrt{a}(\xi_2 - s\xi_1) \right) \right|^2 a^{-3/2} dad s &\lesssim \int_{a > 1} (a|\xi_1|)^{-N} a^{-3/2} \\ &\lesssim |\xi_1|^{-N} \lesssim |\xi|^{-N}. \end{aligned}$$

The last inequality follows since we can always estimate  $|\xi_1|^{-1}$  by  $|\xi|^{-1}$  due to the fact that  $|\xi_2|/|\xi_1| \leq 3/2$ . We turn to the estimation of

$$\int_{a \in \mathbb{R}_+, |s| > 2} |\hat{\psi}(a\xi_1, \sqrt{a}(\xi_2 - s\xi_1))|^2 a^{-3/2} da ds.$$

First we treat the case  $a > 1$  by estimating

$$\begin{aligned} \int_{a > 1, |s| > 2} |\hat{\psi}(a\xi_1, \sqrt{a}(\xi_2 - s\xi_1))|^2 a^{-3/2} da ds &\lesssim \int_{a > 1, |s| > 2} a^{-N} (\xi_2 - s\xi_1)^{-2N} \times \\ &\quad a^{-3/2} da ds \\ &= \int_{a > 1, |s| > 2} |\xi_1|^{-2N} a^{-N} |\xi_2/\xi_1 - s|^{-2N} \times \\ &\quad a^{-3/2} da ds \\ &\leq \int_{a > 1, |s| > 2} |\xi_1|^{-2N} a^{-N} |3/2 - |s||^{-2N} \times \\ &\quad a^{-3/2} da ds \lesssim |\xi|^{-N}. \end{aligned}$$

Now we come to the last case where we will utilize the fact that  $\psi$  possesses infinitely many moments as well as the smoothness of  $\psi$  in the second coordinate.

$$\begin{aligned} \int_{a < 1, |s| > 2} |\hat{\psi}(a\xi_1, \sqrt{a}(\xi_2 - s\xi_1))|^2 a^{-3/2} da ds &\lesssim \int_{a < 1, |s| > 2} a^M |\xi_1|^M a^{-L} |\xi_2 - s\xi_1|^{-2L} \times \\ &\quad a^{-3/2} da ds \\ &\leq \int_{a < 1, |s| > 2} a^M |\xi_1|^{M-2L} a^{-L} |3/2 - |s||^{-2L} \times \\ &\quad a^{-3/2} da ds \end{aligned}$$

for any  $L, M$ , in particular for  $L = N + 2$  and  $M = L + 4$  which gives that

$$\int_{a < 1, |s| > 2} |\hat{\psi}(a\xi_1, \sqrt{a}(\xi_2 - s\xi_1))|^2 a^{-3/2} da ds \lesssim |\xi|^{-N}.$$

Summing up all three estimates proves the lemma.

Now we have collected all the necessary ingredients to prove Theorem 3.

*Proof (of Theorem 3).* By Lemma 1 all we need to show is that any  $\Phi$ , defined by (9) is smooth. But this is established by Lemma 3.

*Remark 2.* The assumptions in 3 can be weakened considerably, see [10]. In this paper it is also shown that it not possible to obtain useful representation formulas without first projecting to a frequency cone. In [13] slightly different continuous representation formulas are considered which are called *atomic decompositions*, see also [22, 2] where similar constructions are introduced for the curvelet transform.

### 3 Resolution of the Wavefront Set

In this section we prove our main result, Theorem 2. The proof turns out to be rather long but nevertheless quite elementary. Intuitively it is not too surprising that the shearlet transform is capable of resolving the wavefront set since every shearlet element only interacts with frequency content which is contained in a cone that gets narrower as the scale increases. The difficult part is to overcome the technical details in making this intuition rigorous. To this end the Radon transform will turn out to be a valuable tool.

We divide this section into three parts. In the first part we prove one half of Theorem 2, namely the fast decay of the shearlet coefficients corresponding to a regular directed point. This turns out to be the easier part. To prove the converse statement we need to study the notion of the wavefront set a little more in the second part before we can tackle the full proof of Theorem 2 in the third part.

In the results that we present here, the choice of parabolic scaling is not essential – it could be replaced by any anisotropic scaling with corresponding dilation matrix  $\text{diag}(a, a^\delta)$ ,  $0 < \delta < 1$ .

#### 3.1 A Direct Theorem

We start by proving one half of Theorem 2, namely we show that if we are given a regular directed point of  $f$ , then only the parameter pair  $(s, t)$  corresponding to this point and direction can possibly have a large interaction with  $f$ . Such statements are usually called direct theorems (or also Jackson theorems).

*Remark 3.* The corresponding result for the wavelet case states that if a univariate function is smooth in a point then the wavelet coefficients of  $f$  associated with the location of that point decay fast with the scale, provided that the underlying wavelet has sufficiently many vanishing moments. The proofs in the wavelet case are considerably simpler, see e.g. [19].

**Theorem 4 (Direct Theorem).** *Assume that  $f \in \mathcal{S}'(\mathbb{R}^2)$  and that  $(t_0, s_0)$  is a regular directed point of  $f$ . Let  $\psi$  be a test function with infinitely many directional vanishing moments. Then there exists a neighborhood  $U_{t_0}$  of  $t_0$  and  $V_{s_0}$  of  $s_0$  such that we have the decay estimate*

$$\mathcal{S}\mathcal{H}_\psi f(a, s, t) = O(a^N) \text{ for all } N \in \mathbb{N}. \quad (13)$$

*Proof.* In the proof we will denote by  $N$  an unspecified and arbitrarily large integer. We can without loss of generality assume that  $f$  is already localized around  $t_0$ , i.e.  $f = \Phi f$  where  $\Phi$  is the cutoff function from the definition of the Wavefront Set which equals 1 around  $t_0$ . To show this we prove that

$$\langle (1 - \Phi)f, \psi_{ast} \rangle = O(a^N). \quad (14)$$

Since we have assumed that  $\psi$  is in the Schwartz class, we have for any  $P > 0$  that

$$|\psi(x)| \lesssim (1 + |x|)^{-P} \quad (15)$$

By definition we have

$$\psi_{ast}(x_1, x_2) = a^{-3/4} \psi \left( \frac{(x_1 - t_1) + s(x_2 - t_2)}{a}, \frac{x_2 - t_2}{a^{1/2}} \right). \quad (16)$$

Now we note that in computing the inner product (14) we can assume that  $|x - t| > \delta$  for some  $\delta > 0$  and  $t$  in a small neighborhood  $U_{t_0}$  of  $t_0$ , since  $(1 - \Phi)f = 0$  around  $t_0$ . By (15) we estimate

$$\begin{aligned} |\psi_{ast}(x)| &\lesssim a^{-3/4} \left( 1 + \left| \begin{pmatrix} a^{-1} & sa^{-1} \\ 0 & a^{-1/2} \end{pmatrix} (x - t) \right| \right)^{-P} \\ &\leq a^{-3/4} \left( 1 + \left\| \begin{pmatrix} a & -sa^{1/2} \\ 0 & a^{1/2} \end{pmatrix} \right\|^{-1} |x - t| \right)^{-P} \\ &\lesssim a^{-3/4} \left( 1 + C(s)a^{-1/2}|x - t| \right)^{-P} = O \left( a^{-3/4+P/2}|x - t|^{-P} \right) \end{aligned}$$

for  $|x - t| > \delta$  and  $C(s) = \left( 1 + \frac{s^2}{2} + (s^2 + \frac{s^2}{4})^{1/2} \right)^{1/2}$  (compare [17, Lemma 5.2]).

Let us for now assume that  $f$  is a slowly growing function (i.e. a function with at most polynomial growth). Then we can estimate

$$\begin{aligned} \langle (1 - \Phi)f, \psi_{ast} \rangle &\lesssim a^{-3/4+P/2} \int_{|x-t| \geq \delta} |x - t|^{-P} |1 - \Phi(x_1, x_2)| |f(x_1, x_2)| dx_1 dx_2 \\ &= O(a^N), \end{aligned} \quad (17)$$

for  $t \in U_{t_0}$  and  $P$  large enough, which yields (14). For a general tempered distribution  $f$  we use the fact that  $f$  can be written as a finite superposition of terms of the form  $D^\beta g$ , where  $g$  has slow growth,  $D$  denotes the total differential and  $\beta \in \mathbb{N}^2$  [21]. Then we can use integration by parts together with the fact that also the derivatives of  $\psi$  obey the decay property (15) to arrive at the general case.

Now, assuming that  $f = \varphi f$  is localized, we go on to estimate the shearlet coefficients  $|\langle f, \psi_{ast} \rangle|$ . To do this we utilize the Fourier transform. Furthermore, we assume that  $f \in L_2(\mathbb{R}^2)$ . The general case can again be handled by repeated partial integrations, at the expense of some (finitely many) powers of  $a$ . First note that the Fourier transform of  $\psi_{ast}$  is given by

$$\hat{\psi}_{ast}(\xi) = a^{3/4} \exp(-2\pi i t \xi) \hat{\psi} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right). \quad (18)$$

Now pick  $\frac{1}{2} < \alpha < 1$  and write

$$\begin{aligned}
|\langle f, \psi_{ast} \rangle| &= |\langle \hat{f}, \hat{\psi}_{ast} \rangle| \leq a^{3/4} \int_{\mathbb{R}^2} |\hat{f}(\xi_1, \xi_2)| \left| \hat{\psi} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right| d\xi \\
&= \underbrace{a^{3/4} \int_{|\xi_1| < a^{-\alpha}}}_{A} + \underbrace{a^{3/4} \int_{|\xi_1| > a^{-\alpha}}}_{B}.
\end{aligned} \tag{19}$$

Since  $\psi$  possesses  $M$  moments in the  $x_1$  direction which means that

$$\hat{\psi}(\xi_1, \xi_2) = \xi_1^M \hat{\theta}(\xi_1, \xi_2)$$

with some  $\theta \in L^2(\mathbb{R}^2)$ , we can estimate  $A$  as

$$\begin{aligned}
A &= a^{3/4} \int_{|\xi_1| < a^{-\alpha}} |\hat{f}(\xi_1, \xi_2)| \left| \hat{\psi} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right| d\xi \\
&= a^{3/4} \int_{|\xi_1| < a^{-\alpha}} a^M |\xi_1|^M |\hat{f}(\xi_1, \xi_2)| \left| \hat{\theta} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right| d\xi \\
&\leq a^{M(1-\alpha)} a^{3/4} \int_{|\xi_1| < a^{-\alpha}} |\hat{f}(\xi_1, \xi_2)| \left| \hat{\theta} \left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1) \right) \right| d\xi \\
&\leq a^{(1-\alpha)M} \langle |\hat{f}|, |\hat{\theta}_{ast}| \rangle \leq a^{(1-\alpha)M} \|\hat{f}\|_2 \|\hat{\theta}_{ast}\|_2 = a^{(1-\alpha)M} \|f\|_2 \|\theta\|_2 \\
&= O(a^N)
\end{aligned} \tag{20}$$

with  $M$  large enough. In order to estimate  $B$  we make the following substitution:

$$\begin{pmatrix} a & 0 \\ -a^{1/2}s & a^{1/2} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \end{pmatrix}, \quad d\xi_1 d\xi_2 = a^{-3/2} d\tilde{\xi}_1 d\tilde{\xi}_2.$$

Then

$$B = a^{-3/4} \int_{\frac{|\tilde{\xi}_1|}{a} > a^{-\alpha}} \left| \hat{f} \left( \frac{\tilde{\xi}_1}{a}, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \right| \left| \hat{\psi} \left( \tilde{\xi}_1, \tilde{\xi}_2 \right) \right| d\xi. \tag{21}$$

Now we shall use that  $(t_0, s_0)$  is a regular directed point of  $f$ . This means that there is a neighborhood  $(s_0 - \varepsilon, s_0 + \varepsilon)$  such that

$$\hat{f}(\eta_1, \eta_2) \lesssim (1 + |\eta|)^{-R} \quad \text{for all } \frac{\eta_2}{\eta_1} \in (s_0 - \varepsilon, s_0 + \varepsilon). \tag{22}$$

Looking at (21) we now consider  $\frac{\eta_2}{\eta_1}$  with

$$\eta_1 := \frac{\tilde{\xi}_1}{a}, \quad \eta_2 := \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \quad \text{and} \quad \frac{\tilde{\xi}_1}{a} > a^{-\alpha}$$

and get the estimate

$$s - a^{\alpha-1/2} \frac{\tilde{\xi}_2}{\tilde{\xi}_1} \leq \frac{\eta_2}{\eta_1} = s + a^{-1/2} \tilde{\xi}_2 \frac{a}{\tilde{\xi}_1} \leq s + a^{\alpha-1/2} \tilde{\xi}_2. \tag{23}$$

By (22) we have that

$$\left| \hat{f} \left( \frac{\tilde{\xi}_1}{a}, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \right| \lesssim \left( 1 + \frac{|\tilde{\xi}_1|}{a} \right)^{-R} \quad (24)$$

for  $s$  in a neighborhood  $V_{s_0}$  of  $s_0$ ,  $\frac{|\tilde{\xi}_1|}{a} > a^{-\alpha}$  and  $|\tilde{\xi}_2| < \varepsilon' a^{1/2-\alpha}$  for some  $\varepsilon' < \varepsilon$ . Now we first split the integral  $B$  according to

$$\begin{aligned} B &= a^{-3/4} \int_{|\tilde{\xi}_1|/a \geq a^{-\alpha}} \left| \hat{f} \left( \tilde{\xi}_1/a, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \right| \left| \hat{\psi}(\tilde{\xi}_1, \tilde{\xi}_2) \right| d\tilde{\xi}_1 d\tilde{\xi}_2 \\ &= \underbrace{a^{-3/4} \int_{|\tilde{\xi}_1|/a \geq a^{-\alpha}}}_{B_1} \underbrace{a^{-3/4} \int_{|\tilde{\xi}_1|/a \geq a^{-\alpha} \quad |\tilde{\xi}_2| < \varepsilon' a^{1/2-\alpha}}}_{B_2} \left| \hat{f} \left( \tilde{\xi}_1/a, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \right| \left| \hat{\psi}(\tilde{\xi}_1, \tilde{\xi}_2) \right| d\tilde{\xi}_1 d\tilde{\xi}_2 \end{aligned} \quad (25)$$

By (24) we can estimate  $B_1$  according to

$$B_1 = O \left( a^{\alpha R - 3/4} \|\hat{\psi}\|_1 \right) = O(a^N) \quad (26)$$

whenever  $R$  is large enough.

It only remains to estimate  $B_2$ . For this we will use the fact that  $\frac{\partial^L}{\partial x_2^L} \psi \in L^2(\mathbb{R}^2)$ . This implies that

$$\begin{aligned} B_2 &\leq a^{-3/4} \int_{|\tilde{\xi}_1|/a \geq a^{-\alpha} \quad |\tilde{\xi}_2| > \varepsilon' a^{1/2-\alpha}} \left| \hat{f} \left( \tilde{\xi}_1/a, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \hat{\psi}(\tilde{\xi}_1, \tilde{\xi}_2) \right| d\tilde{\xi}_1 d\tilde{\xi}_2 \\ &= a^{-3/4} \int \left| \hat{f} \left( \tilde{\xi}_1/a, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \tilde{\xi}_2^{-L} \left( \frac{\partial^L}{\partial x_2^L} \psi \right)^\wedge(\tilde{\xi}_1, \tilde{\xi}_2) \right| d\tilde{\xi}_1 d\tilde{\xi}_2 \\ &\leq (\varepsilon')^{-L} a^{-3/4 + (\alpha - 1/2)L} \times \end{aligned} \quad (27)$$

$$\begin{aligned} &\int_{\mathbb{R}^2} \left| \hat{f} \left( \tilde{\xi}_1/a, \frac{s}{a} \tilde{\xi}_1 + a^{-1/2} \tilde{\xi}_2 \right) \left( \frac{\partial^L}{\partial x_2^L} \psi \right)^\wedge(\tilde{\xi}_1, \tilde{\xi}_2) \right| d\tilde{\xi}_1 d\tilde{\xi}_2 \\ &= (\varepsilon')^{-L} a^{(\alpha - 1/2)L} \left| \left\langle |\hat{f}|, \left( \frac{\partial^L}{\partial x_2^L} \psi_{ast} \right)^\wedge \right\rangle \right| \\ &\leq (\varepsilon')^{-L} a^{(\alpha - 1/2)L} \|f\|_2 \left\| \frac{\partial^L}{\partial x_2^L} \psi \right\|_2 = O(a^N). \end{aligned} \quad (28)$$

Putting together the estimates (17), (20), (26) and (28) we finally arrive at the desired conclusion.

*Remark 4.* Observe that in the proof of the direct theorem it is nowhere essential that we have parabolic scaling of  $a$  in the first and  $a^{1/2}$  in the second coordinate. All the results that we present in this chapter hold equally well for any anisotropic scaling of  $a$  in the first coordinate and  $a^\delta$  in the second coordinate where  $0 < \delta < 1$  is arbitrary, see also the discussion at the end of [17]. This stands in contrast to the results on Fourier integral operators [13] and sparse approximation of cartoon images [12, 18], where the parabolic scaling plays an essential role.

### 3.2 Properties of the Wavefront Set

Here we prove to basic results related to the wavefront set. The first result concerns its well-definedness. Recall that the definition of a regular directed point involves a localization by a bump function. The first thing we need to show is that the property of being a regular directed point does not depend on the choice of such a function. The second result concerns the frequency side and states that a point-direction pair comprises a regular directed point of  $f$  if and only if it is a regular directed point of the frequency projection of  $f$  onto a cone containing the direction of the point-direction pair. Both of these results seem obvious but they need to be proven, nevertheless.

We start with the first statement.

**Lemma 4.** *Assume that  $(t_0, s_0)$  is a regular directed point of  $f$  and  $\varphi$  is a test function. Then  $(t_0, s_0)$  is a regular directed point of  $\varphi f$ .*

*Proof.* Assume that  $(t_0, s_0)$  is a regular directed point of  $f$  and let  $\xi$  be such that  $\xi_2/\xi_1 = s_0$ . Then we can write  $\xi = t e_0$  where  $e_0$  denotes the unit vector with slope  $s_0$  and  $t$  proportional to  $|\xi|$ . What we want to show is that

$$\hat{\varphi} f(t e_0) = O(|t|^{-N}).$$

Since pointwise multiplication transforms into convolution in the Fourier domain, this is equivalent to

$$\hat{\varphi} * \hat{f}(t e_0) = \int_{\mathbb{R}^2} \hat{f}(t e_0 - \xi) \hat{\varphi}(\xi) d\xi = O(|t|^{-N}). \quad (29)$$

Since  $(t_0, s_0)$  is a regular directed point, by definition there exists  $0 < \delta < 1$  such that  $t e_0 + B_\delta$  is still contained in the frequency cone with slopes  $s \in V_{s_0}$  for all  $t \in \mathbb{R}$ . Here,  $B_\delta$  denotes the unit ball in  $\mathbb{R}^2$  with radius  $\delta$  around the origin. After picking  $\delta$  we can split the integral in (29) into

$$\int_{|\xi| < \delta t} \hat{f}(t e_0 - \xi) \hat{\varphi}(\xi) d\xi + \int_{|\xi| > \delta t} \hat{f}(t e_0 - \xi) \hat{\varphi}(\xi) d\xi.$$

We start by estimating the first term: By assumption we then have that

$$\hat{f}(t e_0 - \xi) = O(|t e_0 - \xi|^{-N}) = O(|t|^{-N}),$$

and this suffices to establish that

$$\int_{|\xi| < \delta t} \hat{f}(t e_0 - \xi) \hat{\varphi}(\xi) d\xi = O(|t|^{-N}).$$

Now the second term. As before in the proof of Theorem 4 we assume that  $\hat{f}$  is a slowly growing function. Again, this is no restriction since any tempered distribution is a finite sum of derivatives of slowly growing functions. To get rid of the derivatives



we simply do some integrations by parts in the integral (29) and shift them to  $\hat{\phi}$ . Since  $\hat{\phi}$  is still a test function this does not do any harm. Now we can establish the second part as by estimating

$$\begin{aligned} \int_{|\xi| > \delta t} \hat{f}(te_0 - \xi) \hat{\phi}(\xi) d\xi &\lesssim \int_{|\xi| > \delta t} |te_0 - \xi|^L |\xi|^{-M} d\xi \\ &\lesssim \int_{|\xi| > \delta t} |t|^L |\xi|^L |\xi|^{-M} d\xi \end{aligned}$$

with  $M$  arbitrary and  $L$  the (finite) order of growth of  $\hat{f}$ . Picking  $M$  sufficiently large and using the fact that  $|\xi| \gtrsim |t|$  we arrive at the desired estimate.

The second basic result that we want to establish is that a frequency projection onto a cone does not affect the set of regular directed points.

**Lemma 5.** *Assume that  $(t_0, s_0)$  is a regular directed point of  $f$ . Let  $C_0$  be a cone containing the slope  $s_0$ . Then  $(t_0, s_0)$  is a regular directed point of  $\hat{P}_{C_0} f$ , where  $\hat{P}_{C_0}$  denotes the frequency projection of  $f$  onto the frequency cone  $C_0$ . The converse also holds true.*

*Proof.* To show this we first assume that  $(t_0, s_0)$  is a regular directed point of  $f$ . By definition we then can pick a bump function  $\varphi$  such that  $\varphi f$  has fast Fourier decay in a frequency cone around  $s_0$ , i.e.

$$\hat{\varphi} * \hat{f}(\xi) = O(|\xi|^{-N}), \quad \xi_2/\xi_1 \in V_{s_0}.$$

By shrinking the neighborhood  $V_{s_0}$  of  $s_0$  if necessary, we can assume without loss of generality that for some small  $\delta > 0$  we have the inclusion (see Figure 3.3 right)

$$\{\eta + B_\delta |\eta| : \eta_2/\eta_1 \in V_{s_0}\} \subset C_0. \quad (30)$$

The inclusion (30) implies that

$$\xi \in C_0^c \Rightarrow |\eta - \xi| > \delta |\eta| \quad \eta_2/\eta_1 \in V_{s_0}. \quad (31)$$

Write

$$\hat{\varphi} * \hat{f}(\eta) = \int_{\mathbb{R}^2} \chi_{C_0} \hat{f}(\eta - \xi) \hat{\varphi}(\xi) d\xi + \int_{\mathbb{R}^2} \chi_{C_0^c} \hat{f}(\eta - \xi) \hat{\varphi}(\xi) d\xi.$$

The statement is proven if we can show that

$$\int_{\mathbb{R}^2} \chi_{C_0^c} \hat{f}(\eta - \xi) \hat{\varphi}(\xi) d\xi = O(|\eta|^{-N}), \quad \eta_2/\eta_1 \in V_{s_0}. \quad (32)$$

But this follows by writing (32) as

$$\int_{\mathbb{R}^2} \chi_{C_0^c} \hat{f}(\xi) \hat{\varphi}(\eta - \xi) d\xi$$

and using (31) together with the fact that

$$\hat{\phi}(\eta - \xi) = O(|\eta - \xi|^{-N}).$$

The last result in particular implies that in order to study the Wavefront Set of a tempered distribution  $f$  we can restrict ourselves to studying the Wavefront Sets of the two frequency projections  $\hat{P}_C f$ ,  $\hat{P}_{C^c} f$  separately. This also holds true for the shearlet coefficients of a tempered distribution:

**Lemma 6.** *Assume that  $f$  is a tempered distribution. Let  $(t_0, s_0)$  be a point-direction pair and  $C_0$  a frequency cone around the direction with slope  $s_0$ . Then we have the equivalence*

$$\mathcal{S}\mathcal{H}_\psi f(a, s_0, t_0) = O(a^N) \Leftrightarrow \mathcal{S}\mathcal{H}_\psi (\hat{P}_{C_0} f)(a, s_0, t_0) = (a^N).$$

*Proof.* By linearity of the shearlet transform we have

$$\mathcal{S}\mathcal{H}_\psi f(a, s_0, t_0) - \mathcal{S}\mathcal{H}_\psi (\hat{P}_{C_0} f)(a, s_0, t_0) = \mathcal{S}\mathcal{H}_\psi (\hat{P}_{C_0^c} f)(a, s_0, t_0).$$

But clearly  $(t_0, s_0)$  is a regular directed point of  $\hat{P}_{C_0^c} f$ . Therefore, by Theorem 4 we can establish that

$$\mathcal{S}\mathcal{H}_\psi (\hat{P}_{C_0^c} f)(a, s_0, t_0) = O(a^N)$$

which proves the statement.

### 3.3 Proof of the Main Result

We are almost ready to tackle the second half of Theorem 2. First we need the following localization lemma.

**Lemma 7.** *Consider a tempered distribution  $f$  and a smooth bump function  $\phi$  which is supported in a small neighborhood  $V_{t_0}$  of  $t_0 \in \mathbb{R}^2$ . Let  $U_{t_0}$  be another neighborhood of  $t_0$  with  $V_{t_0} \subset\subset U_{t_0}$ . Consider the function*

$$g(x) = \int_{t \in U_{t_0}^c, s \in [-2, 2], a \in [0, 1]} \langle f, \Psi_{a, s, t} \rangle \phi(x) \Psi_{a, s, t}(x) a^{-3} da ds dt.$$

Then

$$\hat{g}(\xi) = O(|\xi|^{-N}), \quad \xi \in C. \quad (33)$$

*Proof.* Consider for  $s \in [-1, 1]$  the Radon transform

$$I(u) := \mathcal{R}g(u, s).$$

By the projection slice theorem we need to show that

$$I^{(N)}(u) := \left( \frac{d}{du} \right)^N I \in L_1(\mathbb{R}) \quad (34)$$

which implies that

$$\omega^N \hat{I}(\omega) = \omega^N \hat{g}(\omega, s\omega) \lesssim 1,$$

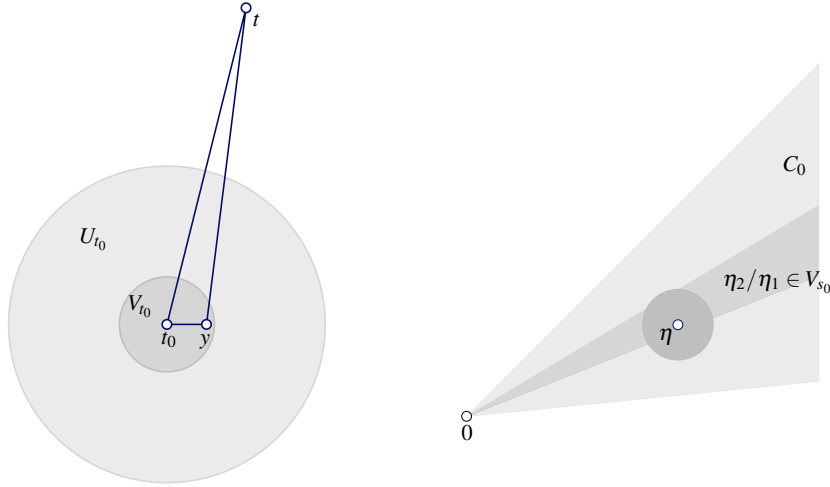
and therefore since  $|s| \leq 1$  this implies (33). By the product rule  $I^{(N)}$  can be written as a sum of terms of the form

$$\int_{t \in U(t_0)^c, s \in [-2, 2], a \in [0, 1]} \langle f, \psi_{a,s,t} \rangle \int_{\mathbb{R}} \left( \frac{d}{dx_1} \right)^{N-j} \varphi(u - sx, s) a^{-j} \left( \left( \frac{d}{dx_1} \right)^{N-j} \psi \right)_{a,s,t} (u - sx, x) dx a^{-3} da ds dt.$$

By the support properties of  $\varphi$ , the points  $y := (u - sx, x)$  must lie in  $V(t_0)$  for this expression to be nonzero. With the same argument as in the beginning of the proof of Theorem 4, leading to (14), we can establish that

$$\left( \left( \frac{d}{dx_1} \right)^{N-j} \psi \right)_{a,s,t} (y) = O(a^N |y - t|^{-N}). \quad (35)$$

Since we can assume that  $y \in V_{t_0}$  and  $t \in U_{t_0}^c$  we obtain (see Figure 3.3 left)



**Fig. 1** Left: Illustration of the proof of (36). Right: Illustration of (30).

$$|y - t| \gtrsim |t - t_0|, \quad (36)$$

which, together with (35) establishes the desired claim. Note that by Fubini's theorem the application of the Radon transform is justified a-posteriori.

**Theorem 5.** *Assume that  $f$  is a tempered distribution and that for  $(s_0, t_0) \in [-1, 1] \times \mathbb{R}^2$  we have in a neighborhood  $U$  of  $(s_0, t_0)$ ,  $|\mathcal{S}\mathcal{H}_\psi f(a, s, t)| = O(a^N)$  for all  $N \in \mathbb{N}$ , with the implied constant uniform over  $U$ . Then  $(s_0, t_0)$  is a regular directed point of  $f$ . An analogous result holds for  $\frac{1}{s_0} \in [-1, 1]$  and the shearlet  $\tilde{\psi}$  for the vertical cone  $\tilde{C}$ .*

*Proof.* First we assume without loss of generality that  $f$  has its Fourier transform supported in  $C$ . Otherwise we continue with the frequency projection  $\hat{P}_C f$  and invoke Lemma 5 and Lemma 6 to arrive at the theorem.

By Theorem 3 we can represent  $f$  as

$$f = \int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 \mathcal{S}\mathcal{H}_\psi f(a, s, t) \psi_{a,s,t} a^{-3} da ds dt + \int_{\mathbb{R}^2} \langle f, \Phi(\cdot - t) \rangle \Phi(\cdot - t) dt,$$

modulo an irrelevant constant. A further simplification can be obtained by noting that  $\int_{\mathbb{R}^2} \langle f, \Phi(\cdot - t) \rangle \Phi(\cdot - t) dt$  is always smooth, since

$$\left( \int_{\mathbb{R}^2} \langle f, \Phi(\cdot - t) \rangle \Phi(\cdot - t) dt \right)^\wedge(\xi) = \hat{f}(\xi) |\hat{\Phi}(\xi)|^2 = O(|\xi|^{-N})$$

by Lemma 3 (this holds if  $\hat{f}$  is a slowly growing function, the general case is handled by integration-by-parts as usual). Therefore, all we need to show is that  $(t_0, s_0)$  is a regular directed point of

$$\int_{\mathbb{R}^2} \int_{-2}^2 \int_0^1 \mathcal{S}\mathcal{H}_\psi f(a, s, t) \psi_{a,s,t} a^{-3} da ds dt.$$

To this end, we multiply this expression by a smooth bump function  $\varphi$  localized around  $t_0$  and note that by Lemma 7 we actually only need to show that  $(t_0, s_0)$  is a regular directed point of

$$h := \int_{U_{t_0}} \int_{-2}^2 \int_0^1 \mathcal{S}\mathcal{H}_\psi f(a, s, t) \varphi \psi_{a,s,t} a^{-3} da ds dt,$$

where  $U_{t_0}$  is a compact neighborhood of  $t_0$ . To show this we will establish that

$$I^{(N)}(u) \in L_1(\mathbb{R}),$$

where

$$I(u) := \mathcal{B}h(u, s_0).$$

With the same computations as in the proof of Lemma 7, we see that  $I^{(N)}$  consists of terms of the form

$$\int_{t \in U_{t_0}, s \in [-2, 2], a \in [0, 1]} \langle f, \psi_{a,s,t} \rangle \int_{\mathbb{R}} \left( \frac{d}{dx_1} \right)^{N-j} \varphi(u - s_0 x, s) a^{-j} \left( \left( \frac{d}{dx_1} \right)^{N-j} \psi \right)_{a,s,t} (u - s_0 x, x) dx a^{-3} da ds dt.$$

By making  $U_{t_0}$  (and the support of  $\varphi$ ) sufficiently small, we can establish the existence of  $\varepsilon > 0$  such that for all  $t \in U_{t_0}$  and  $s \in [s_0 - \varepsilon, s_0 + \varepsilon]$  we have  $(s, t) \in U$ . We now split the above integral according to  $s \in [s_0 - \varepsilon, s_0 + \varepsilon]$  and  $|s - s_0| > \varepsilon$ . For the first part we invoke the fast decay of the shearlet coefficients of  $f$  for  $(s, t) \in U$  to see that

$$\int_{t \in U_{t_0}, s \in [s_0 - \varepsilon, s_0 + \varepsilon], a \in [0, 1]} \langle f, \psi_{a,s,t} \rangle \int_{\mathbb{R}} \left( \frac{d}{dx_1} \right)^{N-j} \varphi(u - s_0 x, s) a^{-j} \left( \left( \frac{d}{dx_1} \right)^{N-j} \psi \right)_{a,s,t} (u - s_0 x, x) dx a^{-3} da ds dt = O(1). \quad (37)$$

In order to handle the case  $|s - s_0| > \varepsilon$  we note that the corresponding integral can be written as

$$\int_{t \in U_{t_0}, s \in [s_0 - \varepsilon, s_0 + \varepsilon]^c, a \in [0, 1]} \langle f, \psi_{a,s,t} \rangle a^{-j} \times \mathcal{R} \left( \left( \frac{d}{dx_1} \right)^{N-j} \varphi \left( \left( \frac{d}{dx_1} \right)^{N-j} \psi \right)_{a,s,t} \right) (u, s_0) a^{-3} da ds dt. \quad (38)$$

Note that we can write

$$\mathcal{R} \left( \left( \frac{d}{dx_1} \right)^{N-j} \varphi \left( \left( \frac{d}{dx_1} \right)^{N-j} \psi \right)_{a,s,t} \right) (u, s_0) = \langle \tilde{\delta}_{u,s_0}, \theta_{a,s,t} \rangle, \quad (39)$$

where

$$\theta := \left( \frac{d}{dx_1} \right)^{N-j} \psi$$

and

$$\tilde{\delta}_{u,s_0} := \left( \frac{d}{dx_1} \right)^{N-j} \varphi \delta_{x_1 = u - s_0 x_2}.$$

The Wavefront Set of  $\tilde{\delta}_{u,s_0}$  is given by

$$\{(x_1, x_2, s) : x_1 = u - s_0 x_2, s = s_0\},$$

as can be seen from the computations in Example 2. Since the function  $\theta$  satisfies the assumptions of Theorem 4 we can apply this result and obtain that

$$\langle \tilde{\delta}_{u,s_0}, \theta_{a,s,t} \rangle = \mathcal{S} \mathcal{H}_\theta \tilde{\delta}_{u,s_0}(a, s, t) = O(a^N).$$

By (39) this implies that also the expression (39) is bounded. Together with (37) this proves that  $I^{(N)}$  is bounded and therefore in  $L_1$  since it is compactly supported. This proves the theorem. The argument for the dual cone follows from obvious modifications.

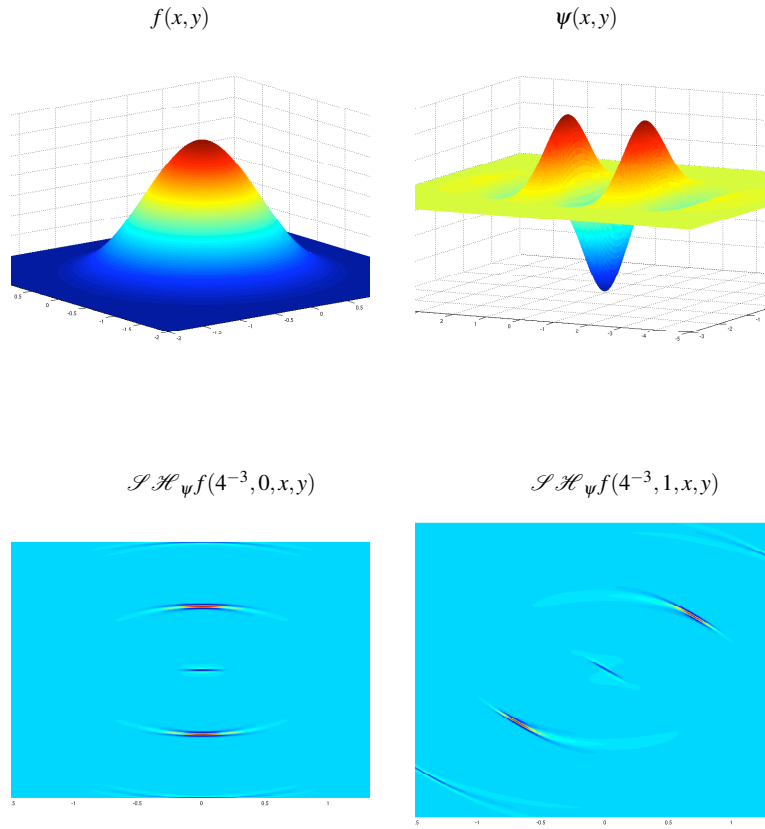
Putting together Theorem 4 and Theorem 5 we have finally proved Theorem 2:

**Corollary 1.** *Theorem 2 holds true.*

*Remark 5.* It is possible to weaken the assumptions in Theorem 2 considerably if one is only interested in determining directional regularity of a finite order, as opposed to our definition where Fourier decay of arbitrary order is asked in the definition of a regular directed point. In that case only finitely many vanishing moments and only finite smoothness of  $\psi$  is required. The details are given in [9].

## References

1. J.-P. Antoine, R. Murenzi, P. Vandergheynst, and S. T. Ali. *Two-Dimensional Wavelets and their Relatives*. Cambridge University Press, 2004.
2. E. Candes and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Communications in Pure and Applied Mathematics*, 58:1472–1528, 2004.
3. E. Candes and D. Donoho. Continuous curvelet transform: I. resolution of the wavefront set. *Applied and Computational Harmonic Analysis*, 19(2):162–197, 2005.
4. E. Candes and D. Donoho. Continuous curvelet transform: II. discretization and frames. *Applied and Computational Harmonic Analysis*, 19(2):198–222, 2005.
5. S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke. The uncertainty principle associated with the continuous shearlet transform. *International Journal of Wavelets Multiresolution and Information Processing*, 6(2):157, 2008.
6. S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke. Shearlet coorbit spaces and associated banach frames. *Applied and Computational Harmonic Analysis*, 27(2):195 – 214, 2009.
7. I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.
8. S. R. Deans. *The Radon Transform and Some of Its Applications*. John Wiley and Sons, 1983.
9. P. Grohs. Continuous shearlet frames and resolution of the wavefront set. *Monatshefte für Mathematik*, 2011. in press.
10. P. Grohs. Continuous shearlet tight frames. *Journal of Fourier Analysis and Applications*, 2011. in press.
11. K. Grchenig. *Foundations of Time-Frequency Analysis*. Birkhuser, 2001.
12. K. Guo and D. Labate. Optimally sparse multidimensional representation using shearlets. *SIAM journal on mathematical analysis*, 39(1):298–318, 2008.
13. K. Guo and D. Labate. Representation of fourier integral operators using shearlets. *Journal of Fourier Analysis and Applications*, 14(3):327–371, 2008.
14. E. Hewitt and K. Ross. *Abstract Harmonic Analysis I*. Springer, 1979.
15. M. Holschneider and P. Tchamitchian. Pointwise analysis of Riemann’s “nondifferentiable” function. *Inventiones Mathematicae*, 105:157–175, 1991.
16. L. Hörmander. *The Analysis of linear Partial Differential Operators*. Springer, 1983.
17. G. Kutyniok and D. Labate. Resolution of the wavefront set using continuous shearlets. *Transactions of the American Mathematical Society*, 361:2719–2754, 2009.



**Fig. 2** This figure illustrates the main result, Theorem 2. On the top left we show the function  $f$  to be analyzed – a rotated quadratic B-spline curve which has curvature discontinuities at integer points. Therefore, the Wavefront Set of  $f$  consists of the origin with all directions attached and the tangents of concentric circles with integer radius. The top right shows the analyzing shearlet which is of tensor product type. The two figures at the bottom show the magnitudes of the shearlet coefficients corresponding to two different directions – the horizontal direction on the bottom left and the diagonal direction with slope 1 on the bottom right. It is evident from these pictures that only the parameters corresponding to points of the Wavefront Set have non-negligible shearlet coefficients.

18. G. Kutyniok and W.-Q. Lim. Compactly supported shearlets are optimally sparse. Technical report, 2010.
19. S. Mallat. *A wavelet tour of signal processing*. Academic Press, 2003.
20. A. Martinez. *An Introduction to Semiclassical and Microlocal Analysis*. Springer, 2002.
21. W. Rudin. *Functional Analysis*. McGraw-Hill, 1991.
22. H. Smith. A hardy space for fourier integral operators. *Journal of Geometric Analysis*, 8:629–653, 1998.



# Research Reports

No.	Authors/Title
11-27	<i>Ph. Grohs</i> Shearlets and microlocal analysis
11-26	<i>H. Kumar</i> Implicit-explicit Runge-Kutta methods for the two-fluid MHD equations
11-25	<i>H. Papasaïka, E. Kokiopoulou, E. Baltasvias, K. Schindler and D. Kressner</i> Sparsity-seeking fusion of digital elevation models
11-24	<i>H. Harbrecht and J. Li</i> A fast deterministic method for stochastic elliptic interface problems based on low-rank approximation
11-23	<i>P. Corti and S. Mishra</i> Stable finite difference schemes for the magnetic induction equation with Hall effect
11-22	<i>H. Kumar and S. Mishra</i> Entropy stable numerical schemes for two-fluid MHD equations
11-21	<i>H. Heumann, R. Hiptmair, K. Li and J. Xu</i> Semi-Lagrangian methods for advection of differential forms
11-20	<i>A. Moiola</i> Plane wave approximation in linear elasticity
11-19	<i>C.J. Gittelsohn</i> Uniformly convergent adaptive methods for parametric operator equations
11-18	<i>E. Kokiopoulou, D. Kressner, M. Zervos and N. Paragios</i> Optimal similarity registration of volumetric images
11-17	<i>D. Marazzina, O. Reichmann and Ch. Schwab</i> $hp$ -DGFEM for Kolmogorov-Fokker-Planck equations of multivariate Lévy processes
11-16	<i>Ch. Schwab and A.M. Stuart</i> Sparse deterministic approximation of Bayesian inverse problems
11-15	<i>A. Barth and A. Lang</i> Almost sure convergence of a Galerkin–Milstein approximation for stochastic partial differential equations
11-14	<i>X. Claeys</i> A single trace integral formulation of the second kind for acoustic scattering