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.

Key words: Microlocal analysis, Radon transform, Representation formulas, Wavefront set

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 [16], where the pointwise smoothness of the Riemann function is studied with a precision that had not been achievable 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

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

G. Kutyniok and D. Labate (eds.), *Shearlets: Multiscale Analysis for Multivariate Data*, Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-8316-0_2,

is the wavefront set which has been defined a literature reference to the introduction chapter of this book. It has its origins in the work of Lars Hörmander on the propagation of singularities of pseudodifferential operators [17, 21].

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[1] [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 of edges in images, compare chapter "Analysis and Identification of Multidimensional Singularities using the Continuous Shearlet Transform" in this volume. For real practical purposes such as analysis and classification of edges, still more refined results are needed [13, 14].

The first proof of this result has been given in [18] for "classical," bandlimited shearlets. In [9], an extension to general shearlet generators has been obtained.

1.1 Notation

Whenever possible we use the notation from chapter "Introduction to Shearlets" of this volume. We write $\mathscr{S}(\mathbb{R}^k)$ for the space of Schwartz functions [22] and $\mathscr{S}(\mathbb{R}^k)'$ for its dual w.r.t. to the canonical pairing $\langle \cdot, \cdot \rangle_{L^2(\mathbb{R}^k)}$, the space of tempered distributions. For $k = 2$, we simply write $\mathscr{S}$, $\mathscr{S}'$. We also denote

$$\mathscr{C} := \left\{ \xi \in \mathbb{R}^2 : |\xi_2|/|\xi_1| \leq 3/2 \right\}$$

and

$$\mathscr{C}' := \left\{ \xi \in \mathbb{R}^2 : |\xi_1|/|\xi_2| \leq 3/2 \right\},$$

the horizontal, resp. vertical frequency cone. For $A \subset \mathbb{R}^2$, we write χ_A for its indicator function, i.e.,

$$\chi_A(\xi) = \begin{cases} 1 \text{ if } \xi \in A \\ 0 \text{ if } \xi \notin A \end{cases}$$

The symbol $\mathbb{T}$ shall denote the one-dimensional torus.

[1] We thank J.-P. Antoine for pointing this out.

1.2 Getting to Know the Wavefront Set

Recall from Sect. 2.2 of chapter "Introduction to Shearlets" in this volume 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 $\varphi f \in C^\infty$, where φ 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 φ with $\varphi \equiv 1$ on U_{t_0}, and a neighborhood V_{s_0} of s_0 such that

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

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

Remark 1. The definition of the wavefront set as given in (1) excludes the case $s_0 = \infty$, or $\eta_1 = 0$. In order to avoid this problem, we can make the same definition with the coordinate directions reversed in this case. Alternatively, we can let the parameter s vary in the projective line $\mathbb{P}^1$. We will not explicitly state this in all places below but would like to remark that we can usually restrict our attention to $s \in [-1, 1]$, the other directions being handled by reversing the coordinate directions. $\diamondsuit$

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

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

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

$\diamondsuit$

Example 2. The line distribution $\delta_{x_2=p+qx_1}$ defined by

$$\langle \delta_{x_2=p+qx_1}, \varphi \rangle := \int_{\mathbb{R}} \varphi(x_1, p+qx_1) \mathrm{d}x_1$$

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

$$\begin{aligned}
\hat{\delta}_{x_2=p+qx_1}(\xi) &= \left\langle \delta_{x_2=p+qx_1}, \exp\left(2\pi \mathrm{i}\langle \xi, x\rangle\right)\right\rangle \\
&= \int_{\mathbb{R}} \exp\left(2\pi \mathrm{i}(\xi_1 x_1 + \xi_2(p+qx_1))\right) \mathrm{d}x_1 \\
&= \mathrm{e}^{2\pi \mathrm{i} p \xi_2} \int_{\mathbb{R}} \exp\left(2\pi \mathrm{i}(\xi_1 + q\xi_2)x_1\right) \mathrm{d}x_1 \\
&= \mathrm{e}^{2\pi \mathrm{i} p \xi_1} \delta_{\xi_1+q\xi_2=0}.
\end{aligned}$$

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 [21]. Essentially, this means that the equality holds in a weak sense, which is appropriate since we are dealing with tempered distributions. It follows that $\hat{\delta}_{x_2=p+qx_1}(\xi)$ is of fast decay, except when $\xi_2/\xi_1 = -1/q$, and therefore

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

$\diamondsuit$

Before we go to the next example we pause to introduce the Radon transform [7]. As we shall see below, it will serve us as a valuable tool in the proofs of the latter sections.

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

$$\mathscr{R}f(u,s) := \int_{\mathbb{R}} f(u - sx_2, x_2)\mathrm{d}x_2, \quad u,s \in \mathbb{R}, \tag{2}$$

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 $\mathscr{F}_1$ denoting the univariate Fourier transform with respect to the first coordinate, we have the equality*

$$\mathscr{F}_1(\mathscr{R}f(u,s))(\omega) = \hat{f}(\omega(1,s)) \quad \forall s \in \mathbb{R}. \tag{3}$$

Proof.

$$\begin{aligned}\mathscr{F}_1(\mathscr{R}f(u,s))(\omega) &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(u-sx_2,x_2)\mathrm{e}^{-2\pi\mathrm{i}u\omega}\mathrm{d}x_2\mathrm{d}u\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(\tilde{u},x_2)\mathrm{e}^{-2\pi\mathrm{i}(\tilde{u}+sx_2)\omega}\mathrm{d}x_2\mathrm{d}\tilde{u} = \hat{f}(\omega(1,s)).\end{aligned}$$

□

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

$$\mathscr{F}_1(\mathscr{R}\varphi f(u,s))(\omega) = O(|\omega|^{-N}) \quad \text{and } s\in V_{s_0}, \text{ for all } N\in\mathbb{N}.$$

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

Example 3. We let $f=\chi_{\mathrm{B}}$ with $B=\{(x_1,x_2):\ x_1^2+x_2^2\le 1\}$. 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 φ around a point $t\in\partial B$ with $t_2/t_1 = s_0$ and look at the Radon transform

$$\mathscr{R}\varphi f(u,s) = \int_{\frac{us-\sqrt{1+s^2-u^2}}{1+s^2}}^{\frac{us+\sqrt{1+s^2-u^2}}{1+s^2}} \varphi(u-sx_2,x_2)\mathrm{d}x_2. \tag{4}$$

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 φ around t. To see this, let us assume that the function φ is supported in the set

$$(t_1+[-\varepsilon,\varepsilon])\times(t_2+[-\varepsilon,\varepsilon]). \tag{5}$$

Therefore, for the expression above to be nonzero, we need to require that

$$x_2\in t_2+[-\varepsilon,\varepsilon]$$

which, by (5) implies that

$$u\in t_1+st_2+(1+s)[-\varepsilon,\varepsilon].$$

By the definition of t we have

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

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 $\mathscr{R}\varphi f$ is C^∞ and therefore (t,s) is a regular directed point for $s \neq s_0$. The same argument implies that $\mathscr{R}\varphi f$ is not smooth for $s = s_0$ and we arrive at

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

$\diamondsuit$

We hope that this last example convinced the reader that indeed the Radon transform is a useful tool for our purposes (compare [3, 18] 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.

1.2.1 Shearlets and the wavefront set

Now we will take a first look at the relation between the decay rate of the shearlet coefficients and directional regularity. Let us consider the classical bandlimited shearlet ψ described in the introduction at hand of the previous examples. First, we briefly recall the classical shearlet construction, see also Definition 1 from chapter "Introduction to Shearlets" of this volume. We pick functions $\psi_1, \, \psi_2 \in \mathscr{S}$ such that supp $\hat{\psi}_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$, supp $\hat{\psi}_2 \subset [-1,1]$ and define

$$\hat{\psi}(\xi) = \hat{\psi}_1(\xi_1)\hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right).$$

We would like to remark that the specific choices for the supports of the functions $\psi_1, \, \psi_2$ have no deeper meaning; the important thing is that $\hat{\psi}_1$ is supported away from zero (or in other words that ψ_1 is a wavelet) and $\hat{\psi}_2$ is supported around zero. In fact, all that is needed is a sufficient number of directional vanishing moments which imposes the crucial frequency localization property needed for the detection of anisotropic structures.

Definition 3. A function ψ possesses N directional vanishing moments in x_1-direction if

$$\frac{\hat{f}(\xi)}{\xi_1^N} \in L^2(\mathbb{R}^2).$$

Directional vanishing moments in other directions can be defined in a similar way.

For simplicity, in the present discussion we stick to the classical bandlimited construction and treat the general case in the following sections. We then have

$$\begin{aligned}\hat{\psi}_{a,s,t}(\xi) &= a^{3/4}\exp\left(-2\pi\mathrm{i}\langle t,\xi\rangle\right)\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\\ &= a^{3/4}\exp\left(-2\pi\mathrm{i}\langle t,\xi\rangle\right)\hat{\psi}_1(a\xi_1)\hat{\psi}_2\left(a^{-1/2}\left(\frac{\xi_2}{\xi_1} - s\right)\right).\end{aligned} \tag{6}$$

It follows that

$$\operatorname{supp}\,\hat{\psi}_{a,s,t} \subset \left\{\xi \in \mathbb{R}^2 :\ \xi_1 \in \left[-\frac{1}{2a}, -\frac{1}{16a}\right] \cup \left[\frac{1}{16a}, \frac{1}{2a}\right],\ \left|\frac{\xi_2}{\xi_1} - s\right| \leq \sqrt{a}\right\}. \tag{7}$$

The following two examples are inspired by the discussions in [18, Sect. 4].

Example 4. Let us consider again the distribution δ_0 and examine its shearlet coefficients. First consider the case $t = 0$ and look at

$$\langle \delta_0, \psi_{a,s,0}\rangle = \psi_{a,s,0}(0) = a^{-3/4}\psi(0).$$

On the other hand, for $t \neq 0$ we use the fact that $\psi \in \mathscr{S}$, which implies that

$$|\psi_{a,s,t}(0)| \lesssim a^{-3/4}\left(1 + \|A_a^{-1}S_s^{-1}t\|^2\right)^{-k}, \quad \forall k \in \mathbb{N}.$$

An elementary calculation shows that, e.g., for $s \in [-1,1]$ and $t \neq 0$ this expression is of order $O(a^k)$ for all $k \in \mathbb{N}$. For other directions with slope greater than 1, we use the shearlet construction $\tilde{\Psi}(\tilde{\psi})$ for the vertical cone and apply the same analysis, see the introduction. From Example 1, it follows that the wavefront set of δ_0 is characterized precisely by the point–direction pairs for which the shearlet transform does not decay faster than any power of a:

Proposition 1. *If $t = 0$, we have*

$$\mathscr{SH}_\psi(\delta_0)(a,s,t) \sim a^{-3/4} \quad \text{as } a \to 0.$$

In all other cases $\mathscr{SH}_\psi(\delta_0)(a,s,t)$ decays rapidly as $a \to 0$.

$\diamondsuit$

Example 5. Here we study the behavior of the shearlet coefficients of the line singularity distribution $\nu = \delta_{x_2 = qx_1}$. As seen in Example 2 we have

$$\hat{\nu} = \hat{\delta}_{x_2 = qx_1} = \delta_{\xi_1 + q\xi_2 = 0},$$

and thus

$$\langle \psi_{a,s,t}, \nu \rangle = \langle \hat{\psi}_{a,s,t}, \hat{\nu} \rangle = \langle \hat{\psi}_{a,s,t}, \delta_{\xi_1+q\xi_2=0} \rangle = \int_{\mathbb{R}} \hat{\psi}_{a,s,t}(-q\xi_2, \xi_2) \mathrm{d}\xi_2.$$

Inserting (6) gives

$$\begin{aligned}
\langle \psi_{a,s,t}, \nu \rangle &= a^{3/4} \int_{\mathbb{R}} \exp\left(-2\pi\mathrm{i}(-qt_1\xi_2 + t_2\xi_2)\right) \hat{\psi}_1(-aq\xi_2) \hat{\psi}_2 \\
&\quad \times \left(a^{-1/2}\left(-\frac{1}{q} - s\right)\right) \mathrm{d}\xi_2 \\
&= -\frac{a^{-1/4}}{q} \int_{\mathbb{R}} \exp\left(-2\pi\mathrm{i}a^{-1}\left(t_1\xi_2 - \frac{1}{q}t_2\xi_2\right)\right) \hat{\psi}_1(\xi_2) \hat{\psi}_2 \\
&\quad \times \left(a^{-1/2}\left(-\frac{1}{q} - s\right)\right) \mathrm{d}\xi_2 \\
&= -\frac{a^{-1/4}}{q} \psi_1\left(a^{-1}\left(t_1 - \frac{1}{q}t_2\right)\right) \hat{\psi}_2\left(a^{-1/2}\left(-\frac{1}{q} - s\right)\right) \qquad (8)
\end{aligned}$$

Since $\psi_1 \in \mathscr{S}(\mathbb{R})$, we have

$$|\psi_1(x)| \leq C_k \left(1 + x^2\right)^k, \quad \forall k \in \mathbb{N}$$

and thus

$$\psi_1\left(a^{-1}\left(t_1 - \frac{1}{q}t_2\right)\right) = O(a^k), \quad \forall k \in \mathbb{N},$$

whenever $t_1 \neq \frac{1}{q}t_2$. Let us now assume that $t_1 = \frac{1}{q}t_2$. We distinguish two cases, namely $s \neq -\frac{1}{q}$ and $s = -\frac{1}{q}$. First assume that $s \neq -\frac{1}{q}$. Then, for a sufficiently small, the expression

$$\hat{\psi}_2\left(a^{-1/2}\left(-\frac{1}{q} - s\right)\right)$$

will always be zero due to the support properties of $\hat{\psi}_2$. By (8) we again have the estimate

$$\langle \nu, \psi_{a,s,t} \rangle = O(a^k), \quad \forall k \in \mathbb{N}.$$

We are left with the final case $t_1 = \frac{1}{q}t_2$ and $s = -\frac{1}{q}$ where we have

$$\langle \nu, \psi_{a,s,t} \rangle = -\frac{a^{-1/4}}{q} \psi_1(0) \hat{\psi}_2(0) \sim a^{-1/4}.$$

Summarizing, we get

Proposition 2. *If $t_1 = \frac{1}{q}t_2$ and $s = -\frac{1}{q}$, we have*

$$\mathscr{SH}_\psi(\nu)(a,s,t) \sim a^{-1/4}.$$

In all other cases $\mathcal{SH}_\psi(\nu)(a,s,t)$ decays rapidly as $a \to 0$. In other words, $WF(\nu)$ is given precisely by those indices (s,t) for which $\mathcal{SH}_\psi(\nu)$ does not decay rapidly with a.

$\diamondsuit$

Example 6. With a little more work, one can show the following result related to Example 3, compare [18, Sect. 4.4]. We let $f = \chi_B$ with B the two-dimensional unit ball. The following result (which we give here without proof) holds:

Proposition 3. *If $(t,s) \in WF(f)$, then we have*

$$\mathcal{SH}_\psi(f)(a,s,t) \sim a^{3/4}.$$

In all other cases $\mathcal{SH}_\psi(f)(a,s,t)$ decays rapidly as $a \to 0$.

$\diamondsuit$

The previous examples have shown that at least for very simple singularity distributions the wavefront set can be precisely described by the asymptotic behavior of the shearlet transform coefficients. It is the purpose of the remaining chapter to formulate and prove this fact in a more general setting. But first we take a look at a simpler transform:

1.2.2 Wavelets and the wavefront set

Also two-dimensional wavelets possess a directional parameter, and one might wonder whether the anisotropic scaling underlying the shearlet transform is really necessary. We would like to illustrate at hand of a simple example that it actually is. The discussion is inspired by [1] and [3, Sect. 6]. Similar to the one-dimensional setting, we can construct a two-dimensional wavelet transform by starting with a function $\psi \in \mathcal{S}$ satisfying the admissibility condition

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

Now we define functions $\psi_{a,\theta,t}$, $(a,\theta,t) \in \mathbb{R}_+ \times \mathbb{T} \times \mathbb{R}^2$ by

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

R_θ denoting rotation by $\theta \in \mathbb{T}$. Define the two-dimensional wavelet transform of a tempered distribution $f \in \mathcal{S}'$ by

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

Using this definition it is possible to get, up to a constant, a representation formula which represents f as a superposition of the $\psi_{a,\theta,t}$'s with the corresponding coefficients given by $\mathscr{W}_\psi^{2D}(f)(a,\theta,t)$. We have

$$\int_{\mathbb{R}^2}\int_{\mathbb{T}}\int_{\mathbb{R}_+}\left|\mathscr{W}_\psi^{2D}(f)(a,\theta,t)\right|^2\frac{\mathrm{d}a}{a^3}\mathrm{d}\theta\mathrm{d}t = C_\psi^{2D}\int_{\mathbb{R}^2}|f(x)|^2\mathrm{d}x,$$

with some constant C_ψ^{2D}, see [1, Proposition 2.2.1]. One possibility for the construction of ψ would be to start with a usual wavelet $\tilde{\psi}$ and define $\psi(x_1,x_2) := \tilde{\psi}(10x_1,x_2/10)$. This gives an anisotropic basis function ψ and the parameter θ in the two-dimensional wavelet transform gives a notion of directionality. A natural question to ask is whether this much simpler transform is able to characterize the wavefront set of a tempered distribution.

Example 7. Let us start with a simple point singularity δ_0. Then we have

$$\mathscr{W}_\psi^{2D}(a,\theta,t) = a^{-1}\psi\left(\frac{R_\theta(-t)}{a}\right) = O(a^k) \quad \forall k\in\mathbb{N},$$

whenever $t\neq 0$. If $t=0$ we have

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

It follows that the two-dimensional wavelet transform is able to describe the wavefront set of δ_0. $\diamondsuit$

Of course, in terms of describing anisotropic notions of regularity, the point distribution δ_0 is irrelevant. Let us see what happens if we analyze the simple line singularity ν from Example 5.

Example 8. We consider the line distribution $\nu = \delta_{x_1=0}$ and its two-dimensional wavelet transform at the point $t=0$ lying on the singularity line. We have

$$\begin{aligned}\langle\nu,\psi_{a,\theta,0}\rangle &= \langle\hat{\nu},\hat{\psi}_{a,\theta,0}\rangle = a\int_{\mathbb{R}}\hat{\psi}\left(a\cos(\theta)\xi_1,-a\sin(\theta)\xi_1\right)\mathrm{d}\xi_1\\ &= \int_{\mathbb{R}}\hat{\psi}\left(\cos(\theta)\xi_1,-\sin(\theta)\xi_1\right)\mathrm{d}\xi_1 := A(\theta).\end{aligned}$$

The function $A(\theta)$ varies smoothly with θ and in particular there is no way to sharply distinguish between the singularity direction corresponding to $\theta=0$ and other directions nearby. In particular for θ near 0, the decay rate of $\mathscr{W}_\psi^{2D}(\nu)(a,\theta,0)$ is only $O(1)$. $\diamondsuit$

The previous example implies that for the description of truly anisotropic phenomena, the two-dimensional wavelet transform is not suitable:

Anisotropic scaling is necessary to describe anisotropic regularity!

Note that wavelets can characterize the singular support of a tempered distribution, see, e.g., [2] and the references therein.

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

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

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 by Theorem 3 from chapter "Introduction to Shearlets" in this volume. In Sect. 2, we develop analogous formulas for arbitrary shearlet generators. Then, using these representations, in Sect. 3 we prove our main result, Theorem 2. In addition, Fig. 2 provides an illustration of the result.

1.4 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

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

where $x, y \in \mathbb{R}^2$ and h is a semiclassical parameter (see [21] for more information on semiclassical analysis) and α_h is some parameter. This transform can be interpreted as a semiclassical version of the Gabor transform [8] where the semiclassical Fourier transform is defined by

$$f \mapsto \hat{f}^h(\xi) := \int_{\mathbb{R}^2} f(x) \exp\left(2\pi \mathrm{i}\langle x, \xi\rangle/h\right).$$

Heisenberg's uncertainty principle asserts that a time–frequency window must have area at least h. Therefore, by letting $h \to 0$ 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 $h \to 0$ determines whether the pair $(x, \xi_2/\xi_1)$ lies in the wavefront set of f [21].

Another transform which—being 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 [23].

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 [18] for classical shearlet generators and further studies can be found in [10]. We will follow this latter work in our exposition.

Example 9. To give some motivation, we mention the continuous wavelet transform which is defined by mapping a function f to its transform coefficients

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

where

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

It is well known that whenever the Calderòn condition

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

holds, we have the reconstruction formula

$$f = \frac{1}{C_\psi^{\mathrm{wav}}}\int_{\mathbb{R}}\int_{\mathbb{R}} \mathscr{W}_\psi(a,b)\overline{\psi_{a,b}}\frac{\mathrm{d}a}{a}\mathrm{d}b.$$

The measure $\frac{\mathrm{d}a}{a}\mathrm{d}b$ 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 [15]. 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 scales a. Under some assumptions on ψ one can show that there exists a smooth function Φ such that

$$f = \frac{1}{C_\psi^{\mathrm{wav}}}\left(\int_{\mathbb{R}}\int_0^1 \mathscr{W}_\psi(a,b)\overline{\psi_{a,b}}\frac{\mathrm{d}a}{a}\mathrm{d}b + \int_{\mathbb{R}}\langle f, \Phi(\cdot - b)\rangle\overline{\Phi(\cdot - b)}\mathrm{d}b\right). \qquad (9)$$

See [6] for more information on wavelets. $\diamondsuit$

We would like to find conditions for a formula similar to (9) 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., chapter

"Multivariate Shearlet Transform, Shearlet Coorbit Spaces and their Structural Properties" in this volume.

Remark 2. The group structure provides us with the natural (left-) invariant measure for the shearlet transform: it is given by $\frac{\mathrm{d}a}{a^3}\mathrm{d}s\mathrm{d}t$. 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]. $\diamondsuit$

In Example 9, we have seen an integral formula which is a C_ψ^{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 ψ 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 ξ_1, ξ_2.

Definition 4. A function ψ is called *admissible* if

$$C_\psi = \int_{\mathbb{R}}\int_{-\infty}^{0} \frac{|\hat{\psi}(\xi)|^2}{|\xi_1|^2}\mathrm{d}\xi_1\mathrm{d}\xi_2 = \int_{\mathbb{R}}\int_{0}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi_1|^2}\mathrm{d}\xi_1\mathrm{d}\xi_2 < \infty. \tag{10}$$

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 (9) for the cone-adapted shearlet transform.

From now on, we will assume that the shearlet ψ possesses infinitely many directional vanishing moments, compare Definition 3. The main result is as follows:

Theorem 3. *We have the representation formula*

$$C_\psi f = \int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} \mathscr{S}\mathscr{H}_\psi f(a,s,t)\psi_{a,s,t}a^{-3}\mathrm{d}a\mathrm{d}s\mathrm{d}t + \int_{\mathbb{R}^2}\langle f, \Phi(\cdot - t)\rangle \Phi(\cdot - t)\mathrm{d}t \tag{11}$$

which is valid for all $f \in L^2(\mathscr{C})^\vee$ with a smooth function Φ and C_ψ being the constant from the shearlet admissibility condition, see Definition 4. An analogous statement is true for the vertical cone $\mathscr{C}'$.

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

$$\Delta_\psi(\xi) := \int_{-2}^{2}\int_{0}^{1} \left|\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right|^2 a^{-3/2}\mathrm{d}a\mathrm{d}s. \tag{12}$$

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

Lemma 1. *The representation (11) holds if and only if*

$$\Delta_\psi(\xi) + \left|\hat{\Phi}(\xi)\right|^2 = C_\psi \quad \textit{for all } \xi \in \mathscr{C}. \tag{13}$$

Proof. First we note that (11) is equivalent to

$$C_\psi^2\|f\|_2^2 = \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 |\langle f, \psi_{a,s,t}\rangle|^2 a^{-3}\mathrm{d}a\mathrm{d}s\mathrm{d}t + \int_{\mathbb{R}^2} |\langle f, \Phi(\cdot - t)\rangle|^2 \mathrm{d}t. \tag{14}$$

This follows from polarization. Taking the Fourier transform of both sides in (14) yields

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

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

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

An application of Parseval's formula yields

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

This implies the statement. □

Due to the previous lemma, the goal in proving Theorem 3 is to show that the (more precisely: any) function Φ defined by (13) is smooth. To this end, it suffices to show that

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

Before we do this, we would like to understand the function Δ_ψ 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_ψ.

Lemma 2. *We have*

$$C_\psi = \int_{\mathbb{R}}\int_{\mathbb{R}_+} \left|\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right|^2 a^{-3/2}\mathrm{d}a\mathrm{d}s. \tag{15}$$

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 $a^{1/2}\xi_1^2 = a^{1/2}\left(\eta_1/a\right)^2 = a^{-3/2}\eta_1^2$ which shows the desired result. □

Now, we can prove the Fourier decay of Φ.

Lemma 3. *We have*

$$\left|\hat{\Phi}(\xi)\right|^2 = O\left(|\xi|^{-N}\right), \quad \textit{for all } N\in\mathbb{N} \textit{ and } |\xi_2|/|\xi_1| \le 3/2. \tag{16}$$

Proof. By Lemma 2, we have that

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

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

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

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

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

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

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

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

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

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

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

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 Φ defined by (13) is smooth. But this is established by Lemma 3. □

Remark 3. The assumptions in Theorem 3 can be weakened considerably, see [10]. In that paper, it is also shown that it not possible to obtain useful representation formulas without first projecting to a frequency cone. In [12] slightly different continuous representation formulas are considered which are called *atomic decompositions*, see also [2, 23] 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 diag $\left(a,a^{\delta}\right)$, $0<\delta<1$.

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 4. 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., [20]. $\diamondsuit$

Theorem 4 (Direct Theorem). *Assume that $f\in\mathscr{S}'$ and that (t_0,s_0) is a regular directed point of f. Let $\psi\in\mathscr{S}$ 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 such that we have the decay estimate*

$$\mathscr{S}\mathscr{H}_{\psi}f(a,s,t)=O\left(a^N\right)\ \textit{for all}\ N\in\mathbb{N}. \tag{17}$$

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., $f=\varphi f$ where φ is the cutoff function from the definition of the wavefront set which equals 1 around t_0. More specifically, we prove that

$$\langle(1-\varphi)f,\psi_{a,s,t}\rangle=O\left(a^N\right). \tag{18}$$

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

$$|\psi(x)| \lesssim (1+|x|)^{-P} \tag{19}$$

By definition we have

$$\psi_{a,s,t}(x_1,x_2) = a^{-3/4}\psi\left(\frac{(x_1-t_1)+s(x_2-t_2)}{a}, \frac{x_2-t_2}{a^{1/2}}\right). \tag{20}$$

Now we note that in computing the inner product (18) 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 (19), we estimate

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

for $|x-t| > \delta$ and $C(s) = \left(1+\frac{s^2}{2}+(s^2+\frac{s^2}{4}\right)^{1/2})^{1/2}$ (compare [18, 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

$$\begin{aligned}
\langle(1-\varphi)f, \psi_{a,s,t}\rangle &\lesssim a^{-3/4+P/2}\int_{|x-t|\geq\delta} |x-t|^{-P}|1-\varphi(x_1,x_2)||f(x_1,x_2)|\mathrm{d}x_1\mathrm{d}x_2 \\
&= O\left(a^N\right),
\end{aligned} \tag{21}$$

for $t \in U_{t_0}$ and P large enough, which yields (18). 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 $\beta \in \mathbb{N}^2$ [22]. Then we can use integration by parts together with the fact that also the derivatives of ψ obey the decay property (19) to arrive at the general case.

Now, assuming that $f = \varphi f$ is localized, we go on to estimate the shearlet coefficients $|\langle f, \psi_{a,s,t}\rangle|$. 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 $\psi_{a,s,t}$ is given by

$$\hat{\psi}_{a,s,t}(\xi) = a^{3/4}\exp(-2\pi \mathrm{i}\langle t,\xi\rangle)\,\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right). \tag{22}$$

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

$$
\begin{aligned}
|\langle f, \psi_{a,s,t}\rangle| = \left|\langle \hat{f}, \hat{\psi}_{a,s,t}\rangle\right| &\le a^{3/4} \int_{\mathbb{R}^2} \left|\hat{f}(\xi_1,\xi_2)\right| \left|\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right| \mathrm{d}\xi \\
&= \underbrace{a^{3/4}\int_{|\xi_1|<a^{-\alpha}}}_{A} + \underbrace{a^{3/4}\int_{|\xi_1|>a^{-\alpha}}}_{B}. \qquad (23)
\end{aligned}
$$

Since ψ 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

$$
\begin{aligned}
A &= a^{3/4} \int_{|\xi_1|<a^{-\alpha}} \left|\hat{f}(\xi_1,\xi_2)\right| \left|\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right| \mathrm{d}\xi \\
&= a^{3/4} \int_{|\xi_1|<a^{-\alpha}} a^M |\xi_1|^M \left|\hat{f}(\xi_1,\xi_2)\right| \left|\hat{\theta}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right| \mathrm{d}\xi \\
&\le a^{M(1-\alpha)} a^{3/4} \int_{|\xi_1|<a^{-\alpha}} \left|\hat{f}(\xi_1,\xi_2)\right| \left|\hat{\theta}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right| \mathrm{d}\xi \\
&\le a^{(1-\alpha)M} \left\langle \left|\hat{f}\right|, \left|\hat{\theta}_{a,s,t}\right|\right\rangle \le a^{(1-\alpha)M} \left\|\hat{f}\right\|_2 \left\|\hat{\theta}_{a,s,t}\right\|_2 = a^{(1-\alpha)M} \|f\|_2 \|\theta\|_2 \\
&= O\left(a^N\right) \qquad (24)
\end{aligned}
$$

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

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

Then

$$
B = a^{-3/4} \int_{\frac{|\tilde{\xi}_1|}{a} > a^{-\alpha}} \left|\hat{f}\left(\frac{\tilde{\xi}_1}{a}, \frac{s}{a}\tilde{\xi}_1 + a^{-1/2}\tilde{\xi}_2\right)\right| \left|\hat{\psi}\left(\tilde{\xi}_1, \tilde{\xi}_2\right)\right| \mathrm{d}\tilde{\xi}. \qquad (25)
$$

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

$$
\hat{f}(\eta_1,\eta_2) \lesssim (1 + |\eta|)^{-R} \quad \text{for all } \frac{\eta_2}{\eta_1} \in (s_0 - \varepsilon, s_0 + \varepsilon). \qquad (26)
$$

Looking at (25), we now consider $\frac{\eta_2}{\eta_1}$ with

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

and get the estimate

$$s - a^{\alpha-1/2}\tilde{\xi}_2 \leq \frac{\eta_2}{\eta_1} = s + a^{-1/2}\tilde{\xi}_2 \frac{a}{\tilde{\xi}_1} \leq s + a^{\alpha-1/2}\tilde{\xi}_2. \tag{27}$$

By (26), we have that

$$\left| \hat{f}\left(\frac{\tilde{\xi}_1}{a}, \frac{s}{a}\tilde{\xi}_1 + a^{-1/2}\tilde{\xi}_2\right)\right| \lesssim \left(1 + \frac{|\tilde{\xi}_1|}{a}\right)^{-R} \tag{28}$$

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

$$\begin{aligned} B &= a^{-3/4}\int_{|\tilde{\xi}_1|/a \geq a^{-\alpha}} \left|\hat{f}\left(\tilde{\xi}_1/a, \frac{s}{a}\tilde{\xi}_1 + a^{-1/2}\tilde{\xi}_2\right)\right| \left|\hat{\psi}\left(\tilde{\xi}_1, \tilde{\xi}_2\right)\right| \mathrm{d}\tilde{\xi}_1 \mathrm{d}\tilde{\xi}_2 \\ &= \underbrace{a^{-3/4}\int_{|\tilde{\xi}_1|/a \geq a^{-\alpha},\ |\tilde{\xi}_2| < \varepsilon' a^{1/2-\alpha}}}_{B_1} + \underbrace{a^{-3/4}\int_{|\tilde{\xi}_1|/a \geq a^{-\alpha}\ |\tilde{\xi}_2| > \varepsilon' a^{1/2-\alpha}}}_{B_2} \end{aligned} \tag{29}$$

By (28), we can estimate B_1 according to

$$B_1 = O\left(a^{\alpha R - 3/4}\|\hat{\psi}\|_1\right) = O\left(a^N\right) \tag{30}$$

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

$$\begin{aligned} B_2 &\leq a^{-3/4}\int_{|\tilde{\xi}_1|/a \geq a^{-\alpha}\ |\tilde{\xi}_2| > \varepsilon' a^{1/2-\alpha}} \left|\hat{f}\left(\tilde{\xi}_1/a, \frac{s}{a}\tilde{\xi}_1 + a^{-1/2}\tilde{\xi}_2)\hat{\psi}(\tilde{\xi}_1, \tilde{\xi}_2\right)\right| \mathrm{d}\tilde{\xi}_1 \mathrm{d}\tilde{\xi}_2 \\ &= a^{-3/4}\int \left|\hat{f}\left(\tilde{\xi}_1/a, \frac{s}{a}\tilde{\xi}_1 + a^{-1/2}\tilde{\xi}_2\right)\tilde{\xi}_2^{-L}\left(\frac{\partial^L}{\partial x_2^L}\psi\right)^{\wedge}(\tilde{\xi}_1, \tilde{\xi}_2)\right| \mathrm{d}\tilde{\xi}_1 \mathrm{d}\tilde{\xi}_2 \\ &\leq (\varepsilon')^{-L} a^{-3/4+(\alpha-1/2)L} \\ &\quad \times \int_{\mathbb{R}^2} \left|\hat{f}\left(\tilde{\xi}_1/a, \frac{s}{a}\tilde{\xi}_1 + a^{-1/2}\tilde{\xi}_2\right)\right| \left|\left(\frac{\partial^L}{\partial x_2^L}\psi\right)^{\wedge}\left(\tilde{\xi}_1, \tilde{\xi}_2\right)\right| \mathrm{d}\tilde{\xi}_1 \mathrm{d}\tilde{\xi}_2 \\ &= (\varepsilon')^{-L} a^{(\alpha-1/2)L}\left|\left\langle |\hat{f}|, \left|\left(\frac{\partial^L}{\partial x_2^L}\psi_{a,s,t}\right)^{\wedge}\right|\right\rangle\right| \end{aligned} \tag{31}$$

$$\leq (\varepsilon')^{-L} a^{(\alpha-1/2)L}\|f\|_2 \left\|\frac{\partial^L}{\partial x_2^L}\psi\right\|_2 = O\left(a^N\right). \tag{32}$$

Putting together the estimates (21), (24), (30), and (32) we finally arrive at the desired conclusion. □

Remark 5. 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^{δ} in the second coordinate where $0 < \delta < 1$ is arbitrary, see also the discussion at the end of [18]. This stands in contrast to the results on Fourier integral operators [12] and sparse approximation of cartoon images [11, 19], where the parabolic scaling plays an essential role. $\diamondsuit$

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. These results might 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 φ is a test function. Then, (t_0, s_0) is a regular directed point of φf.*

Proof. Assume that (t_0, s_0) is a regular directed point of f and let ξ be such that $\xi_2/\xi_1 = s_0$ (if $\xi_1 = 0$ we need to reverse the coordinate directions, compare Remark 1). Then we can write $\xi = te_0$ where e_0 denotes the unit vector with slope s_0 and t proportional to $|\xi|$. What we want to show is that

$$\widehat{\varphi f}(te_0) = O\left(|t|^{-N}\right).$$

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

$$\hat{\varphi} * \hat{f}(te_0) = \int_{\mathbb{R}^2} \hat{f}(te_0 - \xi)\hat{\varphi}(\xi)\mathrm{d}\xi = O\left(|t|^{-N}\right). \tag{33}$$

Since (t_0, s_0) is a regular directed point, by definition there exists $0 < \delta < 1$ such that $te_0 + B_\delta$ is still contained in the frequency cone with slopes $s \in V_{s_0}$ for all $t \in \mathbb{R}$. Here, B_δ denotes the unit ball in $\mathbb{R}^2$ with radius δ around the origin. After picking δ, we can split the integral in (33) into

$$\int_{|\xi|<\delta t} \hat{f}(te_0 - \xi)\hat{\varphi}(\xi)\mathrm{d}\xi + \int_{|\xi|>\delta t} \hat{f}(te_0 - \xi)\hat{\varphi}(\xi)\mathrm{d}\xi.$$

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

$$\hat{f}(te_0-\xi)=O\left(|te_0-\xi|^{-N}\right)=O\left(|t|^{-N}\right),$$

and this suffices to establish that

$$\int_{|\xi|<\delta t}\hat{f}(te_0-\xi)\hat{\varphi}(\xi)\mathrm{d}\xi=O\left(|t|^{-N}\right).$$

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 (33) and shift them to $\hat{\varphi}$. Since $\hat{\varphi}$ is still a test function, this does not do any harm. Now we can establish the second part as by estimating

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

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 $\mathscr{C}_0$ be a cone containing the slope s_0. Then (t_0,s_0) is a regular directed point of $\hat{P}_{\mathscr{C}_0}f$, where $\hat{P}_{\mathscr{C}_0}$ denotes the frequency projection of f onto the frequency cone $\mathscr{C}_0$. 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 φ such that φf has fast Fourier decay in a frequency cone around s_0, i.e.,

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

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 Fig. 1b)

$$\left\{\eta+B_\delta|\eta|:\ \eta_2/\eta_1\in V_{s_0}\right\}\subset\mathscr{C}_0. \tag{34}$$

The inclusion (34) implies that

$$\xi\in\mathscr{C}_0^c\Rightarrow|\eta-\xi|>\delta|\eta|\ \eta_2/\eta_1\in V_{s_0}. \tag{35}$$

Write

$$\hat{\varphi} * \hat{f}(\eta) = \int_{\mathbb{R}^2} \chi_{\mathscr{C}_0} \hat{f}(\eta - \xi)\hat{\varphi}(\xi)\mathrm{d}\xi + \int_{\mathbb{R}^2} \chi_{\mathscr{C}_0^c} \hat{f}(\eta - \xi)\hat{\varphi}(\xi)\mathrm{d}\xi.$$

The statement is proven if we can show that

$$\int_{\mathbb{R}^2} \chi_{\mathscr{C}_0^c} \hat{f}(\eta - \xi)\hat{\varphi}(\xi)\mathrm{d}\xi = O\left(|\eta|^{-N}\right), \text{ for } \eta_2/\eta_1 \in V_{s_0}. \tag{36}$$

But this follows by writing (36) as

$$\int_{\mathbb{R}^2} \chi_{\mathscr{C}_0^c} \hat{f}(\xi)\hat{\varphi}(\eta - \xi)\mathrm{d}\xi$$

and using (35) together with the fact that

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

□

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 $\hat{P}_{\mathscr{C}} f$, $\hat{P}_{\mathscr{C}'} f$ 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 $\mathscr{C}_0$ a frequency cone around the direction with slope s_0. Then we have the equivalence*

$$\mathscr{S}\mathscr{H}_\psi f(a, s_0, t_0) = O\left(a^N\right) \Leftrightarrow \mathscr{S}\mathscr{H}_\psi\left(\hat{P}_{\mathscr{C}_0} f\right)(a, s_0, t_0) = O\left(a^N\right).$$

Proof. By linearity of the shearlet transform, we have

$$\mathscr{S}\mathscr{H}_\psi f(a, s_0, t_0) - \mathscr{S}\mathscr{H}_\psi\left(\hat{P}_{\mathscr{C}_0} f\right)(a, s_0, t_0) = \mathscr{S}\mathscr{H}_\psi\left(\hat{P}_{\mathscr{C}_0^c} f\right)(a, s_0, t_0).$$

But clearly (t_0, s_0) is a regular directed point of $\hat{P}_{\mathscr{C}_0^c} f$. Therefore, by Theorem 4 we can establish that

$$\mathscr{S}\mathscr{H}_\psi\left(\hat{P}_{\mathscr{C}_0^c} f\right)(a, s_0, t_0) = O\left(a^N\right)$$

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 φ 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 $V_{t_0} \subset\subset U_{t_0}$*[2]*. Consider the function*

$$g(x) = \int_{t \in U_{t_0}^c,\, s \in [-2,2],\, a \in [0,1]} \langle f, \psi_{a,s,t} \rangle \varphi(x) \psi_{a,s,t}(x) a^{-3} \mathrm{d}a \mathrm{d}s \mathrm{d}t.$$

Then

$$\hat{g}(\xi) = O\left(|\xi|^{-N}\right), \quad \xi \in \mathscr{C}. \tag{37}$$

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

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

By the projection slice theorem we need to show that

$$I^{(N)}(u) := \left(\frac{\mathrm{d}}{\mathrm{d}u}\right)^N I \in L^1(\mathbb{R}) \tag{38}$$

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 (37). By the product rule, $I^{(N)}$ can be written as a sum of terms of the form

$$\int_{t \in U(t_0)^c,\, s \in [-2,2],\, a \in [0,1]} \langle f, \psi_{a,s,t} \rangle \int_{\mathbb{R}} \left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \varphi(u - sx, s) a^{-j} \left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \psi\right)_{a,s,t} (u - sx, x) \mathrm{d}x a^{-3} \mathrm{d}a \mathrm{d}s \mathrm{d}t.$$

By the support properties of φ, 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 (18), we can establish that

$$\left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \psi\right)_{a,s,t} (y) = O\left(a^N |y - t|^{-N}\right). \tag{39}$$

Since we can assume that $y \in V_{t_0}$ and $t \in U_{t_0}^c$, we obtain (see Fig. 1a)

$$|y - t| \gtrsim |t - t_0|, \tag{40}$$

which, together with (39), establishes the desired claim. Note that by Fubini's theorem the application of the Radon transform is justified a posteriori. □

---

[2] With the notation $A \subset\subset B$, we mean that a tube of diameter δ around A is still contained in B for some $\delta > 0$.

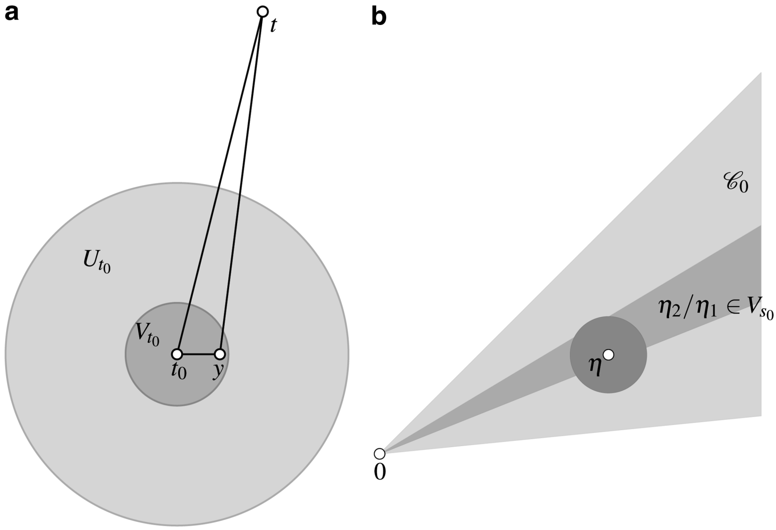

Fig. 1 (**a**) Illustration of the proof of (40). (**b**) Illustration of (34)

Theorem 5. *Assume that $f \in \mathscr{S}'$ 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), $|\mathscr{S}\mathscr{H}_\psi f(a,s,t)| = O(a^N)$ 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 $\mathscr{C}'$.*

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

By Theorem 3, we can represent f as

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

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

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

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

$$\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 \mathscr{S}\mathscr{H}_\psi f(a,s,t)\psi_{a,s,t}a^{-3}\mathrm{d}a\mathrm{d}s\mathrm{d}t.$$

To this end, we multiply this expression by a smooth bump function φ 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

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

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

$$I(u) := \mathscr{R}h(u, s_0).$$

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

$$\int_{t\in U_{t_0},\, s\in[-2,2],\, a\in[0,1]} \langle f, \psi_{a,s,t}\rangle \int_{\mathbb{R}} \left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \varphi(u - s_0 x, s) a^{-j} \left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \psi\right)_{a,s,t} (u - s_0 x, x)\mathrm{d}x a^{-3} \mathrm{d}a\mathrm{d}s\mathrm{d}t.$$

By making U_{t_0} (and the support of φ) 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

$$\int_{t\in Ut_0,\, s\in[s_0-\varepsilon,s_0+\varepsilon],\, a\in[0,1]} \langle f, \psi_{a,s,t}\rangle \int_{\mathbb{R}} \left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \varphi(u - s_0 x, s) a^{-j} \left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \psi\right)_{a,s,t} (u - s_0 x, x)\mathrm{d}x a^{-3} \mathrm{d}a\mathrm{d}s\mathrm{d}t = O(1). \quad (41)$$

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

$$\int_{t\in U_{t_0},\, s\in[s_0-\varepsilon,s_0+\varepsilon]^c,\, a\in[0,1]} \langle f, \psi_{a,s,t}\rangle a^{-j} \times \mathscr{R}\left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \varphi \left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j} \psi\right)_{a,s,t}\right)(u, s_0) a^{-3} \mathrm{d}a\mathrm{d}s\mathrm{d}t. \quad (42)$$

Note that we can write

$$\mathscr{R}\left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j}\varphi\left(\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j}\psi\right)_{a,s,t}\right)(u,s_0)=\langle\tilde{\delta}_{u,s_0},\theta_{a,s,t}\rangle, \tag{43}$$

where

$$\theta:=\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j}\psi$$

and

$$\tilde{\delta}_{u,s_0}:=\left(\frac{\mathrm{d}}{\mathrm{d}x_1}\right)^{N-j}\varphi\delta_{x_1=u-s_0x_2}.$$

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

$$\{(x_1,x_2,s):\ x_1=u-s_0x_2,\ s=s_0\},$$

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

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

By (43), this implies that also the expression (43) is bounded. Together with (41), 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 Theorems 4 and 5, we have finally proved Theorem 2:

Corollary 1. *Theorem 2 holds true.*

Remark 6. 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 ψ are required. The details are given in [9]. ♢

Remark 7. The shearlet transform (and similar related transform like for instance the curvelet transform) are able to characterize even finer notions of microlocal smoothness. We say that that $f\in\mathscr{S}'$ belongs to the microlocal Sobolev space $H^\alpha(t_0,s_0)$ if there exists a smooth bump function $\varphi\in\mathscr{S}$ around t_0 and a frequency cone $\mathscr{C}_{s_0}$ around s_0 such that

$$\int_{\mathscr{C}_{s_0}}|\xi|^{2\alpha}\left|(\varphi f)^\wedge(\xi)\right|^2\mathrm{d}\xi<\infty.$$

Define the shearlet square function

$$S_2^\alpha(f)(t,s):=\left(\int_0^1\left|\mathscr{S}\mathscr{H}_\psi(f)(a,s,t)\right|a^{-2\alpha}\frac{\mathrm{d}a}{a^3}\right)^{1/2}.$$

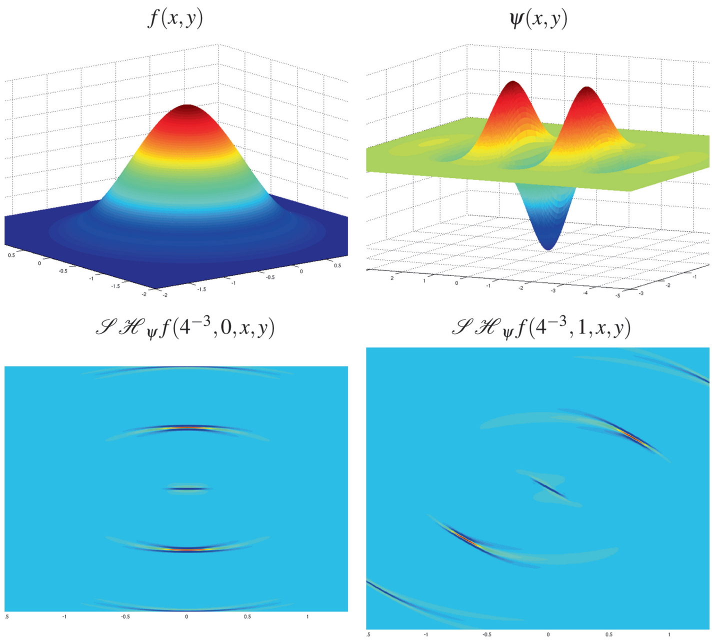

Fig. 2 This figure illustrates the main result, Theorem 2. On the *top left*, we show the function f to be analyzed—a rotated quadratic B-spline curve which has curvature discontinuities at integer points. Therefore, the wavefront set of f consists of the origin with all directions attached and the tangents 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 nonnegligible shearlet coefficients

Then it holds that

$$f \in H^{\alpha}(t_0, s_0) \Leftrightarrow S_2^{\alpha}(f) \in L^2(\mathscr{N}),$$

for a neighborhood $\mathscr{N}$ of (t_0, s_0). The proof is similar to the proof of the corresponding curvelet result [3] and will be given elsewhere. $\diamondsuit$

Acknowledgments

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

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. I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.
7. S. R. Deans. *The Radon Transform and Some of Its Applications*. John Wiley and Sons, 1983.
8. K. Gröchenig. *Foundations of Time-Frequency Analysis*. Birkhäuser, 2001.
9. P. Grohs. Continuous shearlet frames and resolution of the wavefront set. *Monatshefte für Mathematik*, 164(4):393–426, 2011.
10. P. Grohs. Continuous shearlet tight frames. *Journal of Fourier Analysis and Applications*, 17(3):506–518, 2011.
11. K. Guo and D. Labate. Optimally sparse multidimensional representation using shearlets. *SIAM journal on mathematical analysis*, 39(1):298–318, 2008.
12. K. Guo and D. Labate. Representation of Fourier integral operators using shearlets. *Journal of Fourier Analysis and Applications*, 14(3):327–371, 2008.
13. K. Guo and D. Labate. Characterization and analysis of edges using the continuous shearlet transform. *SIAM journal on Imaging Sciences*, 2:959–986, 2009.
14. K. Guo, D. Labate, and W.-Q. Lim. Edge analysis and identification using the continuous shearlet transform. *Applied and Computational Harmonic Analysis*, 27(1):24–46, 2009.
15. E. Hewitt and K. Ross. *Abstract Harmonic Analysis I*. Springer, 1979.
16. M. Holschneider and P. Tchamitchian. Pointwise analysis of Riemann's "nondifferentiable" function. *Inventiones Mathematicae*, 105:157–175, 1991.
17. L. Hörmander. *The Analysis of linear Partial Differential Operators*. Springer, 1983.
18. G. Kutyniok and D. Labate. Resolution of the wavefront set using continuous shearlets. *Transactions of the American Mathematical Society*, 361:2719–2754, 2009.
19. G. Kutyniok and W.-Q. Lim. Compactly supported shearlets are optimally sparse. Technical report, 2010.
20. S. Mallat. *A wavelet tour of signal processing*. Academic Press, 2003.
21. A. Martinez. *An Introduction to Semiclassical and Microlocal Analysis*. Springer, 2002.
22. W. Rudin. *Functional Analysis*. Mc Graw-Hill, 1991.
23. H. Smith. A Hardy space for Fourier integral operators. *Journal of Geometric Analysis*, 8:629–653, 1998.

http://www.springer.com/978-0-8176-8315-3

Shearlets
Multiscale Analysis for Multivariate Data
Kutyniok, G.; labate, d. (Eds.)
2012, XIX, 328 p. 50 illus., 19 illus. in color., Hardcover
ISBN: 978-0-8176-8315-3
A product of Birkhäuser Basel