TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 361, Number 5, May 2009, Pages 2719–2754
S 0002-9947(08)04700-4
Article electronically published on October 24, 2008

RESOLUTION OF THE WAVEFRONT SET USING CONTINUOUS SHEARLETS

GITTA KUTYNIOK AND DEMETRIO LABATE

ABSTRACT. It is known that the Continuous Wavelet Transform of a distribution f decays rapidly near the points where f is smooth, while it decays slowly near the irregular points. This property allows the identification of the singular support of f. However, the Continuous Wavelet Transform is unable to describe the geometry of the set of singularities of f and, in particular, identify the wavefront set of a distribution. In this paper, we employ the same framework of affine systems which is at the core of the construction of the wavelet transform to introduce the Continuous Shearlet Transform. This is defined by $\mathcal{SH}_\psi f(a,s,t) = \langle f \psi_{ast} \rangle$, where the analyzing elements ψ_{ast} are dilated and translated copies of a single generating function ψ. The dilation matrices form a two-parameter matrix group consisting of products of parabolic scaling and shear matrices. We show that the elements $\{\psi_{ast}\}$ form a system of smooth functions at continuous scales $a > 0$, locations $t \in \mathbb{R}^2$, and oriented along lines of slope $s \in \mathbb{R}$ in the frequency domain. We then prove that the Continuous Shearlet Transform does exactly resolve the wavefront set of a distribution f.

1. Introduction

It is well known that, provided ψ is a "nice" continuous wavelet on $\mathbb{R}^n$ and f is a distribution that is smooth apart from a discontinuity at a point $x_0 \in \mathbb{R}^n$, the Continuous Wavelet Transform

$$\mathcal{W}_\psi f(a,t) = a^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x)\, \psi\left(a^{-1}(x-t)\right)\, dx, \quad a > 0, t \in \mathbb{R}^n$$

decays rapidly as $a \to 0$ unless t is near x_0 [19, 25]. As a consequence, the Continuous Wavelet Transform is able to resolve the *singular support* of a distribution f, i.e., to identify the set of points where f is not regular. However, the transform $\mathcal{W}_\psi f(a,t)$ is unable to provide additional information about the *geometry* of the singular support. In many situations, it is essential to not only identify the location of a certain distributed singularity, but also its orientation in the sense of resolving the wavefront set. This is, for instance, particularly useful in the study of the propagation of singularities associated with partial differential equations [20, 27].

Received by the editors April 24, 2006 and, in revised form, November 1, 2007.
2000 *Mathematics Subject Classification.* Primary 42C15; Secondary 42C40.
Key words and phrases. Analysis of singularities, continuous wavelets, curvelets, directional wavelets, shearlets, wavefront set, wavelets.
The first author acknowledges support from Deutsche Forschungsgemeinschaft (DFG), Grant KU 1446/5-1.
The second author acknowledges support from NSF Grant DMS 0604561.

Historically, the idea of using continuous transforms to identify both the location and the geometry of the set of singularities of a distribution can be traced back to the notion of *wave packet transforms*, introduced independently by Bros and Iagolnitzer [1] and Córdoba and Fefferman [10]. More recently, Smith [26] and Candès and Donoho [6, 7] have introduced continuous transforms, which use parabolic scaling and rotations in polar coordinates and have the ability to resolve the wavefront set of a distribution. In particular, the Continuous Curvelet Transform of Candès and Donoho is closely related to the successful discrete curvelet construction [5]. However, the Continuous Curvelet Transform does not have the simple mathematical structure of the wavelet transform. For instance, it requires infinitely many generators, thereby losing useful properties of the Continuous Wavelet Transform such as being associated with an affine group structure.[1] This raises the question, whether it is possible to construct a genuinely "wavelet-like" continuous transform, which is capable of precisely resolving the wavefront set of distributions while being equipped with the same simple affine structure of the Continuous Wavelet Transform.

Another motivation for our investigation and the use of the framework of affine systems comes from the study of discrete wavelets, and, more specifically, their ability to approximate efficiently smooth functions with singularities. This property is closely related to the micro-local properties of the Continuous Wavelet Transform. To illustrate this point, consider a one-dimensional function f that is smooth apart from a discontinuity at a point x_0 and consider its wavelet representation:

$$f = \sum_{j,k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle\, \psi_{j,k},$$

where $\psi_{j,k}(x) = 2^{j/2}\,\psi(2^j x - k)$ and ψ is a "nice" wavelet. Notice that the coefficients of the representation are just samples of the Continuous Wavelet Transform at points $(2^{-j}, 2^{-j}k)$, for $j,k \in \mathbb{Z}$, that is, $\mathcal{W}_\psi f(2^{-j}, 2^{-j}k) = \langle f, \psi_{j,k}\rangle$. Since the elements $\mathcal{W}_\psi f(2^{-j}, 2^{-j}k)$ decay rapidly for $j \to \infty$ unless k is near x_0, it follows that one can approximate f accurately by using very few coefficients of the wavelet representation. Indeed, the wavelet representation is optimally sparse for this type of functions (cf. [24, Ch.9]). However, the situation is significantly different in higher dimensions, where more general discontinuities are usually present or even dominant, and traditional wavelets are not equally effective. Consider, for example, the wavelet representation of a two-dimensional function that is smooth away from a discontinuity along a curve. Because the discontinuity is spatially distributed, it interacts extensively with the elements of the wavelet basis, and thus "many" wavelet coefficients are needed to represent the function accurately. In fact, this is a manifestation of the fact that the Continuous Wavelet Transform is unable to deal with distributed discontinuities effectively. As pointed out by several authors (see [4, 5]), to overcome this limitation, one needs a transform with the ability to capture the geometry of multidimensional phenomena. In this paper, we will show that this can be achieved by properly reexamining the notion of a continuous wavelet transform in higher dimensions.

Indeed, the *Continuous Shearlet Transform*, which is introduced in this paper, fully exploits the framework of the affine group on $\mathbb{R}^2$ to precisely capture the

[1]Recall that also the corresponding discrete curvelets have no affine structure and are not associated to a Multiresolution Analysis.

geometric information of two-dimensional functions. More precisely, the Continuous Shearlet Transform maps a tempered distribution $f \in \mathcal{S}'(\mathbb{R}^2)$ to $\mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{ast} \rangle$ on the transform domain $\{(a,s,t) : a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2\}$, where the analyzing elements ψ_{ast} are dilated and translated copies of a single generating function $\psi \in L^2(\mathbb{R}^2)$. This generator ψ is called a *continuous shearlet*, and is chosen to be arbitrarily smooth with compact support in the frequency domain. The dilation matrices consist of the product of a parabolic scaling matrix associated with some $a > 0$ and a shear matrix associated with some $s \in \mathbb{R}$. As a result, the elements ψ_{ast} constitute an *affine system* of well-localized waveforms at various scales a, orientations controlled by s and spatial locations t. Due to the parabolic scaling, the elements ψ_{ast} become increasingly thin as $a \to 0$, and this anisotropic behavior allows them to detect the singularities along curves. As a result, the Continuous Shearlet Transform is able to identify not only the singular support of a distribution f, but also the orientation of distributed singularities along curves. In particular, the decay properties of the Continuous Shearlet Transform as $a \to 0$ precisely characterize the *wavefront set* of f (see Section 5) with the translation parameter detecting the location and the shear parameter s detecting the orientation of a singularity.

We would also like to mention that the study of the discrete analog of the Continuous Shearlet Transform is currently being developed by the authors and their collaborators. In particular, by employing the advantageous properties of the shear operator over the rotation operator, it was recently shown that discrete shearlets are associated with a Multiresolution Analysis and with directional subdivision schemes generalizing those of traditional wavelets. This is very relevant for the development of fast algorithmic implementations. In addition, shearlets provide optimally sparse representations for bivariate functions with discontinuities along curves. We refer to [13, 15, 16, 17, 18, 22, 21] for more detail about the research about shearlets.

Note added: Following submission of this paper, it was shown in [11] that the Continuous Shearlet Transform is related with a locally compact group, the so-called *Shearlet group*, in the sense of the continuous shearlet systems being generated by a strongly continuous, irreducible, square-integrable representation of this group. This additional rich mathematical structure enables, for instance, the application of uncertainty principles to tune the accuracy of the transform [11], and of the coorbit theory to study smoothness spaces, so-called *Shearlet Coorbit Spaces*, associated with the decay of the shearlet coefficients [12].

The paper is organized as follows. In Section 2 we recall the basic properties of affine systems on $\mathbb{R}^n$ and the Continuous Wavelet Transform, and then introduce the Continuous Shearlet Transform (Section 3). In Section 4 we apply this new transform to several examples of distributions containing different types of singularities. The main result of this paper is proved in Section 5, where we show that the Continuous Shearlet Transform exactly characterizes the wavefront set of a distribution. Finally, in Section 6, we discuss several variants and generalizations of our construction.

1.1. **Notation and definitions.** We adopt the convention that $x \in \mathbb{R}^n$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^n$ is a row vector, i.e., $\xi = (\xi_1, \ldots, \xi_n)$. A vector x multiplying a matrix $a \in GL_n(\mathbb{R})$ on the right is understood to be a column vector, while a vector ξ multiplying a on the left is a row vector. Thus, $ax \in \mathbb{R}^n$ and $\xi a \in \widehat{\mathbb{R}}^n$. The Fourier transform is defined as

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \xi x}\, dx,$$

where $\xi \in \widehat{\mathbb{R}}^n$, and the inverse Fourier transform is

$$\check{f}(x) = \int_{\widehat{\mathbb{R}}^n} f(\xi)\, e^{2\pi i \xi x}\, d\xi.$$

We consider three fundamental operators on $L^2(\mathbb{R}^n)$: the *translations* T_y : $(T_y f)(x) = f(x-y)$, where $y \in \mathbb{R}^n$; the *dilations* D_A : $(D_A f)(x) = |\det A|^{-1/2} f(A^{-1}x)$, where $A \in GL_n(\mathbb{R})$; and the *modulations* $M_z : (M_z \hat{f})(\xi) = e^{2\pi i \xi z} \hat{f}(\xi)$, where $z \in \mathbb{R}^n$.

The following proposition, which is easily verified, states some basic properties of the translation and dilation operators.

Proposition 1.1. *Let $G = \{U = D_A\, T_y : (A, y) \in GL_n(\mathbb{R}) \times \mathbb{R}^n\}$. Then G is a subgroup of the group of unitary operators on $L^2(\mathbb{R}^n)$ which is preserved by the action of the operator $U \mapsto \widehat{U}$, where $\widehat{U}\, \hat{f} = (Uf)^\wedge$. In particular, we have:*

(i) $D_A\, T_y = T_{Ay}\, D_A$;
(ii) $D_{A_1}\, D_{A_2} = D_{A_1 A_2}$, *for each* $A_1, A_2 \in GL_n(\mathbb{R})$;
(iii) *for* $U = D_A\, T_y$, *then* $\widehat{U} = \widehat{D}_A\, M_{-y}$, *where* $\widehat{D}_A \hat{f}(\xi) = |\det A|^{1/2}\, \hat{f}(\xi\, A)$;
(iv) *for* $S \subset \widehat{\mathbb{R}}^n$ *a measurable set, and* $L^2(S) = \{\hat{f} \in L^2(\widehat{\mathbb{R}}^n) : \operatorname{supp} \hat{f} \subseteq S\}$, *we have:* $\widehat{D}_A\, L^2(S) = L^2(SA^{-1})$.

Recall that a countable collection $\{\psi_i\}_{i\in I}$ in a Hilbert space $\mathcal{H}$ is a *Parseval frame* (sometimes called a *tight frame*) for $\mathcal{H}$ if

$$\sum_{i\in I} |\langle f, \psi_i\rangle|^2 = \|f\|^2, \qquad \text{for all } f \in \mathcal{H}.$$

This is equivalent to the reproducing formula $f = \sum_i \langle f, \psi_i\rangle\, \psi_i$, for all $f \in \mathcal{H}$, where the series converges in the norm of $\mathcal{H}$. This shows that a Parseval frame provides a basis-like representation even though a Parseval frame need not be a basis in general. We refer the reader to [8, 9] for more details about frames.

For any $E \subset \widehat{\mathbb{R}}^n$, we denote by $L^2(E)^\vee$ the space $\{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset E\}$.

2. Affine systems and wavelets

2.1. **One-dimensional continuous wavelet transform.** Let $\mathbb{A}_1$ be the *affine group* associated with $\mathbb{R}$, consisting of all pairs (a, t), $a, t \in \mathbb{R}, a > 0$, with group operation $(a, t) \cdot (a', t') = (aa', t + at')$. The *(continuous) affine systems* generated by $\psi \in L^2(\mathbb{R})$ are obtained from the action of the *quasi-regular representation* $\pi_{(a,t)}$ of $\mathbb{A}_1$ on $L^2(\mathbb{R})$; that is,

$$\big\{\psi_{a,t}(x) = \pi_{(a,t)}\, \psi(x) = T_t\, D_a\, \psi(x) : (a, t) \in \mathbb{A}_1\big\},$$

where the *translation operator* T_t is defined by $T_t \psi(x) = \psi(x - t)$ and the *dilation operator* D_a is defined by $D_a \psi(x) = a^{-1/2} \psi(a^{-1}x)$.

It was observed by Calderón [2] that, if ψ satisfies the *admissibility* condition

$$\int_0^\infty |\hat{\psi}(a\xi)|^2 \frac{da}{a} = 1 \quad \text{for a.e. } \xi \in \mathbb{R}, \tag{2.1}$$

then any $f \in L^2(\mathbb{R})$ can be recovered via the reproducing formula:

$$f = \int_{\mathbb{A}_1} \langle f, \psi_{a,t} \rangle \, \psi_{a,t} \, d\mu(a,t),$$

where $d\mu(a,t) = dt \, \frac{da}{a^2}$ is the left Haar measure of $\mathbb{A}_1$. Here the Fourier transform is defined by $\hat{\psi}(\xi) = \int \psi(x)\, e^{-2\pi i \xi x}\, dx$. As usual, $\check{\psi}$ will denote the inverse Fourier transform. The function ψ is called a *continuous wavelet*, if ψ satisfies (2.1), and $\mathcal{W}_\psi f(a,t) = \langle f, \psi_{a,t} \rangle$ is the *Continuous Wavelet Transform* of f. We refer to [14] for more details about this.

Discrete affine systems and wavelets are obtained by "discretizing" appropriately the corresponding continuous systems. In fact, by replacing $(a,t) \in \mathbb{A}_1$ with the discrete set $(2^j, 2^j m)$, $j, m \in \mathbb{Z}$, one obtains the discrete dyadic affine system

$$\{\psi_{j,m}(x) = T_{2^j m}\, D_2^j \, \psi(x) = D_2^j \, T_m \, \psi(x) : \; j, m \in \mathbb{Z}\}, \tag{2.2}$$

and ψ is called a *wavelet* if (2.2) is an orthonormal basis or, more generally, a Parseval frame for $L^2(\mathbb{R})$.

Recall that a countable collection $\{\psi_i\}_{i \in I}$ in a Hilbert space $\mathcal{H}$ is a *Parseval frame* (sometimes called a *tight frame*) for $\mathcal{H}$ if $\sum_{i\in I} |\langle f, \psi_i \rangle|^2 = \|f\|^2$ for all $f \in \mathcal{H}$. This is equivalent to the reproducing formula $f = \sum_{i \in I} \langle f, \psi_i \rangle \, \psi_i$ for all $f \in \mathcal{H}$, where the series converges in the norm of $\mathcal{H}$. Thus Parseval frames provide basis-like representations even though a Parseval frame need not be a basis in general. We refer the reader to [8, 9] for more details about frames.

2.2. **Higher-dimensional continuous wavelet transform.** The natural way of extending the theory of affine systems to higher dimensions is by replacing $\mathbb{A}_1$ with the *full affine group of motions on* $\mathbb{R}^n$, $\mathbb{A}_n$, consisting of the pairs $(M,t) \in GL_n(\mathbb{R}) \times \mathbb{R}^n$ with the group operation $(M,t) \cdot (M', t') = (MM', t + Mt')$. Similarly to the one-dimensional case, the affine systems generated by $\psi \in L^2(\mathbb{R}^n)$ are given by

$$\{\psi_{M,t}(x) = T_t \, D_M \, \psi(x) : \; (M,t) \in \mathbb{A}_n\},$$

where here the *dilation operator* D_M is defined by $D_M \, \psi(x) = |\det M|^{-\frac{1}{2}} \psi(M^{-1}x)$. The generalization of the Calderón admissibility condition to higher dimensions and the construction of multidimensional wavelets is a far more complex task than the corresponding one-dimensional problem, and yet not fully understood. We refer to [23, 29] for more details.

Now let G be a subset of $GL_n(\mathbb{R})$ and define $\Lambda \subseteq \mathbb{A}_n$ by $\Lambda = \{(M,t) : M \in G, t \in \mathbb{R}^n\}$. If there exists a function $\psi \in L^2(\mathbb{R}^n)$ such that, for all $f \in L^2(\mathbb{R}^n)$, we have:

$$f = \int_{\mathbb{R}^n} \int_G \langle f, T_t \, D_M \, \psi \rangle \, T_t \, D_M \, \psi \, d\lambda(M) \, dt, \tag{2.3}$$

where λ is a measure on G, then ψ is a *continuous wavelet* with respect to Λ. The following result, that is a simple modification of Theorem 2.1 in [29], gives an exact characterization of all those $\psi \in L^2(\mathbb{R}^n)$ that are continuous wavelets with respect to Λ. The proof of this theorem is reported in the Appendix.

Theorem 2.1. *Equality* (2.3) *is valid for all* $f \in L^2(\mathbb{R}^n)$ *if and only if, for all* $\xi \in \mathbb{R}^n \setminus \{0\}$,

$$\Delta(\psi)(\xi) = \int_G |\hat{\psi}(M^t \xi)|^2 \, |\det M| \, d\lambda(M) = 1. \tag{2.4}$$

The choice of the measure λ on G is not unique. If G is not simply a subset of $GL_n(\mathbb{R})$, but also a subgroup, then we can use the left Haar measure on G which is unique up to a multiplicative constant. Also, observe that Theorem 2.1 extends to functions on subspaces of $L^2(\mathbb{R}^n)$ of the form

$$L^2(V)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset V\}.$$

2.3. **Localization of wavelets.** The decay properties of the functions $\psi_{M,t} = T_t \, D_M \, \psi$, where $\hat{\psi} \in C_0^\infty$, are described by the following proposition.

Proposition 2.2. *Suppose that* $\psi \in L^2(\mathbb{R}^n)$ *is such that* $\hat{\psi} \in C_0^\infty(R)$, *where* $R = \operatorname{supp} \hat{\psi} \subset \mathbb{R}^n$. *Then, for each* $k \in \mathbb{N}$, *there is a constant* C_k *such that, for any* $x \in \mathbb{R}^n$, *we have*

$$|\psi_{M,t}(x)| \le C_k \, |\det M|^{-\frac{1}{2}} \, (1 + |M^{-1}(x - t)|^2)^{-k}.$$

In particular, $C_k = k \, m(R) \left(\|\hat{\psi}\|_\infty + \|\triangle^k \hat{\psi}\|_\infty \right)$, *where* $\triangle = \sum_{i=1}^n \frac{\partial^2}{\partial \xi_i^2}$ *is the frequency domain Laplacian operator and* $m(R)$ *is the Lebesgue measure of* R.

The proof of this proposition relies on the following known observation, whose proof is included for completeness.

Lemma 2.3. *Let* g *be such that* $\hat{g} \in C_0^\infty(R)$, *where* $R \subset \mathbb{R}^n$ *is the* $\operatorname{supp} \hat{g}$. *Then, for each* $k \in \mathbb{N}$, *there is a constant* C_k *such that for any* $x \in \mathbb{R}^n$,

$$|g(x)| \le C_k \, (1 + |x|^2)^{-k}.$$

In particular, $C_k = k \, m(R) \left(\|\hat{g}\|_\infty + \|\triangle^k \hat{g}\|_\infty \right)$.

Proof. Since $g(x) = \int_R \hat{g}(\xi) \, e^{2\pi i \xi x} \, d\xi$, then, for every $x \in \mathbb{R}^2$,

$$|g(x)| \le m(R) \, \|\hat{g}\|_\infty. \tag{2.5}$$

An integration by parts shows that

$$\int_R \triangle \hat{g}(\xi) \, e^{2\pi i \xi x} \, d\xi = -(2\pi)^2 \, |x|^2 \, g(x)$$

and thus, for every $x \in \mathbb{R}^2$,

$$(2\pi \, |x|)^{2k} \, |g(x)| \le m(R) \, \|\triangle^k \hat{g}\|_\infty. \tag{2.6}$$

Using (2.5) and (2.6), we have

$$\left(1 + (2\pi \, |x|)^{2k}\right) |g(x)| \le m(R) \left(\|\hat{g}\|_\infty + \|\triangle^k \hat{g}\|_\infty \right). \tag{2.7}$$

Observe that, for each $k \in \mathbb{N}$,

$$(1 + |x|^2)^k \le \left(1 + (2\pi)^2 \, |x|^2\right)^k \le k \left(1 + (2\pi \, |x|)^{2k}\right).$$

Using this last inequality and (2.7), we have that for each $x \in \mathbb{R}^n$,

$$|g(x)| \le k\, m(R)\, (1+|x|^2)^{-k} \left(\|\hat{g}\|_\infty + \|\triangle^k \hat{g}\|_\infty\right).$$

□

A simple re-scaling argument now proves Proposition 2.2.

Proof of Proposition 2.2. A direct computation gives:

$$\begin{aligned} \psi(M^{-1}(x-t)) &= \int_R \hat{\psi}(\xi)\, e^{2\pi i M^{-1}(x-t)\xi}\, d\xi \\ &= \int_R \hat{\psi}(\xi)\, e^{2\pi i (x-t) M^{-t}\xi}\, d\xi \\ &= \int_{(M^t)^{-1}R} \hat{\psi}(M^t\eta)\, e^{2\pi i (x-t)\eta}\, |\det M|\, d\eta. \end{aligned}$$

It follows that

$$|\psi(M^{-1}(x-t))| \le m((M^t)^{-1}R)\, |\det M|\, \|\hat{\psi}(M^t\cdot)\|_\infty = m(R)\, \|\hat{\psi}\|_\infty.$$

Using a simple modification of the argument in Lemma 2.3, we have that

$$(2\pi\, |M^{-1}(x-t)|)^{2k}\, |\psi(M^{-1}(x-t))| \le m(R)\, \|\triangle^k \hat{\psi}\|_\infty.$$

Next, arguing again as in Lemma 2.3 we have that

$$|\psi(M^{-1}(x-t))| \le k\, m(R)\, (1+|M^{-1}(x-t)|^2)^{-k} \left(\|\hat{\psi}\|_\infty + \|\triangle^k \hat{\psi}\|_\infty\right).$$

This completes the proof. □

3. Continuous shearlet transform

3.1. **Definition.** In this paper, we will be interested in the affine systems obtained when Λ is a subset of $\mathbb{A}_2$ of the form

$$\Lambda = \{(M,t) :\, M \in G, t \in \mathbb{R}^2\}, \tag{3.1}$$

and $G \subset GL_2(\mathbb{R})$ is the set of matrices:

$$G = \left\{ M = M_{as} = \begin{pmatrix} a & -\sqrt{a}\, s \\ 0 & \sqrt{a} \end{pmatrix}, \quad a \in I,\, s \in S \right\}, \tag{3.2}$$

where $I \subset \mathbb{R}^+$, $S \subset \mathbb{R}$. It is useful to notice that the matrices M can be factorized as $M = B\,A$, where B is the *shear matrix* $B = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$ and A is the diagonal matrix $A = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$. In particular, A produces *parabolic scaling*; that is, $f(Ax) = f\left(A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)$ leaves invariant the parabola $x_1 = x_2^2$. Thus, the matrix M can be interpreted as the superposition of parabolic scaling and shear transformation.

We will now consider those functions, which satisfy (2.1) for the subset Λ of the affine group, given by (3.1). In order to distinguish these functions from a general continuous wavelet, in the following we will refer to them as *continuous shearlets*. We will consider two situations, corresponding to $I = \mathbb{R}^+$, $S = \mathbb{R}$ or $I = \{a : 0 \le a \le 1\}$, $S = \{s \in \mathbb{R} : |s| \le s_0\}$, for some $s_0 > 0$.

For $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, $\xi_2 \neq 0$, let ψ be given by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\tfrac{\xi_2}{\xi_1}). \tag{3.3}$$

Proposition 3.1. *Let Λ be given by* (3.1) *and* (3.2) *with $I = \mathbb{R}^+$, $S = \mathbb{R}$, and $\psi \in L^2(\mathbb{R}^2)$ be given by* (3.3) *where:*

(i) $\psi_1 \in L^2(\mathbb{R})$ *satisfies the Calderón condition* (2.1)*;*
(ii) $\|\psi_2\|_{L^2} = 1$.

Then ψ is a continuous shearlet for $L^2(\mathbb{R}^2)$ with respect to Λ.

Proof. A direct computation shows that $M^t(\xi_1,\xi_2)^t = (a\xi_1, a^{1/2}(\xi_2 - s\xi_1))^t$. By choosing as measure $d\lambda(M) = \frac{da}{|\det M|^2}\, ds$, the admissibility condition (2.4) becomes

$$\Delta(\psi)(\xi) = \int_{\mathbb{R}} \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2\, |\hat{\psi}_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1} - s))|^2\; a^{-\frac{3}{2}} da\, ds = 1. \tag{3.4}$$

Thus, by Theorem 2.1, to show that ψ is a continuous shearlet it is sufficient to show that (3.4) is satisfied. Using the assumption on ψ_1 and ψ_2, we have:

$$\begin{aligned} \Delta(\psi)(\xi) &= \int_{\mathbb{R}} \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2\, |\hat{\psi}_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1} - s))|^2\; a^{-\frac{3}{2}} da\, ds \\ &= \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2 \left(\int_{\mathbb{R}} |\hat{\psi}_2(a^{-\frac{1}{2}}\, \tfrac{\xi_2}{\xi_1} - s)|^2\; ds \right) \frac{da}{a} \\ &= \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2\, \frac{da}{a} = 1 \quad \text{for a.e. } \xi = (\xi_1,\xi_2) \in \mathbb{R}^2. \end{aligned}$$

This shows that equality (3.4) is satisfied and, hence, ψ is a continuous shearlet. □

If the set S is not all of $\mathbb{R}$, then we need some additional assumptions on ψ. Consider the subspace of $L^2(\mathbb{R}^2)$ given by $L^2(C)^\vee = \{f \in L^2(\mathbb{R}^2) : \operatorname{supp} \hat{f} \subset C\}$, where

$$C = \{(\xi_1,\xi_2) \in \mathbb{R}^2 : |\xi_1| \geq 2 \text{ and } |\tfrac{\xi_2}{\xi_1}| \leq 1\}.$$

We have the following result.

Proposition 3.2. *Let Λ be given by* (3.1) *and* (3.2) *with $I = \{a : 0 \leq a \leq 1\}$, $S = \{s \in \mathbb{R} : |s| \leq 2\}$, and $\psi \in L^2(\mathbb{R}^2)$ be given by* (3.3) *where:*

(i) $\psi_1 \in L^2(\mathbb{R})$ *satisfies the Calderón condition* (2.1)*, and* $\operatorname{supp} \hat{\psi}_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$*;*
(ii) $\|\psi_2\|_{L^2} = 1$ *and* $\operatorname{supp} \hat{\psi}_2 \subset [-1, 1]$.

Then ψ is a continuous shearlet for $L^2(C)^\vee$ with respect to Λ, that is, for all $f \in L^2(C)^\vee$,

$$f(x) = \int_{\mathbb{R}^2} \int_{-2}^{2} \int_0^1 \langle f, \psi_{ast} \rangle\, \psi_{ast}(x)\, \frac{da}{a^3}\, ds\, dt.$$

Proof. We apply again Theorem 2.1 to functions on $L^2(C)^\vee$. Using the assumptions on ψ_2, S and I we have that, for $\xi \in C$:

$$\int_{\frac{1}{\sqrt{a}}(\frac{\xi_2}{\xi_1} - 2)}^{\frac{1}{\sqrt{a}}(\frac{\xi_2}{\xi_1} + 2)} |\hat{\psi}_2(s)|^2\, ds = \int_{-1}^{1} |\hat{\psi}_2(s)|^2\, ds = 1.$$

Thus, for a.e. $\xi \in C$ we have that

$$\begin{aligned}\Delta(\psi)(\xi) &= \int_{-2}^{2}\int_{0}^{1} |\hat\psi(M_{as}^t\xi)|^2\, a^{-\frac{3}{2}}\, da\, ds\\ &= \int_{-2}^{2}\int_{0}^{1} |\hat\psi_1(a\xi_1)|^2\, |\hat\psi_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1}-s))|^2 a^{-\frac{3}{2}}\, da\, ds\\ &= \int_0^1 |\hat\psi_1(a\xi_1)|^2 \int_{\frac{1}{\sqrt{a}}(\frac{\xi_2}{\xi_1}-2)}^{\frac{1}{\sqrt{a}}(\frac{\xi_2}{\xi_1}+2)} |\hat\psi_2(s)|^2\, ds\, \frac{da}{a}\\ &= \int_0^1 |\hat\psi_1(a\xi_1)|^2\, \frac{da}{a}.\end{aligned}$$

Since $\xi_1 \geq 2$, using the assumptions on the support of $\hat\psi_1$ and condition (2.1), from the last expression we have that, for a.e. $\xi \in C$,

$$\Delta(\psi)(\xi) = \int_0^{\xi_1} |\hat\psi_1(a)|^2\, \frac{da}{a} = \int_{\frac{1}{2}}^{2} |\hat\psi_1(a)|^2\, \frac{da}{a} = \int_0^{\infty} |\hat\psi_1(a)|^2\, \frac{da}{a} = 1.$$

This shows that the admissibility condition (2.4) for this system is satisfied and this completes the proof. □

There are several examples of functions ψ_1 and ψ_2 satisfying the assumptions of Proposition 3.1 as well as Proposition 3.2. In addition, we can choose ψ_1, ψ_2 such that $\hat\psi_1$, $\hat\psi_2$ are real-valued and belong to C_0^∞ (see [15, 18] for the construction of these functions).

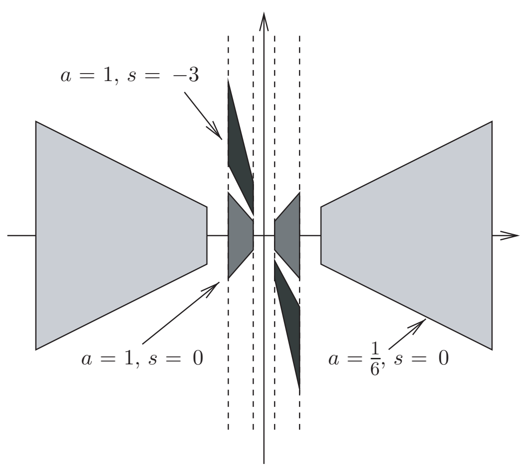

FIGURE 1. Support of the shearlets $\hat\psi_{ast}$ (in the frequency domain) for different values of a and s.

Now we can define the *Continuous Shearlet Transform:*

Definition 3.3. Let $\psi \in L^2(\mathbb{R}^2)$ be given by (3.3) where:

(i) $\psi_1 \in L^2(\mathbb{R})$ satisfies the Calderón condition (2.1), and $\hat\psi_1 \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp}\hat\psi_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$;

(ii) $\|\psi_2\|_{L^2} = 1$, and $\hat\psi_2 \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp}\hat\psi_2 \subset [-1, 1]$ and $\hat\psi_2 > 0$ on $(-1, 1)$.

The set of functions generated by ψ under the action of Λ, namely:

$$\{\psi_{ast} = T_t D_{M_{as}} \psi = a^{-\frac{3}{4}}\, \psi\left(M_{as}^{-1}(\cdot\, - t)\right) : a \in I \subset \mathbb{R}^+, s \in S \subset \mathbb{R}, t \in \mathbb{R}^2\},$$

where M_{as} was defined in (3.2), is called a *continuous shearlet system.* The *Continuous Shearlet Transform* of f is defined by

$$\mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{ast} \rangle, \quad a \in I \subset \mathbb{R}^+, s \in S \subset \mathbb{R}, t \in \mathbb{R}^2.$$

Observe that, unlike the traditional wavelet transform which depends only on scale and translation, the shearlet transform is a function of three variables, that is, the scale a, the shear s and the translation t. Many properties of the continuous shearlets are more evident in the frequency domain. A direct computation shows that

$$\begin{aligned} \hat{\psi}_{ast}(\xi) &= a^{\frac{3}{4}}\, e^{-2\pi i \xi t}\, \hat{\psi}(a\,\xi_1, \sqrt{a}(\xi_2 - s\,\xi_1)) \\ &= a^{\frac{3}{4}}\, e^{-2\pi i \xi t}\, \hat{\psi}_1(a\,\xi_1)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1} - s)). \end{aligned}$$

Thus, each function $\hat{\psi}_{ast}$ is supported on the set:

$$\operatorname{supp} \hat{\psi}_{ast} \subset \{(\xi_1, \xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}],\, |\tfrac{\xi_2}{\xi_1} - s| \le \sqrt{a}\}.$$

As illustrated in Figure 1, each continuous shearlet ψ_{ast} has frequency support on a pair of trapezoids, symmetric with respect to the origin, oriented along a line of slope s. The support becomes increasingly thin as $a \to 0$.

When $S = \mathbb{R}$ and $I = \mathbb{R}^+$, by Proposition 3.1, the Continuous Shearlet Transform provides a reproducing formula (2.3) for all $f \in L^2(\mathbb{R}^2)$:

$$\|f\|^2 = \int_{\mathbb{R}^2} \int_{-\infty}^{\infty} \int_0^{\infty} |\mathcal{SH}_\psi f(a,s,t)|^2\, \frac{da}{a^3}\, ds\, dt.$$

On the other hand, if S, I are bounded sets, by Proposition 3.2, the Continuous Shearlet Transform provides a reproducing formula only for functions in a proper subspace of $L^2(\mathbb{R}^2)$. However, even when S, I are bounded, it is possible to obtain a reproducing formula for all $f \in L^2(\mathbb{R}^2)$ as follows. Let

$$\hat{\psi}^{(v)}(\xi) = \hat{\psi}^{(v)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2(\tfrac{\xi_1}{\xi_2}),$$

where $\hat{\psi}_1$, $\hat{\psi}_2$ are defined as in Definition 3.3, and let $\Lambda^{(v)} = \{(M,t) : M \in G^{(v)}, t \in \mathbb{R}^2\}$, where

$$G^{(v)} = \left\{ M = M_{as} = \begin{pmatrix} \sqrt{a} & 0 \\ -\sqrt{a}\, s & a \end{pmatrix}, \quad a \in I,\, s \in S \right\}. \tag{3.5}$$

Then, proceeding as above, it is easy to show that $\psi^{(v)}$ is a continuous shearlet for $L^2(C^{(v)})^\vee$ with respect to $\Lambda^{(v)}$, where $C^{(v)}$ is the vertical cone:

$$C^{(v)} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \ge 2 \text{ and } |\tfrac{\xi_2}{\xi_1}| > 1\}.$$

Accordingly, we introduce the shearlet system $\{\psi_{ast}^{(v)} = T_t\, D_M \psi^{(v)} : a \in I,\, s \in S,\, t \in \mathbb{R}^2\}$, for $(M,t) \in \Lambda^{(v)}$, and the associated Continuous Shearlet Transform $\mathcal{SH}_\psi^{(v)} f(a,s,t) = \langle f, \psi_{ast}^{(v)} \rangle$. Finally, let $W(x)$ be such that $\hat{W}(\xi) \in C^\infty(\mathbb{R}^2)$ and

$$|\hat{W}(\xi)|^2 + \chi_{C_1}(\xi) \int_0^1 |\hat{\psi}_1(a\xi_1)|^2\, \frac{da}{a} + \chi_{C_2}(\xi) \int_0^1 |\hat{\psi}_1(a\xi_2)|^2\, \frac{da}{a} = 1, \tag{3.6}$$

for a.e. $\xi \in \mathbb{R}^2$, where $C_1 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\frac{\xi_2}{\xi_1}| \leq 1\}$, $C_2 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\frac{\xi_2}{\xi_1}| > 1\}$. Then it follows that W is a C^∞-window function in $\mathbb{R}^2$ with $\hat{W}(\xi) = 1$ for $\xi \in [-1/2, 1/2]^2$, $\hat{W}(\xi) = 0$ outside the box $\{\xi \in [-2, 2]^2\}$. Finally, let $(P_{C_1} f)^\wedge = \hat{f}\,\chi_{C_1}$ and $(P_{C_2} f)^\wedge = \hat{f}\,\chi_{C_2}$. Then, for each $f \in L^2(\mathbb{R}^2)$ we have:

$$\begin{aligned} \|f\|^2 &= \int_{\mathbb{R}^2} |\langle f, T_t\, W\rangle|^2\, dt + \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 |\mathcal{SH}_{(P_{C_1} f)}(a,s,t)|^2\, \frac{da}{a^3}\, ds\, dt \\ &+ \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 |\mathcal{SH}^{(v)}_{(P_{C_2} f)}(a,s,t)|^2\, \frac{da}{a^3}\, ds\, dt. \end{aligned} \tag{3.7}$$

The proof of this equality is reported in the Appendix. Equation (3.7) shows that f is continuously reproduced by using isotropic window functions at coarse scales, and two sets of continuous shearlet systems at fine scales: one set corresponding to the horizontal cone C (in the frequency domain) and another set corresponding to the vertical cone $C^{(v)}$. The advantage of this construction, with respect to the simpler one where $S = \mathbb{R}$, is that in this case the set S associated with the shear variable is the closed interval $S = \{s : |s| \leq 2\}$. This property will be important in Subsection 4.4 and Section 5.

There are other choices of the subset Λ, given by (3.1), generating affine systems with properties similar to the continuous shearlet systems. Variants and generalizations of this construction will be discussed in Section 6.

3.2. **Localization of shearlets.** Since the continuous shearlets ψ we constructed in the previous subsection satisfy $\hat{\psi} \in C_0^\infty(\mathbb{R}^2)$, it follows that the analyzing elements of the associated continuous shearlet systems *decay rapidly* as $|x| \to \infty$; that is,

$$\psi_{ast}(x) = O(|x|^{-k}) \quad \text{as } |x| \to \infty, \quad \text{for every } k \geq 0.$$

More precisely, we have the following result.

Proposition 3.4. *Let $\psi \in L^2(\mathbb{R}^2)$ be a continuous shearlet satisfying $\hat{\psi} \in C_0^\infty(\mathbb{R}^2)$, and let M be defined as in* (3.2). *Then, for each $k \in \mathbb{N}$, there is a constant C_k such that, for any $x \in \mathbb{R}^2$, we have*

$$\begin{aligned} |\psi_{ast}(x)| &\leq C_k\, |\det M|^{-\frac{1}{2}}\, (1 + |M^{-1}(x-t)|^2)^{-k} \\ &= C_k\, a^{-\frac{3}{4}}\, (1 + a^{-2}(x_1 - t_1)^2 + 2a^{-2} s(x_1 - t_1)(x_2 - t_2) \\ &\quad + a^{-1}(1 + a^{-1}s^2)(x_2 - t_2)^2)^{-k}. \end{aligned}$$

In particular, $C_k = k\,\frac{15}{2}\left(\|\hat{\psi}\|_\infty + \|\triangle^k \hat{\psi}\|_\infty\right)$, where $\triangle = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2}$ is the frequency domain Laplacian operator.

Proof. Observe that, for $t = \binom{t_1}{t_2}$ and $x = \binom{x_1}{x_2}$ in $\mathbb{R}^2$, we have:

$$\psi_{ast}(x) = |\det M|^{-\frac{1}{2}}\, \psi(M^{-1}(x-t)) = a^{-\frac{3}{4}}\, \psi\begin{pmatrix} a^{-1}\,(x_1 - t_1) + s\, a^{-1}(x_2 - t_2) \\ a^{-\frac{1}{2}}(x_2 - t_2) \end{pmatrix}.$$

The proof then follows from Proposition 2.2, where

$$R = \{(\xi_1, \xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}],\ |s - \tfrac{\xi_2}{\xi_1}| \leq \sqrt{a}\}.$$

It is easy to check that $m(R) = \frac{15}{2}$. ∎

4. Analysis of singularities

As observed above, the continuous shearlet ψ, constructed in Section 3, satisfies $\hat{\psi} \in C_0^\infty(\mathbb{R}^2)$. It follows that $\psi \in \mathcal{S}(\mathbb{R}^2)$ and, therefore, the Continuous Shearlet Transform $\mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{ast} \rangle$, $a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2$, is well defined for all tempered distributions $f \in \mathcal{S}'$.

In the following, we will examine the behavior of the Continuous Shearlet Transform of several distributions containing different types of singularities. This will be useful to illustrate the basic properties of the shearlet transform, before stating a more general result in the next section. Indeed, the rate of decay of the Continuous Shearlet Transform exactly describes the location and orientation of the singularities. Interestingly, despite the different mathematical structure, the decay rates found for the Continuous Shearlet Transform are consistent with those found using the Continuous Curvelet Transform in [6].

In order to state our results, it will be useful to introduce the following notation to distinguish between the following two different behaviors of the Continuous Shearlet Transform.

Definition 4.1. Let f be a distribution on $\mathbb{R}^2$, $\mathcal{SH}_\psi f(a,s,t)$ be defined as in Definition 3.3, and let $r \in \mathbb{R}$. Then $\mathcal{SH}_\psi f(a,s,t)$ *decays rapidly as* $a \to 0$ if

$$\mathcal{SH}_\psi f(a,s,t) = O(a^k) \quad \text{as } a \to 0, \text{ for every } k \geq 0.$$

We use the notation: $\mathcal{SH}_\psi f(a,s,t) \sim a^r$ *as* $a \to 0$ if there exist constants $0 < \alpha \leq \beta < \infty$ such that

$$\alpha\, a^r \leq \mathcal{SH}_\psi f(a,s,t) \leq \beta\, a^r \quad \text{as } a \to 0.$$

4.1. **Point singularities.** We start by examining the decay properties of the Continuous Shearlet Transform of the Dirac δ.

Proposition 4.2. *If* $t = 0$*, we have*

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

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

Proof. For $t = 0$ we have

$$\langle \delta, \psi_{ast} \rangle = \psi_{as0}(0) = a^{-\frac{3}{4}}\, \psi(0) \sim a^{-\frac{3}{4}} \text{ as } a \to 0.$$

Next let $t \neq 0$. Then

$$\langle \delta, \psi_{ast} \rangle = \psi_{ast}(0),$$

and, by Proposition 3.4, for each $k \in \mathbb{N}$, we have

$$|\psi_{ast}(0)| \leq C_k\, a^{-\frac{3}{4}} (1 + a^{-2} t_1^2 + 2a^{-2} s t_1 t_2 + (1 + a^{-1} s^2) a^{-1} t_2^2)^{-k}.$$

Thus, if $t_2 \neq 0$, then $|\psi_{ast}(0)| = O(a^{k-3/4})$ as $a \to 0$. Otherwise, if $t_2 = 0$, $t_1 \neq 0$, then $|\psi_{ast}(0)| = O(a^{2k-3/4})$ as $a \to 0$. □

Next let us consider the point singularity $\sigma_\alpha(x) = |x|^\alpha$ for $-2 < \alpha < \infty$. The Continuous Shearlet Transform shows the following decay.

Proposition 4.3. *Let* $\mathcal{SH}_{\sigma_\alpha}(a,s,t)$ *be defined as in Definition* 3.3*. If* $t = 0$*, we have*

$$\mathcal{SH}_{\sigma_\alpha}(a,s,t) \sim a^{\frac{5}{4}+\alpha} \quad \text{as } a \to 0.$$

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

Proof. First observe that $\hat{\sigma}_\alpha(x) = C_\alpha \, |\xi|^{-2-\alpha}$. Using Fubini, we compute

$$\begin{aligned}
\langle \hat{\sigma}_\alpha, \hat{\psi}_{ast}\rangle &= C_\alpha \int_{\mathbb{R}^2} |\xi|^{-2-\alpha} \hat{\psi}_{ast}(\xi)\, d\xi \\
&= C_\alpha \, a^{\frac{3}{4}} \int_{\mathbb{R}^2} \hat{\psi}_1(a\xi_1)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1} - s))\, (\xi_1^2 + \xi_2^2)^{-1-\frac{\alpha}{2}} e^{-2\pi i \xi t}\, d\xi_1\, d\xi_2 \\
&= C_\alpha \, a^{-\frac{1}{4}} \int_{\mathbb{R}^2} \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\sqrt{a}\,\tfrac{\xi_2}{\xi_1} - s))(a^{-2}\xi_1^2 + \xi_2^2)^{-1-\frac{\alpha}{2}}\, e^{-2\pi i (a^{-1}\xi_1, \xi_2) t}\, d\xi_1\, d\xi_2 \\
&= C_\alpha \, a^{-\frac{3}{4}} \int_{\mathbb{R}^2} \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\xi_2)\, (a^{-2}\xi_1^2 + \xi_1^2 a^{-1}(\xi_2 + a^{-\frac{1}{2}} s)^2)^{-1-\alpha/2} \\
&\quad \cdot e^{-2\pi i a^{-1/2}(a^{-1/2}\xi_1, \xi_1(\xi_2 + a^{-1/2}s))t}\, \xi_1\, d\xi_1\, d\xi_2 \\
&= C_\alpha \, a^{\frac{1}{4}+\frac{\alpha}{2}} \int_{\mathbb{R}^2} (a^{-1} + (\xi_2 + a^{-\frac{1}{2}} s)^2)^{-1-\frac{\alpha}{2}} \\
&\quad \cdot e^{-2\pi i a^{-1}\xi_1(t_1 + s t_2)}\, e^{-2\pi i a^{-1/2}\xi_1\xi_2 t_2} d\xi_1\, d\xi_2.
\end{aligned}$$

Let $t = 0$. Then

$$\langle \hat{\sigma}_\alpha, \hat{\psi}_{ast}\rangle = C_\alpha \, a^{\frac{1}{4}+\frac{\alpha}{2}} \int_{\mathbb{R}} \frac{\hat{\psi}_1(\xi_1)}{\xi_1^{1+\alpha}} \int_{\mathbb{R}} \frac{\hat{\psi}_2(\xi_2)}{(a^{-1} + (\xi_2 + a^{-\frac{1}{2}} s)^2)^{1+\frac{\alpha}{2}}}\, d\xi_2\, d\xi_1.$$

For $a \ll 1$, we obtain

$$\begin{aligned}
\left| \int_{\mathbb{R}} \frac{\hat{\psi}_2(\xi_2)}{(a^{-1} + (\xi_2 + a^{-\frac{1}{2}} s)^2)^{1+\frac{\alpha}{2}}}\, d\xi_2 \right| &= a^{1+\frac{\alpha}{2}} \int_{\mathbb{R}} \left| \frac{\hat{\psi}_2(\xi_2)}{(1 + a(\xi_2 + a^{-\frac{1}{2}} s)^2)^{1+\frac{\alpha}{2}}} \right| d\xi_2 \\
&\sim a^{1+\frac{\alpha}{2}} \int_{\mathbb{R}} |\hat{\psi}_2(\xi_2)|\, d\xi_2. \qquad (4.1)
\end{aligned}$$

Thus,

$$\left| \langle \hat{\sigma}_\alpha, \hat{\psi}_{ast}\rangle \right| \le C_\alpha \, a^{\frac{5}{4}+\alpha} \int_{\mathbb{R}} \left| \frac{\hat{\psi}_1(\xi_1)}{\xi_1^{1+\alpha}} \right| d\xi_1 \int_{\mathbb{R}} \left| \hat{\psi}_2(\xi_2) \right| d\xi_2 \sim a^{\frac{5}{4}+\alpha} \text{ as } a \to 0.$$

Next let $t \neq 0$. If $t_1 + s\, t_2 \neq 0$, using again (4.1), we observe that

$$\begin{aligned}
&\left| \langle \hat{\sigma}_\alpha, \hat{\psi}_{ast}\rangle \right| \\
&\le C_\alpha \, a^{\frac{1}{4}+\frac{\alpha}{2}} \left| \int_{\mathbb{R}} \hat{\psi}_1(\xi_1)\, \xi_1^{-1-\alpha}\, e^{-2\pi i a^{-1}\xi_1(t_1 + s t_2)} \int_{\mathbb{R}} \frac{\hat{\psi}_2(\xi_2)\, e^{-2\pi i a^{-1/2}\xi_1\xi_2 t_2}}{(a^{-1} + (\xi_2 + a^{-\frac{1}{2}} s)^2)^{1+\frac{\alpha}{2}}}\, d\xi_2\, d\xi_1 \right| \\
&\qquad \le C'_\alpha \, a^{\frac{5}{4}+\alpha} \, |\widetilde{\psi}_1(\tfrac{t_1 + s t_2}{a})| \int_{\mathbb{R}} \left| \hat{\psi}_2(\xi_2) \right| d\xi_2,
\end{aligned}$$

where $\widetilde{\psi}_1(u) = \int_{\mathbb{R}} \xi_1^{-1-\alpha}\, \hat{\psi}_1(\xi_1)\, e^{-2\pi i u \xi_1}\, d\xi_1$ is a band-limited function decaying rapidly for $|u| \to \infty$ (since ψ_1 is band-limited and C^∞, its behavior is similar when $\hat{\psi}_1(u)$ is divided by $u^{1+\alpha}$). If $t_1 + s\, t_2 = 0$, then one uses a similar estimate employing the exponential function $e^{-2\pi i a^{-1/2}\xi_1\xi_2 t_2}$. □

Observe that, for $\alpha = -2$, which corresponds to the Dirac delta, we have the same rate of convergence $a^{-3/4}$ as computed above.

4.2. **Linear singularities.** Next we will consider the linear delta distribution $\nu_p(x_1, x_2) = \delta(x_1 + p\,x_2)$, $p \in \mathbb{R}$, defined by

$$\langle \nu_p, f \rangle = \int_{\mathbb{R}} f(-p\,x_2, x_2)\,dx_2.$$

The following result shows that the Continuous Shearlet Transform precisely determines both the position and the orientation of the linear singularity, in the sense that the transform $\mathcal{SH}_\psi \nu_p(a,s,t)$ always decays rapidly as $a \to 0$ *except* when t is on the singularity and $s = p$, i.e., the direction perpendicular to the singularity or, in other words, in which the singularity occurs.

Proposition 4.4. *If $t_1 = -p\,t_2$ and $s = p$, we have*

$$\mathcal{SH}_\psi \nu_p(a,s,t) \sim a^{-\frac{1}{4}} \quad \text{as } a \to 0.$$

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

Proof. The following heuristic argument gives

$$\begin{aligned}\hat{\nu}_p(\xi_1,\xi_2) &= \int\int \delta(x_1 + p\,x_2)\,e^{-2\pi i \xi x}\,dx_2\,dx_1 \\ &= \int e^{-2\pi i x_2(\xi_2 - p\,\xi_1)}\,dx_2 = \delta(\xi_2 - p\,\xi_1) = \nu_{(-\frac{1}{p})}(\xi_1,\xi_2).\end{aligned}$$

That is, the Fourier transform of the linear delta on $\mathbb{R}^2$ is another linear delta on $\mathbb{R}^2$, where the slope $-\frac{1}{p}$ is replaced by the slope p. A direct computation gives:

$$\begin{aligned}\langle \hat{\nu}_p\,, \hat{\psi}_{ast} \rangle &= \int_{\mathbb{R}} \overline{\hat{\psi}_{ast}(\xi_1, p\xi_1)}\,d\xi_1 \\ &= a^{\frac{3}{4}} \int_{\mathbb{R}} \hat{\psi}(a\xi_1, \sqrt{a}p\xi_1 - \sqrt{a}s\xi_1)\,e^{2\pi i \xi_1(t_1 + pt_2)}\,d\xi_1 \\ &= a^{-\frac{1}{4}} \int_{\mathbb{R}} \hat{\psi}(\xi_1, a^{-\frac{1}{2}}p\xi_1 - a^{-\frac{1}{2}}s\xi_1)\,e^{2\pi i a^{-1}\xi_1(t_1+pt_2)}\,d\xi_1 \\ &= a^{-\frac{1}{4}} \int_{\mathbb{R}} \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(a^{-\frac{1}{2}}(p-s))\,e^{2\pi i a^{-1}\xi_1(t_1+pt_2)}\,d\xi_1 \\ &= a^{-\frac{1}{4}}\,\hat{\psi}_2(a^{-\frac{1}{2}}(p-s))\,\psi_1(a^{-1}(t_1+pt_2)).\end{aligned}$$

If $s \neq p$, then there exists some $a > 0$ such that $|p - s| > \sqrt{a}$. This implies that $\hat{\psi}_2(a^{-1/2}(p-s)) = 0$, and so $\langle \hat{\nu}_p\,, \hat{\psi}_{ast} \rangle = 0$. On the other hand, if $t_1 = -p\,t_2$ and $s = p$, then $\hat{\psi}_2(a^{-1/2}(p-s)) = \hat{\psi}_2(0) \neq 0$, and

$$\langle \hat{\nu}_p\,, \hat{\psi}_{ast} \rangle = a^{-\frac{1}{4}}\,\hat{\psi}_2(a^{-\frac{1}{2}}(p-s))\,\psi_1(0) \sim a^{-\frac{1}{4}} \quad \text{as } a \to 0.$$

If $t_1 \neq -p\,t_2$, by Proposition 2.2, we observe that, for all $k \in \mathbb{N}$,

$$\begin{aligned}\langle \hat{\nu}_p\,, \hat{\psi}_{ast} \rangle &\leq a^{-\frac{1}{4}}\,\hat{\psi}_2(a^{-\frac{1}{2}}(p-s))\,|\psi_1(a^{-1}(t_1+pt_2))| \\ &\leq C_k\,a^{-\frac{1}{4}}\,\hat{\psi}_2(a^{-\frac{1}{2}}(p-s))\,(1 + a^{-2}(t_1+pt_2)^2)^{-k} = O(a^{2k-\frac{1}{4}}) \text{ as } a \to 0.\end{aligned}$$

□

Next, let us consider the two-dimensional Heaviside function $H(x_1,x_2) = \chi_{\{x_1 \geq 0\}}(x_1,x_2)$.

Proposition 4.5. *If $t_1 = 0$ and $|s| \le \sqrt{a}$, we have*

$$\mathcal{SH}_H(a,s,t) \sim a^{\frac{3}{4}} \quad \text{as } a \to 0.$$

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

Proof. We will use the relation between H and the derivative of the delta distribution. Since $\nu_0(x_1,x_2) = \delta(x_1) = \frac{\partial}{\partial x_1} H(x_1,x_2)$, then $\hat{H}(\xi_1,\xi_2) = -i\frac{1}{\xi_1}\hat{\nu}(\xi_1,\xi_2)$. It follows that:

$$\begin{aligned}
\langle \hat{H}, \hat{\psi}_{ast}\rangle &= i\int_{\mathbb{R}} \frac{1}{\xi_1}\,\hat{\psi}_{ast}(\xi_1,0)\,d\xi_1 \\
&= i a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{1}{\xi_1}\,\hat{\psi}(a\xi_1, -\sqrt{a}s\xi_1) e^{-2\pi i t_1\xi_1}\,d\xi_1 \\
&= i a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{1}{\xi_1}\,\hat{\psi}(\xi_1, -a^{-\frac{1}{2}}s\xi_1)\, e^{-2\pi i t_1 a^{-1}\xi_1}\,d\xi_1 \\
&= i a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{1}{\xi_1}\,\hat{\psi}_1(\xi_1)\,\hat{\psi}_2(-a^{-\frac{1}{2}}s)\, e^{-2\pi i t_1 a^{-1}\xi_1}\,d\xi_1 \\
&= i\,\hat{\psi}_2(-a^{-\frac{1}{2}}s)\, a^{\frac{3}{4}} \int_{\mathbb{R}} \frac{1}{\xi_1}\,\hat{\psi}_1(\xi_1)\, e^{-2\pi i t_1 a^{-1}\xi_1}\,d\xi_1.
\end{aligned}$$

Similarly to the proof of Proposition 4.4, if $|s| > \sqrt{a}$, then $\langle \hat{H}, \hat{\psi}_{ast}\rangle = 0$, and thus, $\mathcal{SH}_H(a,s,t)$ decays rapidly as $a \to 0$. On the other hand, if $t_1 = 0$ and $|s| \le \sqrt{a}$, we obtain

$$|\langle \hat{H}, \hat{\psi}_{ast}\rangle| \le \max_{\xi_2} |\hat{\psi}_2(\xi_2)|\, a^{\frac{3}{4}} \int_{\mathbb{R}} |\tfrac{1}{\xi_1}\,\hat{\psi}_1(\xi_1)|\,d\xi_1 \sim a^{\frac{3}{4}} \text{ as } a \to 0.$$

Finally, in case $t_1 \ne 0$,

$$|\langle \hat{H}, \hat{\psi}_{ast}\rangle| \le \max_{\xi_2} |\hat{\psi}_2(\xi_2)|\, a^{\frac{3}{4}}\, \widetilde{\psi}(\tfrac{t_1}{a}),$$

where $\widetilde{\psi}(u) = \int_{\mathbb{R}} \frac{1}{\xi_1}\,\hat{\psi}_1(\xi_1)\, e^{-2\pi i u\xi_1}\,d\xi_1$ is a band-limited function decaying rapidly as $a \to 0$. □

4.3. **Polygonal singularities.** Here we consider the characteristic function χ_V of the cone $V = \{(x_1,x_2) : x_1 \ge 0,\ qx_1 \le x_2 \le px_1\}$, where $0 < q \le p < \infty$. We have the following result.

Proposition 4.6. *For $t = 0$, if $s = -\frac{1}{p}$ or $s = -\frac{1}{q}$, we have*

$$\mathcal{SH}_\psi \chi_V(a,s,t) \sim a^{\frac{3}{4}} \quad \text{as } a \to 0,$$

and if $s \ne -\frac{1}{p}$ and $s \ne -\frac{1}{q}$, we have

$$\mathcal{SH}_\psi \chi_V(a,s,t) \sim a^{\frac{5}{4}} \quad \text{as } a \to 0.$$

For $t \ne 0$, if $s = -\frac{1}{p}$ or $s = -\frac{1}{q}$, we have

$$\mathcal{SH}_\psi \chi_V(a,s,t) \sim a^{\frac{3}{4}} \quad \text{as } a \to 0.$$

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

The decay of the Continuous Shearlet Transform of χ_V is illustrated in Figure 2. As shown in the figure, the decay of $\mathcal{SH}_\psi \chi_V(a,s,t)$ exactly identifies the location and orientation of the singularities. It is interesting to notice that the orientation of

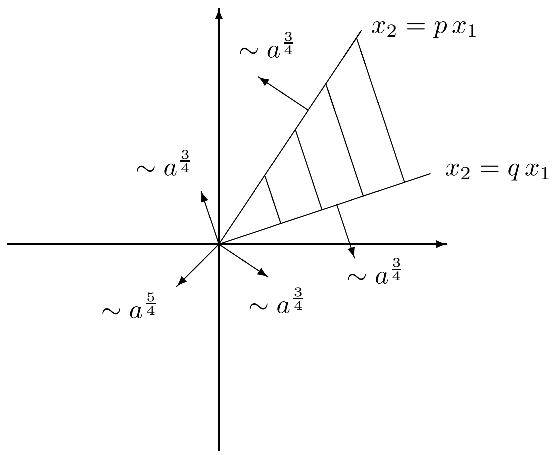

FIGURE 2. Decay properties of the Continuous Shearlet Transform $\mathcal{SH}_\psi \chi_V(a,s,t)$

the linear singularities can even be detected considering only the "point singularity" at the origin.

Proof of Proposition 4.6. The Fourier transform of χ_V can be computed to be

$$\hat{\chi}_V(\xi_1,\xi_2) = C\frac{1}{(\xi_1+q\xi_2)(\xi_1+p\xi_2)}, \quad \text{where } C = \frac{(p+q)^2}{(2\pi)^2}.$$

A direct computation gives:

$$\begin{aligned}
&\langle \hat{\chi}_V, \hat{\psi}_{ast}\rangle \\
&= Ca^{\frac{3}{4}} \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{(\xi_1+q\xi_2)(\xi_1+p\xi_2)}\,\hat{\psi}_1(a\xi_1)\,\hat{\psi}_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1}-s))\,e^{2\pi i\xi t}\,d\xi_1\,d\xi_2 \\
&= Ca^{-\frac{1}{4}} \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{(a^{-1}\xi_1+q\xi_2)(a^{-1}\xi_1+p\xi_2)}\,\hat{\psi}_1(\xi_1)\,\hat{\psi}_2(a^{-\frac{1}{2}}(a\tfrac{\xi_2}{\xi_1}-s)) \\
&\quad \cdot e^{2\pi i(a^{-1}\xi_1,\xi_2)t}\,d\xi_1\,d\xi_2 \\
&= Ca^{-\frac{3}{4}} \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{\xi_1}{(a^{-1}\xi_1+q\xi_1(a^{-1/2}\xi_2+a^{-1}s))(a^{-1}\xi_1+p\xi_1(a^{-1/2}\xi_2+a^{-1}s))} \\
&\quad \cdot\hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\xi_2)\,e^{2\pi i\xi_1(a^{-1}(t_1+st_2)+a^{-1/2}\xi_2t_2)}\,d\xi_1\,d\xi_2 \\
&= Ca^{\frac{1}{4}} \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{\xi_1}{(a^{-1/2}\xi_1(1+sq)+q\xi_1\xi_2)(a^{-1/2}\xi_1(1+sp)+p\xi_1\xi_2)} \\
&\quad \cdot\hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\xi_2)\,e^{2\pi i\xi_1(a^{-1}(t_1+st_2)+a^{-1/2}\xi_2t_2)}\,d\xi_1\,d\xi_2.
\end{aligned}$$

Let us first consider the case $t=0$. By the previous computation we can rewrite $\langle \hat{\chi}_V, \hat{\psi}_{as0}\rangle$ as

$$Ca^{\frac{1}{4}} \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{\xi_1}{(a^{-1/2}\xi_1(1+sq)+q\xi_1\xi_2)(a^{-1/2}\xi_1(1+sp)+p\xi_1\xi_2)}\,\hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\xi_2)\,d\xi_1\,d\xi_2.$$

If $s \neq -\frac{1}{p}$ and $s \neq -\frac{1}{q}$, for $a \ll 1$ we can rewrite $\langle \hat{\chi}_V, \hat{\psi}_{as0} \rangle$ as

$$\begin{aligned} &Ca^{\frac{5}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\xi_1}{(\xi_1(1+sq)+a^{1/2}q\xi_1\xi_2)(\xi_1(1+sp)+a^{1/2}p\xi_1\xi_2)} \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\xi_2)\, d\xi_1\, d\xi_2 \\ &\quad \sim\ C' a^{\frac{5}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\xi_1}{(\xi_1(1+sq))(\xi_1(1+sp))} \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\xi_2)\, d\xi_1\, d\xi_2; \end{aligned}$$

hence

$$\langle \hat{\chi}_V, \hat{\psi}_{as0} \rangle \sim a^{\frac{5}{4}} \quad \text{as } a \to 0.$$

The above computation also shows that if $t = 0$ and $s = -\frac{1}{p}$ or $s = -\frac{1}{q}$, we have

$$\langle \hat{\chi}_V, \hat{\psi}_{as0} \rangle \sim a^{\frac{3}{4}} \quad \text{as } a \to 0.$$

Next, let us consider the situation, where t lies on one singularity but $t \neq 0$, i.e., $t_2 = pt_1$ or $t_2 = qt_1$. Here we will only examine the first case. The second one can be treated similarly. First let $s = -\frac{1}{p}$, i.e., s is perpendicular to the linear boundary of the cone $x_2 = px_1$. For $a \ll 1$ we have

$$\begin{aligned} \langle \hat{\chi}_V, \hat{\psi}_{ast} \rangle &= Ca^{\frac{1}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\xi_1}{(a^{-1/2}\xi_1(1-q/p)+q\xi_1\xi_2)p\xi_1\xi_2} \hat{\psi}_1(\xi_1)\hat{\psi}_2(\xi_2) \\ &\quad \cdot e^{2\pi i a^{-1/2} p t_1 \xi_1 \xi_2} d\xi_1 d\xi_2 \\ &= Ca^{\frac{3}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\xi_1}{(\xi_1(1-q/p)+a^{1/2}q\xi_1\xi_2)p\xi_1\xi_2} \\ &\quad \cdot \hat{\psi}_1(\xi_1)\hat{\psi}_2(\xi_2) e^{2\pi i a^{-1/2} p t_1 \xi_1 \xi_2} d\xi_1 d\xi_2 \\ &\sim a^{\frac{3}{4}} \quad \text{as } a \to 0. \end{aligned}$$

Secondly, let $s \neq -\frac{1}{p}$. We have:

$$\begin{aligned} &\langle \hat{\chi}_V, \hat{\psi}_{ast} \rangle \\ &\quad = Ca^{\frac{1}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\xi_1}{(a^{-1/2}\xi_1(1+sq)+q\xi_1\xi_2)(a^{-1/2}\xi_1(1+sp)+a^{-1/2}p\xi_1\xi_2)} \\ &\qquad \cdot \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\xi_2)\, e^{2\pi i \xi_1 t_1 (a^{-1}(1+sp)+a^{-1/2}p\xi_2)}\, d\xi_1\, d\xi_2 \\ &\quad = Ca^{\frac{1}{4}} \int_{\mathbb{R}} \varphi(\xi_1)\, \hat{\psi}_1(\xi_1)\, e^{2\pi i a^{-1} t_1 (1+sp)\xi_1}\, d\xi_1, \end{aligned}$$

where

$$\begin{aligned} \varphi(\xi_1) &= \int_{\mathbb{R}} \frac{\xi_1}{(a^{-1/2}\xi_1(1+sq)+q\xi_1\xi_2)(a^{-1/2}\xi_1(1+sp)+p\xi_1\xi_2)} \hat{\psi}_2(\xi_2) \\ &\quad \cdot e^{2\pi i a^{-1/2} t_1 p \xi_1 \xi_2} d\xi_2. \end{aligned}$$

Since ψ_1 and ψ_2 are band-limited, the function φ has compact support; hence $(\varphi\hat{\psi}_1)^\vee$ is of rapid decay towards infinity. Thus

$$\langle \hat{\chi}_V, \hat{\psi}_{ast} \rangle = C\, a^{\frac{1}{4}} (\varphi\hat{\psi}_1)^\vee (a^{-1}t_1(1-sp)) = O(a^k) \quad \text{as } a \to 0.$$

Finally, in case $t_2 \neq pt_1$, $t_2 \neq qt_1$ and $t_1 \neq 0$, a similar argument to the one above shows that $\langle \hat{\chi}_V, \hat{\psi}_{ast} \rangle$ decays rapidly also in this case. □

4.4. **Curvilinear singularities.** We will now examine the behavior of the Continuous Shearlet Transform of a distribution having a discontinuity along a curve. Throughout this section we will assume that the shearlet $\hat{\psi}(\xi_1,\xi_2)=\hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\frac{\xi_1}{\xi_2})$ satisfies the following additional assumptions: $\hat{\psi}_1$ is odd, $\hat{\psi}_2$ is even and strictly decreasing. This assumption will be used in the proof of Proposition 4.7.

Let $B(x_1,x_2)=\chi_D(x_1,x_2)$, where $D=\{(x_1,x_2)\in\mathbb{R}^2 : x_1^2+x_2^2\le 1\}$. We have the following:

Proposition 4.7. *If* $t_1^2+t_2^2=1$ *and* $s=\frac{t_2}{t_1}$, $t_1\neq 0$, *we have*

$$\mathcal{SH}_\psi B(a,s,t)\sim a^{\frac{3}{4}} \quad \text{as } a\to 0.$$

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

The assumption $t_1\neq 0$ shows that the shearlet transform $\mathcal{SH}_\psi B(a,s,t)$ is unable to handle the vertical direction $s\to\infty$. To provide a complete analysis of the singularities of B, we need to use both $\mathcal{SH}_\psi B(a,s,t)$ and $\mathcal{SH}_\psi B^{(v)}(a,s,t)$ (as defined in Section 3). Since the shearlets $\psi_{ast}^{(v)}$ are defined on the vertical cone $C^{(v)}$, using $\mathcal{SH}_\psi B^{(v)}(a,s,t)$ one can obtain a similar result to Proposition 4.7, for $s=\frac{t_1}{t_2}$, $t_2\neq 0$. Since the argument for both cases is exactly the same, we will only examine the transform $\mathcal{SH}_\psi B(a,s,t)$.

In order to prove Proposition 4.7, we need to recall the following facts. First, we recall the asymptotic behavior of Bessel functions, that is given by the following lemma (cf. [28]):

Lemma 4.8. *There exists a constant* C_0 *such that*

$$J_1(2\pi\lambda)\sim C_0\,\lambda^{-\frac{1}{2}}(e^{2\pi i\lambda}+e^{-2\pi i\lambda}) \quad \text{as } \lambda\to\infty,$$

and, for $N=1,2,\dots,$ *there are constants* C_N *satisfying*

$$\Big(\frac{d}{d\lambda}\Big)^N J_1(2\pi\lambda)\sim C_N\,\lambda^{-\frac{1}{2}}\,(e^{2\pi i\lambda}\pm_N e^{-2\pi i\lambda}) \quad \text{as } \lambda\to\infty,$$

where the sign in $\pm_N$ *depends on* N *and* J_1 *is the Bessel function of order* 1.

Secondly, we recall the following fact concerning oscillatory integrals of the First Kind, that can be found in [28, Ch.8]:

Lemma 4.9. *Let* $A\in C_0^\infty(\mathbb{R})$ *and* $\Phi\in C^1(\mathbb{R})$, *with* $\Phi'(t)\neq 0$ *on* $\operatorname{supp} A$. *Then*

$$I(\lambda)=\int_{\mathbb{R}}A(t)\,e^{2\pi i\lambda\Phi(t)}dt=\frac{(-1)^N}{(2\pi i\lambda)^N}\int_{\mathbb{R}}D^N\big(A(t)\big)\,e^{2\pi i\lambda\Phi(t)}dt,$$

for $N=1,2,\dots,$ *where* $D\big(A(t)\big)=\frac{d}{dt}\big(\frac{A(t)}{\Phi'(t)}\big)$.

We can now prove Proposition 4.7.

Proof of Proposition 4.7. The Continuous Shearlet Transform of $B(x)$ is given by:
(4.2)

$$\mathcal{SH}_\psi B(a,s,t)=\langle B,\psi_{ast}\rangle=a^{\frac{3}{4}}\int_{\mathbb{R}}\int_{\mathbb{R}}\hat{\psi}_1(a\xi_1)\,\hat{\psi}_2(a^{-\frac{1}{2}}(\tfrac{\xi_2}{\xi_1}-s))\,e^{2\pi i\xi t}\,\hat{B}(\xi)\,d\xi_1\,d\xi_2.$$

The Fourier transform $\hat{B}(\xi_1,\xi_2)$ is the radial function:

$$\hat{B}(\xi_1,\xi_2)=2\int_{-1}^{1}\sqrt{1-x^2}\,e^{2\pi ix\sqrt{\xi_1^2+\xi_2^2}}\,dx=|\xi|^{-1}\,J_1(2\pi|\xi|),$$

where J_1 is the Bessel function of order 1. Therefore, the asymptotic behavior of $\hat{B}(\lambda)$ follows from Lemma 4.8, with the factor $\lambda^{-1/2}$ replaced by $\lambda^{-3/2}$.

Because of the radial symmetry, it is convenient to convert (4.2) into polar coordinates:

$$\begin{aligned} &\mathcal{SH}_\psi B(a,s,t) \\ &= a^{\frac{3}{4}} \iint \hat{\psi}_1(a\rho\cos\theta)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tan\theta - s))\, e^{2\pi i\rho(t_1\cos\theta + t_2\sin\theta)}\, \hat{B}(\rho)\, \rho\, d\rho\, d\theta \\ &= a^{-\frac{5}{4}} \iint \hat{\psi}_1(\rho\cos\theta)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tan\theta - s))\, e^{2\pi i\frac{\rho}{a}(t_1\cos\theta + t_2\sin\theta)}\, \hat{B}(\tfrac{\rho}{a})\, \rho\, d\rho\, d\theta. \end{aligned} \tag{4.3}$$

We will now examine the asymptotic decay of the function $\mathcal{SH}_\psi B(a,s,t)$ along the curve ∂B for $a \to 0$. Thus, we set $t_1^2 + t_2^2 = 1$ and, without loss of generality, assume $a < 1$. As we will show, the decay depends on whether the direction associated with s is normal to the curve ∂B or not.

Let us begin by considering the non-normal case $s \neq t_2/t_1$. From (4.3), we have:

$$\mathcal{SH}_\psi B(a,s,t) = a^{-\frac{5}{4}} \int I(a,\rho)\, \hat{B}(\tfrac{\rho}{a})\, \rho\, d\rho,$$

where (using the conditions on the support of $\hat{\psi}_2$)

$$I(a,\rho) = \int_{|\tan\theta - s| < \sqrt{a}} \hat{\psi}_1(\rho\cos\theta)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tan\theta - s))\, e^{2\pi i\frac{\rho}{a}(t_1\cos\theta + t_2\sin\theta)}\, d\theta.$$

Observe that the domain of integration is the cone $|\tan\theta - s| < \sqrt{a}$ about the direction $\tan\theta = s$, with $a < 1$. This implies that θ ranges over an interval. Since the conditions on the support of $\hat{\psi}_1$ imply that $|\rho\cos\theta| \subset [\frac{1}{2}, 2]$, it follows that ρ also ranges over an interval and, as a consequence, $I(a,\rho)$ is compactly supported in ρ.

We will show that $I(a,\rho)$ is an oscillatory integral of the First Kind that decays rapidly for $a \to 0$ for each ρ. To show that this is the case, we will apply Lemma 4.9 to $I(a,\rho)$, where $A(\theta;\rho) = \hat{\psi}_1(\rho\cos\theta)\, \hat{\psi}_2(a^{-1/2}(\tan\theta - s))$, $\Phi(\theta;\rho) = \rho(t_1\cos\theta + t_2\sin\theta)$ and $\lambda = a^{-1}$ and ρ is a fixed parameter. Observe that $\Phi'(\theta;\rho) = \rho(-t_1\sin\theta + t_2\cos\theta)$ and $\Phi'(\theta;\rho) \neq 0$ for $\tan\theta \neq \frac{t_2}{t_1}$. Thus, provided $|s - \frac{t_2}{t_1}| \geq \sqrt{a}$, it follows that $\Phi'(\theta;\rho) \neq 0$ on $\operatorname{supp} A$. A direct computation gives

$$\begin{aligned} D\big(A(\theta;\rho)\big) &= \frac{\partial}{\partial\theta} \frac{\hat{\psi}_1(\rho\cos\theta)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tan\theta - s))}{\rho(-t_1\sin\theta + t_2\cos\theta)} \\ &= \frac{\sin\theta}{t_1\sin\theta - t_2\cos\theta} \hat{\psi}_1'(\rho\cos\theta)\, \hat{\psi}_2(a^{-\frac{1}{2}}(\tan\theta - s)) \\ &\quad + a^{-\frac{1}{2}} \frac{\sec^2\theta}{\rho(t_2\cos\theta - t_1\sin\theta)} \hat{\psi}_1(\rho\cos\theta)\, \hat{\psi}_2'(a^{-\frac{1}{2}}(\tan\theta - s)) \\ &\quad + \frac{t_2\sin\theta + t_1\cos\theta}{\rho^2(t_2\cos\theta - t_1\sin\theta)^2}\, \hat{\psi}_1(\rho\cos\theta)\hat{\psi}_2(a^{-\frac{1}{2}}(\tan\theta - s)). \end{aligned}$$

Thus, since $\tan\theta \neq \frac{t_2}{t_1}$, using the assumptions on $\hat{\psi}_1, \hat{\psi}_2$, we obtain

$$\big|D\big(A(\theta;\rho)\big)\big| < a^{-\frac{1}{2}}\, C(\theta,\rho)\, \big(\|\hat{\psi}_1'\hat{\psi}_2\|_\infty + \|\hat{\psi}_1\hat{\psi}_2'\|_\infty + \|\hat{\psi}_1\hat{\psi}_2\|_\infty\big).$$

As observed above, the assumptions on the support of $\hat{\psi}_1, \hat{\psi}_2$ imply that $D\big(A(\theta;\rho)\big)$ is compactly supported in ρ away from $\rho = 0$. Using this observation and $\Phi'(\theta) \neq 0$, it follows that

$$\|D(A)\|_\infty < C\, a^{-\frac{1}{2}}.$$

Applying the same estimate repeatedly, we have that for each $N \in \mathbb{N}$,

$$\|D^N(A)\|_\infty < C_N\, a^{-\frac{N}{2}}.$$

Thus, using Lemma 4.9 with $\lambda = a^{-1}$, we conclude that for each $N \in \mathbb{N}$ there is a constant $C_N > 0$ such that

$$\sup_\rho |I(a,\rho)| < C_N\, a^{\frac{N}{2}}.$$

This implies that, under the assumption that we made for $t = (t_1, t_2)$ and s, the function $\mathcal{SH}_\psi B(a,s,t)$ decays rapidly for $a \to 0$.

Let us now consider the function $|\langle \hat{B}, \hat{\psi}_{ast}\rangle|$, where $t_1^2 + t_2^2 = 1$ and $s = t_2/t_1$ (corresponding to the direction normal to ∂B). For simplicity, let $(t_1, t_2) = (1, 0)$. The general case follows using a similar argument. From (4.3), using the change of variables $u = a^{-1/2}\sin\theta$, we obtain

$$\langle \hat{B}, \hat{\psi}_{a0(1,0)}\rangle = a^{-\frac{3}{4}} \int \hat{B}(\tfrac{\rho}{a})\, \eta_a(\rho)\; e^{2\pi i \frac{\rho}{a}}\, \rho\, d\rho, \tag{4.4}$$

where

$$\eta_a(\rho) = \int_{-(1+a)^{-1/2}}^{(1+a)^{-1/2}} \hat{\psi}_1(\rho\sqrt{1-au^2})\, \hat{\psi}_2(\frac{u}{\sqrt{1-au^2}})\, e^{2\pi i \frac{\rho}{a}(\sqrt{1-au^2}-1)}\, \frac{du}{\sqrt{1-au^2}}.$$

The assumptions on the support of $\hat{\psi}_2$ and $\hat{\psi}_1$ imply that $|u| < (1+a)^{1/2}$ and that $|\rho\sqrt{1-au^2}| \subset [\frac{1}{2}, 2]$, respectively. Thus, ρ ranges over a closed interval and, as a consequence, the functions $\eta_a(\rho)$ are compactly supported. For $0 < a < 1$, the functions

$$h_a(u) = \hat{\psi}_1(\rho\sqrt{1-au^2})\, \hat{\psi}_2(\frac{u}{\sqrt{1-au^2}})\, \frac{e^{2\pi i \frac{\rho}{a}(\sqrt{1-au^2}-1)}}{\sqrt{1-au^2}}$$

are equicontinuous and they converge uniformly:

$$\lim_{a\to 0} h_a(u) = h_0(u) = \hat{\psi}_1(\rho)\, \hat{\psi}_2(u)\, e^{-\pi i \rho u^2}.$$

Thus, we have the uniform limit:

$$\lim_{a\to 0} \eta_a(\rho) = \eta_0(\rho) = \int_{-1}^{1} \hat{\psi}_1(\rho)\, \hat{\psi}_2(u)\, e^{-\pi i \rho u^2}\, du,$$

and the same convergence holds for all u-derivatives. In particular, $\|\eta_a\|_\infty < C$, for all $a < 1$.

Using the asymptotic estimate given by Lemma 4.8 into (4.4), for a small, we have:

$$\begin{aligned} |\langle \hat{B}, \hat{\psi}_{a0(1,0)}\rangle| \;&\sim\; C\, a^{-\frac{3}{4}} \Big(\int \Big(\frac{a}{\rho}\Big)^{\frac{3}{2}} \eta_a(\rho)\, e^{4\pi i \frac{\rho}{a}}\, \rho\, d\rho + \int \Big(\frac{a}{\rho}\Big)^{\frac{3}{2}} \eta_a(\rho)\, \rho\, d\rho \Big) \\ &=\; C\, a^{\frac{3}{4}} \Big(\hat{F}_a\Big(-\frac{2}{a}\Big) + \int F_a(\rho)\, d\rho \Big), \end{aligned}$$

where $F_a(\rho) = \eta_a(\rho)\, \rho^{-1/2}$. The family of functions $\{F_a : 0 < a < 1\}$ has all its ρ derivatives bounded uniformly in a, and so $\hat{F}_a(-\frac{2}{a})$ decays rapidly as $a \to 0$. On the other hand, $\int F_a(\rho)\, d\rho$ tends to $\int \eta_0(\rho)\, \rho^{-1/2}\, d\rho$ as $a \to 0$.

One can show that there is a constant $C > 0$ such that $\int \eta_0(\rho)\,\rho^{-1/2}\,d\rho > C$. To do that, and ensure that the integral does not vanish, the functions ψ_1 and ψ_2 have to be chosen appropriately. This can be done, for example, by choosing $\hat{\psi}_1$ odd and $\hat{\psi}_2$ even and decreasing. The proof of this fact is omitted.

Using these observations, we conclude that

$$|\langle \hat{B}, \hat{\psi}_{a0(1,0)}\rangle| \sim a^{\frac{3}{4}} \quad \text{as } a \to 0.$$

Finally, if t is not on ∂B, then one can show that $\mathcal{SH}_\psi B(a,s,t)$ has rapid decay. This follows from the general analysis given in Section 5. □

5. Characterization of the wavefront set using the Shearlet Transform

The examples described in Section 4 suggest that the set of singularities of a distribution on $\mathbb{R}^2$ can be characterized using the Continuous Shearlet Transform. In this section, we will show that this is indeed the case. In order to do this, it will be useful to introduce the notions of singular support and wavefront set.

For a distribution u, we say that $x \in \mathbb{R}^2$ is a *regular point* of u if there exists some $\phi \in C_0^\infty(U_x)$, where U_x is a neighborhood of x and $\phi(x) \neq 0$, such that $\phi\, u \in C_0^\infty(\mathbb{R}^n)$. Recall that the condition $\phi\, u \in C_0^\infty$ is equivalent to $(\phi\, u)^\wedge$ being rapidly decreasing. The complement of the regular points of u is called the *singular support* of u and is denoted by $\operatorname{sing\,supp}(u)$. It is easy to see that the singular support of u is a closed subset of $\operatorname{supp}(u)$.

The wavefront set of u consists of certain $(x,\lambda) \in \mathbb{R}^2 \times \mathbb{R}$, with $x \in \operatorname{sing\,supp}(u)$. For a distribution u, a point $(x,\lambda) \in \mathbb{R}^2 \times \mathbb{R}$ is a *regular directed point* for u if there are neighborhoods U_x of x and V_λ of λ, and a function $\phi \in C_0^\infty(\mathbb{R}^2)$, with $\phi = 1$ on U_x, so that, for each $N > 0$, there is a constant C_N with

$$|(u\,\phi)^\wedge(\eta)| \le C_N\,(1 + |\eta|)^{-N},$$

for all $\eta = (\eta_1, \eta_2) \in \mathbb{R}^2$ satisfying $\frac{\eta_2}{\eta_1} \in V_\lambda$. The complement in $\mathbb{R}^2 \times \mathbb{R}$ of the regular directed points for u is called the *wavefront set* of u and is denoted by $WF(u)$. Thus, the singular support is measuring the location of the singularities and λ is measuring the direction perpendicular to the singularity.[2]

In the examples presented in Section 4, one can verify the following:

(i) *Point Singularity* $\delta(x)$*:* $\operatorname{sing\,supp}(\delta) = \{0\}$ and $WF(\delta) = \{0\} \times \mathbb{R}$.

(ii) *Linear Singularity* $\nu_p(x)$*:* $\operatorname{sing\,supp}(\nu_p) = \{(-p x_2, x_2) : x_2 \in \mathbb{R}\}$ and $WF(\nu_p) = \{((-p x_2, x_2), p) : x_2 \in \mathbb{R}\}$.

(iii) *Curvilinear Singularity* $B(x)$*:* $\operatorname{sing\,supp}(B) = \{(x_1,x_2) : x_1^2 + x_2^2 = 1\}$ and $WF(B) = \{((x_1,x_2),\lambda) : x_1^2 + x_2^2 = 1,\ \lambda = \frac{x_2}{x_1}\}$.

As observed in Section 4, all these sets are exactly identified by the decay properties of the Continuous Shearlet Transform. Indeed, we have the following general result:

Theorem 5.1. (i) *Let* $\mathcal{R} = \{t_0 \in \mathbb{R}^2$ *: for* t *in a neighborhood* U *of* t_0*,* $|\mathcal{SH}_\psi f(a,s,t)| = O(a^k)$ *and* $|\mathcal{SH}_\psi^{(v)} f(a,s,t)| = O(a^k)$ *as* $a \to 0$*, for all*

[2] This definition is consistent with [6], where the direction of the singularity is described by the angle θ. Observe that our approach does not distinguish between θ and $\theta + \pi$, since the continuous shearlets have frequency support that is symmetric with respect to the origin. However, in Section 6 we discuss a variant of the Continuous Shearlet Transform, which can distinguish these cases.

$k \in \mathbb{N}$, with the $O(\cdot)$–terms uniform over $(s,t) \in [-1,1] \times U\}$. Then

$$\operatorname{sing\,supp}(f)^c = \mathcal{R}.$$

(ii) *Let $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$, where $\mathcal{D}_1 = \{(t_0, s_0) \in \mathbb{R}^2 \times [-1,1]$: for (s,t) in a neighborhood U of (s_0,t_0), $|\mathcal{SH}_\psi f(a,s,t)| = O(a^k)$ as $a \to 0$, for all $k \in \mathbb{N}$, with the $O(\cdot)$–term uniform over $(s,t) \in U\}$ and $\mathcal{D}_2 = \{(t_0,s_0) \in \mathbb{R}^2 \times [1,\infty)$: for $(\frac{1}{s}, t)$ in a neighborhood U of (s_0,t_0), $|\mathcal{SH}_f^{(v)}(a,s,t)| = O(a^k)$ as $a \to 0$, for all $k \in \mathbb{N}$, with the $O(\cdot)$–term uniform over $(\frac{1}{s},t) \in U\}$. Then*

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

The statement (ii) of the theorem shows that the Continuous Shearlet Transform $\mathcal{SH}_\psi f(a,s,t)$ identifies the wavefront set for directions s such that $|s| = |\frac{\xi_2}{\xi_1}| \leq 1$ (in the frequency domain). The Continuous Shearlet Transform $\mathcal{SH}_\psi^{(v)} f(a,s,t)$ identifies the wavefront set for directions s such that $|s| = |\frac{\xi_1}{\xi_2}| \leq 1$, corresponding to $|\frac{\xi_2}{\xi_1}| \geq 1$. This result is in part inspired by Theorems 5.1 and 5.2 in [6] where a similar result is proved for the Continuous Curvelet Transform. Several lemmata will be needed in order to prove Theorem 5.1. In these proofs, some ideas from [6] will be adapted to the distinct mathematical structure of the shearlet approach.

The following lemma shows that if t is outside the support of a function g, then the Continuous Shearlet Transform of g decays rapidly as $a \to 0$.

Lemma 5.2. *Let $g \in L^2(\mathbb{R}^2)$ with $\|g\|_\infty < \infty$, and $a < 1$. If $\operatorname{supp}(g) \subset \mathcal{B} \subset \mathbb{R}^2$, then for all $k > 1$,*

$$|\mathcal{SH}_\psi g(a,s,t)| = |\langle g, \psi_{ast}\rangle| \leq C_k\, C(s)^2\, \|g\|_\infty\, a^{\frac{1}{4}}\, \left(1 + C(s)^{-1} a^{-1} d(t,\mathcal{B})^2\right)^{-k},$$

where $C(s) = \left(1 + \frac{s^2}{2} + \left(s^2 + \frac{s^4}{4}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}$ and C_k is as in Proposition 3.4.

Proof. Since $\|g\|_\infty < \infty$, by Proposition 3.4, for all $k \in \mathbb{N}$, there is a $C_k > 0$ such that:

$$\begin{aligned} |\langle g, \psi_{ast}\rangle| &\leq \|g\|_\infty \int_{\mathcal{B}} |\psi_{ast}(x)| dx \\ &\leq C_k\, \|g\|_\infty\, a^{-\frac{3}{4}} \int_{\mathcal{B}} \left(1 + \|M^{-1}(x-t)\|^2\right)^{-k} dx \\ &= C_k\, \|g\|_\infty\, a^{-\frac{3}{4}} \int_{\mathcal{B}+t} \left(1 + \|M^{-1}x\|^2\right)^{-k} dx, \end{aligned} \tag{5.1}$$

where $M = \begin{pmatrix} a & -\sqrt{a}\, s \\ 0 & \sqrt{a} \end{pmatrix}$. Observe that $\|x\| = \|MM^{-1}x\| \leq \|M\|_{op}\, \|M^{-1}x\|$, and, thus,

$$\|M^{-1}x\| \geq \frac{1}{\|M\|_{op}}\, \|x\|.$$

Since $M = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$ and $\| \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} \|_{op} = \sqrt{a}$, then

$$\|M^{-1}x\| \geq C(s)^{-1}\, a^{-\frac{1}{2}}\, \|x\|,$$

where $C(s) = \| \left(\begin{smallmatrix} 1 & -s \\ 0 & 1 \end{smallmatrix}\right) \|_{op}$. Using these observation in (5.1), we have that

$$\begin{aligned} |\langle g, \psi_{ast} \rangle| &\leq C_k \, \|g\|_\infty \, a^{-\frac{3}{4}} \int_{\mathcal{B}+t} \Big(1 + C(s)^{-2} a^{-1} \|x\|^2\Big)^{-k} dx \\ &\leq C_k \, C(s)^2 \, \|g\|_\infty \, a^{\frac{1}{4}} \int_{C(s)^{-1} a^{-1/2} d(t,\mathcal{B})}^{\infty} (1+r^2)^{-k} r \, dr \\ &= C_k \, C(s)^2 \, \|g\|_\infty \, a^{\frac{1}{4}} \Big(1 + C(s)^{-2} a^{-1} d(t,\mathcal{B})^2\Big)^{-k}. \end{aligned}$$

At last, we compute $C(s)$. We have $\left(\begin{smallmatrix} 1 & -s \\ 0 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 \\ -s & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1+s^2 & -s \\ -s & 1 \end{smallmatrix}\right)$. The largest eigenvalue of this matrix is $\lambda_{\max} = 1 + \frac{s^2}{2} + \Big(s^2 + \frac{s^4}{4}\Big)^{\frac{1}{2}}$. Thus we have

$$C(s) = \left(1 + \frac{s^2}{2} + \left(s^2 + \frac{s^4}{4}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} \quad \text{for all } s \in \mathbb{R}.$$

□

Remark. The dependence on s in Lemma 5.2 cannot be removed. If this were the case, then, for some subset $\mathcal{C}$ of $\mathbb{R}^2$ with positive distance to 0, we would need that

$$B_s \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + sy \\ y \end{pmatrix}$$

has a positive distance from 0 for all $(x, y) \in \mathcal{C}$ independent of s. For this, we only need to consider the "inner" boundary of $\mathcal{C}$, which "separates" it from 0. Let $\epsilon > 0$ and consider the point $(x_\epsilon, \epsilon) \in \mathcal{C}$. S_s maps this point to $(x_\epsilon + s\epsilon, \epsilon)$. Obviously, there exists an $s \in \mathbb{R}$ with $x_\epsilon + s\epsilon = 0$. Since ϵ is arbitrary, we can always find a shear s, which maps a point of $\mathcal{C}$ arbitrarily close to 0. The application of the parabolic scaling via a afterwards doesn't change this fact. Thus, in order to obtain a uniform estimate, it is sufficient to assume that $s \in S \subset \mathbb{R}$, where S is compact.

We can now prove the following inclusions.

Proposition 5.3. *Let $\mathcal{R}$ and $\mathcal{D}$ be defined as in Theorem* 5.1*. Then:*

(i) $\operatorname{sing\,supp}(f)^c \subseteq \mathcal{R}$,
(ii) $WF(f)^c \subseteq \mathcal{D}$.

Proof. (i) Let t_0 be a regular point of f. Then there exists $\phi \in C_0^\infty(\mathbb{R}^2)$ with $\phi(t_0) \equiv 1$ on $B(t_0, \delta)$, which is the ball centered at t_0 with radius δ, such that $\phi f \in C^\infty(\mathbb{R}^2)$. We will show that $t_0 \in \mathcal{R}$. For this, we decompose $\mathcal{SH}_\psi f(a,s,t)$ as

$$\mathcal{SH}_\psi f(a,s,t) = \langle \psi_{ast}, \phi f \rangle + \langle \psi_{ast}, (1-\phi) f \rangle. \tag{5.2}$$

Observe that

$$|\langle \psi_{ast}, \phi f \rangle| \leq a^{\frac{3}{4}} \int_{\mathbb{R}^2} |\hat{\psi}_1(a\xi_1)| \, |\hat{\psi}_2(\tfrac{1}{\sqrt{a}}(\tfrac{\xi_2}{\xi_1} - s))| \, |\widehat{\phi f}(\xi_1, \xi_2)| \, d\xi_1 \, d\xi_2 = I^+ + I^-,$$

where I^+ is the integral restricted to $\xi_1 \geq 0$ and I^- is the integral restricted to $\xi_1 < 0$.

Since $\phi \in C_0^\infty(\mathbb{R}^2)$, then $|\widehat{\phi f}(\xi)| = O(1+|\xi|)^{-k}$. Using this fact together with the assumptions on the support of $\hat\psi_{ast}$, for $k > 2$, we have:

$$\begin{aligned}
I^+ &= a^{\frac{3}{4}} \int_{\mathbb{R}^+\times\mathbb{R}} |\hat\psi_1(a\xi_1)|\,|\hat\psi_2(\tfrac{1}{\sqrt{a}}(\tfrac{\xi_2}{\xi_1}-s))|\,|\widehat{\phi f}(\xi_1,\xi_2)|\,d\xi_1\,d\xi_2 \\
&\le C_k\|\hat\psi\|_\infty a^{\frac{3}{4}} \int_{\frac{1}{2a}}^{\frac{2}{a}} \int_{(s-\sqrt{a})\xi_1}^{(s+\sqrt{a})\xi_1} (1+\xi_1^2+\xi_2^2)^{-k/2}\,d\xi_2\,d\xi_1 \\
&\le C_k\|\hat\psi\|_\infty a^{\frac{3}{4}} \int_{\frac{1}{2a}}^{\frac{2}{a}} (1+\xi_1)^{-k} \int_{(s-\sqrt{a})\xi_1}^{(s+\sqrt{a})\xi_1} d\xi_2\,d\xi_1 \\
&= 2\,C_k\|\hat\psi\|_\infty a^{\frac{3}{4}} \int_{\frac{1}{2a}}^{\frac{2}{a}} (1+\xi_1)^{-k}\sqrt{a}\xi_1\,d\xi_1 \\
&\le 2\,C_k\|\hat\psi\|_\infty a^{\frac{5}{4}} \int_{\frac{1}{2a}}^{\frac{2}{a}} (1+\xi_1)^{-k+1}\,d\xi_1 \\
&\le \frac{2\,C_k\|\hat\psi\|_\infty a^{\frac{5}{4}}}{k-2}\,(2\,a)^{k-2}.
\end{aligned}$$

Thus, I^+ decays rapidly as $a \to 0$, uniformly over $(t,s) \in B(t_0,\frac{\delta}{2})\times\mathbb{R}$. The estimate for I^- is similar and, hence, the first term on the RHS of (5.2) decays rapidly as $a \to 0$, uniformly over $(t,s) \in B(t_0,\frac{\delta}{2})\times\mathbb{R}$.

Using Lemma 5.2, we estimate the second term on the RHS of (5.2) as:

$$|\langle\psi_{ast},(1-\phi)f\rangle| \le C_k\,C(s)^2\,\|(1-\phi)f\|_\infty\,a^{\frac{1}{4}}\,(1+C(s)^{-1}a^{-1}d(t,B(t_0,\delta)^c)^2)^{-k},$$

where $k \in \mathbb{N}$ is arbitrary. Since $\|(1-\phi)f\|_\infty < \infty$ and s is bounded, this yields

$$|\langle\psi_{ast},(1-\phi)f\rangle| = O(a^k) \quad \text{as } a \to 0,$$

uniformly over $(t,s)\in B(t_0,\frac{\delta}{2})\times[-1,1]$. A similar estimate holds when $\mathcal{SH}_\psi f(a,s,t)$ is replaced by $\mathcal{SH}_\psi^{(v)} f(a,s,t)$. This proves (i).

(ii) Let (t_0,s_0) be a regular directed point of f, with $s_0 \in [-1,1]$. Then there exists a $\phi \in C_0^\infty(\mathbb{R}^2)$ with $\phi(t_0) \equiv 1$ on a ball $B(t_0,\delta_1)$ such that, for each $k \in \mathbb{N}$, we have $|\widehat{\phi f}(\xi)| = O((1+|\xi|)^{-k})$ for all $\xi \in \mathbb{R}^2$ satisfying $\frac{\xi_2}{\xi_1} \in B(s_0,\delta_2)$. We will prove that $(t_0,s_0) \in \mathcal{D}$. For this, we decompose $\mathcal{SH}_\psi f(a,s,t)$ as in (5.2). The second term on the RHS of (5.2) can be estimated as in the case (i). For the first term of (5.2), we only need to show that $\operatorname{supp}\hat\psi_{ast} \subset \{\xi \in \mathbb{R}^2 : \frac{\xi_2}{\xi_1} \in B(s_0,\delta_2)\}$ for all $(s,t) \in B(s_0,\delta_2)\times B(t_0,\delta_1)$, since in this cone $\widehat{\phi f}$ decays rapidly. As above, we only consider the case $\xi_1 > 0$; the case $\xi_1 \le 0$ is similar. The support of $\hat\psi_{ast}$ in this half-plane is given by

$$\{(\xi_1,\xi_2) : \xi_1 \in [\tfrac{1}{2a},\tfrac{2}{a}], \xi_2 \in \xi_1[s-\sqrt{a},s+\sqrt{a}]\}.$$

Let $(s,t) \in B(s_0,\delta_2)\times B(t_0,\delta_1)$. The cone $\{\xi \in \mathbb{R}^2 : \frac{\xi_2}{\xi_1} \in B(s_0,\delta_2)\}$ is bounded by the lines $\xi_2 = (s_0-\delta_2)\xi_1$ and $\xi_2 = (s_0+\delta_2)\xi_1$. Now let $(\xi_1,\xi_2) \in \operatorname{supp}\hat\psi_{ast}$. Then, for a small enough, we have

$$|\tfrac{\xi_2}{\xi_1} - s_0| \le \sqrt{a} \le \delta_2,$$

and this completes the proof in this case. In case $|s_0| \geq 1$ (this corresponds to $|\frac{\xi_2}{\xi_1}| \leq 1$), we proceed exactly as above, using the transform $\mathcal{SH}_\psi^{(v)} f(a,s,t)$ rather than $\mathcal{SH}_\psi f(a,s,t)$. □

For the converse inclusions we need some additional lemmata. For simplicity of notation, in the following proofs the symbols C' and C_k are generic constants and may vary from expression to expression (in the case of C_k, the constant depends on k).

Lemma 5.4. *Let $S \subset \mathbb{R}$ be a compact set, and let $g \in L^2(\mathbb{R}^2)$ with $\|g\|_\infty < \infty$. Suppose that* $\operatorname{supp} g \subset \mathcal{B}$ *for some $\mathcal{B} \subset \mathbb{R}^2$ and define $(\mathcal{B}^\eta)^c = \{x \in \mathbb{R}^2 : d(x,\mathcal{B}) > \eta\}$. Further define $h \in L^2(\mathbb{R}^2)$ by*

$$\hat{h}(\xi) = \int_0^\infty \int_{(\mathcal{B}^\eta)^c} \int_S \mathcal{SH}_\psi g(a,s,t)\, \hat{\psi}_{ast}(\xi)\, ds\, dt\, \frac{da}{a^3}.$$

Then $\hat{h}(\xi)$ decays rapidly as $|\xi| \to \infty$ with constants dependent only on $\|g\|_\infty$ and η.

Proof. Using the fact that S is compact, Lemma 5.2 implies that, for each $k > 0$,

$$|\mathcal{SH}_\psi g(a,s,t)| \leq C_k\, a^{\frac{1}{4}}\, (1 + a^{-1} d(t,\mathcal{B})^2)^{-k},$$

where C_k depends on $\|g\|_\infty$ but not on s. By definition, the support of $\hat{\psi}_{ast}$ is contained in the set

$$\Gamma(a,s) = \{\xi \in \mathbb{R}^2 : \tfrac{1}{2} \leq a|\xi| \leq 2,\ |s - \tfrac{\xi_2}{\xi_1}| \leq \sqrt{a}\}. \tag{5.3}$$

Thus, $|\hat{\psi}_{ast}(\xi)| \leq C' a^{\frac{3}{4}} \chi_{\Gamma(a,s)}(\xi)$ and

$$\int_S \chi_{\Gamma(a,s)}(\xi)\, ds \leq \int_{S \cap \left[\frac{\xi_2}{\xi_1} - \sqrt{a}, \frac{\xi_2}{\xi_1} + \sqrt{a}\right]} ds \leq C' \sqrt{a}. \tag{5.4}$$

Collecting the above arguments,

$$\begin{aligned}
\hat{h}(\xi) &\leq \int_0^\infty \int_{(\mathcal{B}^\eta)^c} \int_S |\mathcal{SH}_\psi g(a,s,t)|\, |\hat{\psi}_{ast}(\xi)|\, ds\, dt\, \frac{da}{a^3} \\
&\leq C_k \int_0^\infty \int_{(\mathcal{B}^\eta)^c} \int_S a\, \chi_{\Gamma(a,s)}(\xi)\, (1 + a^{-1} d(t,\mathcal{B})^2)^{-k}\, ds\, dt\, \frac{da}{a^3} \\
&\leq C_k \int_0^\infty \int_{(\mathcal{B}^\eta)^c} \int_S \chi_{\Gamma(a,s)}(\xi)\, ds\, a\, (1 + a^{-1} d(t,\mathcal{B})^2)^{-k}\, dt\, \frac{da}{a^3} \\
&\leq C_k \int_{\frac{1}{2|\xi|}}^{\frac{2}{|\xi|}} a^{-\frac{3}{2}} \int_{(\mathcal{B}^\eta)^c} (1 + a^{-1} d(t,\mathcal{B})^2)^{-k}\, dt\, da \\
&\leq C_k \int_{\frac{1}{2|\xi|}}^{\frac{2}{|\xi|}} a^{-\frac{3}{2}} \int_\eta^\infty (1 + a^{-1} r^2)^{-k}\, r\, dr\, da \\
&\leq C_k \int_{\frac{1}{2|\xi|}}^{\frac{2}{|\xi|}} a^{-\frac{1}{2}}\, (1 + a^{-1}\eta^2)^{-k+2}\, da \\
&\leq C_k\, |\xi|^{-\frac{1}{2}}\, (1 + |\xi|\, \eta^2)^{-k+2}.
\end{aligned}$$

Since this holds for each $k > 0$, $\hat{h}(\xi)$ decays rapidly as $|\xi| \to \infty$. □

Lemma 5.5. *Let $S \subset \mathbb{R}$ and $\mathcal{B} \subset \mathbb{R}^2$ be compact sets. Suppose that $G(a,s,t)$ decays rapidly as $a \to 0$ uniformly for $(s,t) \in S \times \mathcal{B}$. Define $h \in L^2(\mathbb{R}^2)$ by*

$$\hat{h}(\xi) = \int_0^\infty \int_{\mathcal{B}} \int_S G(a,s,t)\, \hat{\psi}_{ast}(\xi)\, ds\, dt\, \frac{da}{a^3}.$$

Then $\hat{h}(\xi)$ decays rapidly as $|\xi| \to \infty$.

Proof. As in Lemma 5.4, we will use the fact that $|\hat{\psi}_{ast}(\xi)| \le C'\, a^{\frac{3}{4}} \chi_{\Gamma(a,s)}(\xi)$, where $\Gamma(a,s)$ is given by (5.3) and the estimate (5.4). Also, by hypothesis, for each $k > 0$ and $a > 0$ we have

$$\sup\{|G(a,s,t)| : |\xi| \in [\tfrac{1}{2a}, \tfrac{2}{a}],\ t \in \mathcal{B}\} \le C_k\, a^k.$$

Using all these observations, we have that, for each $k > 0$,

$$\begin{aligned} |\hat{h}(\xi)| &\le \int_0^\infty \int_{\mathcal{B}} \int_S |G(a,s,t)|\, |\hat{\psi}_{ast}(\xi)|\, ds\, dt\, \frac{da}{a^3} \\ &\le C_k \int_0^\infty \int_{\mathcal{B}} \int_S \chi_{\Gamma(a,s)}(\xi)\, a^{k-\frac{9}{4}}\, ds\, dt\, da \\ &\le C_k \int_{\frac{1}{2|\xi|}}^{\frac{2}{|\xi|}} a^{k-\frac{7}{4}}\, da \\ &\le C_k\, |\xi|^{-k+\frac{3}{4}}. \end{aligned}$$

□

The proof of the following lemma adapts several ideas from [3, Lemma 2.3].

Lemma 5.6. *Suppose $0 \le a_0 \le a_1 < 1$ and $|s| \le s_0$. Then for $K > 1$, there is a constant C_K, dependent on K only, such that:*

$$|\langle \psi_{a_0 st}, \psi_{a_1 s' t'} \rangle| \le C_K \left(1 + \frac{a_1}{a_0}\right)^{-K} \left(1 + \frac{|s-s'|^2}{a_1}\right)^{-K} \left(1 + \frac{\|(t-t')\|^2}{a_1}\right)^{-K}.$$

Proof. By the properties of ψ, for $\|\xi\| > \frac{1}{2}$ and any $k > 0$, there exists a corresponding constant C_k such that

$$|\hat{\psi}(\xi)| \le C_k \frac{1}{(1 + |\xi_1| + |\xi_2|)^k}.$$

Further, $\hat{\psi}(\xi) = 0$ for $\|\xi\| < \frac{1}{2}$. Thus, observing that $M_{as}^t \xi = (a\,\xi_1, \sqrt{a}\,\xi_2 - \sqrt{a}\,s\,\xi_1)$, we obtain

$$|\hat{\psi}_{ast}(\xi)| \le C_k \frac{a^{\frac{3}{4}}}{(1 + a\,|\xi_1| + \sqrt{a}\,|\xi_2 - s\,\xi_1|)^k}.$$

Using polar coordinates, by writing $\xi_1 = r\cos\theta$ and $\xi_2 = r\sin\theta$, this expression can be written as

$$|\hat{\psi}_{ast}(r,\theta)| \le C_k \frac{a^{\frac{3}{4}}}{(1 + a\,r|\cos\theta| + \sqrt{a}\,r|\sin\theta - s\,\cos\theta|)^k}.$$

For $|\theta| \le \beta$, with $\beta < \pi/2$, using the assumption $|s| \le s_0$, the last expression can be controlled by

$$|\hat{\psi}_{ast}(r,\theta)| \le C_k \frac{a^{\frac{3}{4}}}{(1 + a\,r + \sqrt{a}\,r\,|\sin\theta - s|)^k}. \tag{5.5}$$

In addition, since $\sin\theta \sim \theta$ on $|\theta| \le \pi/2$, we can replace $\sin\theta$ with θ in the above expression.

Let $\Delta s = s - s'$, $a_0 = \min(a, a')$ and $a_1 = \max(a, a')$. Using (5.5), and applying the same argument on $|\theta| \le \pi/2$ and $\pi/2 < |\theta| \le \pi$, it follows that

$$\begin{aligned}
&\int_{|\frac{\xi_2}{\xi_1}|<\tan\beta} \left|\hat\psi_{a's't'}(\xi)\,\hat\psi_{ast}(\xi)\right| d\xi \\
&\le C_k \int_{\frac{1}{a_0}}^{\infty}\int_{-\pi}^{\pi} \frac{(a\,a')^{\frac34}\, r}{(1+a\,r+\sqrt{a}\,r\,|\theta-s|)^k\left(1+a'\,r+\sqrt{a'}\,r\,|\theta-s'|\right)^k}\, d\theta\, dr \\
&\le C_k \int_{a_0^{-1}}^{\infty} \frac{(a\,a')^{\frac34} r}{(1+a\,r)^{k'}\,(1+a'\,r)^k} \int_{-\infty}^{\infty} \frac{1}{(1+\alpha\,|\theta|)^k\,(1+\alpha'\,|\theta+\Delta s|)^k}\, d\theta\, dr,
\end{aligned}$$

where $\alpha = \frac{\sqrt{a}\,r}{1+a\,r}$, $\alpha' = \frac{\sqrt{a'}\,r}{1+a'\,r}$. Using a simple calculation, for $\alpha > \alpha' > 1$, $k > 1$,

$$\int_{-\infty}^{\infty} \frac{1}{(1+\alpha\,|\theta|)^k\,(1+\alpha'\,|\theta+\Delta s|)^k}\, d\theta \le C_k\, \frac{1}{\alpha\,(1+\alpha'\,|\Delta s|)^k}.$$

From the definition of α', for $r \ge 1/a'$, we obtain $\frac12 \frac{1}{\sqrt{a'}} \le \alpha' \le \frac{1}{\sqrt{a'}}$. Thus, for $r \ge 1/a'$, provided $k > 1$, the last expression gives

$$\int_{-\infty}^{\infty} \frac{1}{(1+\alpha\,|\theta|)^k\,(1+\alpha'\,|\theta+\Delta s|)^k}\, d\theta \le C_k\, \frac{(1+a\,r)}{\sqrt{a}\,r}\left(1+\frac{|\Delta s|}{\sqrt{a'}}\right)^{-k}. \tag{5.6}$$

Another simple estimate, provided $k' > 1$, yields

$$\int_{\frac{1}{a_0}}^{\infty} \frac{1}{(1+a_0\,r)^{k'}\,(1+a_1\,r)^k}\, dr \le C_{k'}\, \frac{1}{a_0}\left(1+\frac{a_1}{a_0}\right)^{-k}. \tag{5.7}$$

Thus, using (5.6) and (5.7), for $\alpha > \alpha' > 1$, $a_0 = a'$ and $a_1 = a$, we obtain

$$\begin{aligned}
\int_{|\frac{\xi_2}{\xi_1}|<\tan\beta} \left|\hat\psi_{a's't'}(\xi)\,\hat\psi_{ast}(\xi)\right| d\xi &\le C_k \left(\frac{a_1}{a_0}\right)^{\frac14}\left(1+\frac{a_1}{a_0}\right)^{-k+1}\left(1+\frac{|\Delta s|}{\sqrt{a_0}}\right)^{-k} \\
&\le C_k \left(1+\frac{a_1}{a_0}\right)^{-k+2}\left(1+\frac{|\Delta s|}{\sqrt{a_0}}\right)^{-k}.
\end{aligned}$$

Similarly, for $\alpha > \alpha' > 1$, $a_1 = a'$ and $a_0 = a$, a similar calculation gives

$$\begin{aligned}
\int_{|\frac{\xi_2}{\xi_1}|<\tan\beta} \left|\hat\psi_{a's't'}(\xi)\,\hat\psi_{ast}(\xi)\right| d\xi &\le C_k \left(\frac{a_1}{a_0}\right)^{\frac34}\left(1+\frac{a_1}{a_0}\right)^{-k}\left(1+\frac{|\Delta s|}{\sqrt{a_1}}\right)^{-k} \\
&\le C_k \left(1+\frac{a_1}{a_0}\right)^{-k+1}\left(1+\frac{|\Delta s|}{\sqrt{a_1}}\right)^{-k}.
\end{aligned}$$

In general, renaming the index k, we can show that

$$\int_{|\frac{\xi_2}{\xi_1}|<\tan\beta} \left|\hat\psi_{a's't'}(\xi)\,\hat\psi_{ast}(\xi)\right| d\xi \le C_k \left(1+\frac{a_1}{a_0}\right)^{-k}\left(1+\frac{|\Delta s|}{\sqrt{a_1}}\right)^{-k}. \tag{5.8}$$

To complete the proof, we will employ the formulas

$$\frac{\partial}{\partial \xi_1}\hat{\psi}_{ast}(\xi) = (a - \sqrt{a}\, s)\, \hat{\psi}_{ast}(\xi), \quad \frac{\partial}{\partial \xi_2}\hat{\psi}_{ast}(\xi) = \sqrt{a}\, \hat{\psi}_{ast}(\xi),$$
$$\frac{\partial^2}{\partial \xi_1^2}\hat{\psi}_{ast}(\xi) = (a - \sqrt{a}\, s)^2\, \hat{\psi}_{ast}(\xi), \quad \frac{\partial^2}{\partial \xi_2^2}\hat{\psi}_{ast}(\xi) = a\, \hat{\psi}_{ast}(\xi).$$

Thus, observing that $a, a' < 1$ and $|s| < s_0$ yields

$$\left|\Delta_\xi\, \hat{\psi}_{ast}(\xi)\, \overline{\hat{\psi}_{a's't'}(\xi)}\right| \le C'\, a_1\, |\hat{\psi}_{ast}(\xi)|\, |\hat{\psi}_{a's't'}(\xi)|.$$

Set

$$L = I - \frac{\Delta_\xi}{(2\pi)^2\, a_1}.$$

On the one hand, for each k, we have

$$\left|L^k \left(\hat{\psi}_{ast}\, \overline{\hat{\psi}_{a's't'}}\right)(\xi)\right| \le C'\, |\hat{\psi}_{ast}(\xi)|\, |\hat{\psi}_{a's't'}(\xi)|. \tag{5.9}$$

On the other hand,

$$L^k(e^{-2\pi i \xi (t-t')}) = \left(1 + \frac{\|t-t'\|^2}{a_1}\right)^k e^{-2\pi i \xi (t-t')}. \tag{5.10}$$

Repeated integrations by parts give

$$\begin{aligned}
\langle \psi_{ast}, \psi_{a's't'}\rangle &= \int_{|\frac{\xi_2}{\xi_1}| < \tan\beta} \hat{\psi}_{ast}(\xi)\, \overline{\hat{\psi}_{a's't'}(\xi)}\, d\xi \\
&= \int_{|\frac{\xi_2}{\xi_1}| < \tan\beta} (a\, a')^{3/4}\, \hat{\psi}(M_{as}^t \xi)\, \overline{\hat{\psi}(M_{a's'}^t \xi)}\, e^{-2\pi i \xi (t-t')}\, d\xi \\
&= \int_{|\frac{\xi_2}{\xi_1}| < \tan\beta} L^k \left((a\, a')^{3/4}\, \hat{\psi}(M_{as}^t \xi)\, \overline{\hat{\psi}(M_{a's'}^t \xi)}\right) L^{-k}\left(e^{-2\pi i \xi (t-t')}\right)\, d\xi.
\end{aligned}$$

Therefore, from the last expression, using (5.8)–(5.10), it follows that

$$|\langle \psi_{ast}, \psi_{a's't'}\rangle| \le C_k\, \left(1 + \frac{\|t-t'\|^2}{a_1}\right)^{-k} \left(1 + \frac{a_1}{a_0}\right)^{-k} \left(1 + \frac{|\Delta s|}{\sqrt{a_1}}\right)^{-k}.$$

The proof is completed recalling that, for $m > 0$, $(1+|x|)^{-2m} \sim (1+|x|^2)^{-m}$. That is, there are constants $C_1, C_2 > 0$ such that $C_1\,(1+|x|^2)^{-m} \le (1+|x|)^{-2m} \le C_2\,(1+|x|^2)^{-m}$. □

From Lemma 5.6, the following result can be easily deduced.

Lemma 5.7. *Let $\phi_1 \in C^\infty(\mathbb{R}^2)$ be supported in $B(0,1)$, and define $\phi(x) = \phi_1(a_\phi^{-1}(x-t))$.*

(i) *Suppose $0 \le \sqrt{a_0} \le \sqrt{a_1} \le a_\phi < 1$. Then for $K > 0$,*

$$|\langle \phi\, \psi_{a_0 st}, \psi_{a_1 s't'}\rangle| \le C_K\, \left(1 + \frac{a_1}{a_0}\right)^{-K} \left(1 + \frac{|s-s'|^2}{a_1}\right)^{-K} \left(1 + \frac{\|(t-t')\|^2}{a_1}\right)^{-K}.$$

(ii) *Suppose $0 \le \sqrt{a_0} \le a_\phi \le \sqrt{a_1} < 1$, $a_1 \le a_\phi$. Then for $K > 0$,*

$$|\langle \phi\, \psi_{a_0 st}, \psi_{a_1 s't'}\rangle| \le C_K\, \left(1 + \frac{a_1}{a_0}\right)^{-K} \left(1 + \frac{|s-s'|^2}{a_\phi^2}\right)^{-K} \left(1 + \frac{\|(t-t')\|^2}{a_1}\right)^{-K}.$$

(iii) *Suppose* $0 \le \sqrt{a_0} \le a_\phi \le a_1 \le \sqrt{a_1} < 1$. *Then for* $K > 0$,

$$|\langle \phi\, \psi_{a_0 s t}, \psi_{a_1 s' t'}\rangle| \le C_K \left(1 + \frac{a_\phi}{a_0}\right)^{-K} \left(1 + \frac{\|(t-t')\|^2}{a_\phi^2}\right)^{-K}.$$

Now we can complete the proof of Theorem 5.1.

Proof of Theorem 5.1. Since one direction was proved by Proposition 5.3, we only have to prove the inclusions:

(i) $\mathcal{R} \subseteq \operatorname{sing\,supp}(f)^c$;
(ii) $\mathcal{D} \subseteq \mathrm{WF}(f)^c$.

First we prove (i). Let $t_0 \in \mathcal{R}$. Then for all $t \in B(t_0,\delta)$, where $B(t_0,\delta)$ denotes a ball centered at t_0 of radius δ, we have that $|\mathcal{SH}_\psi f(a,s,t)| = O(a^k)$ as $a \to 0$, for all $k \in \mathbb{N}$ with the $O(\cdot)$–term uniform over $(t,s) \in B(t_0,\delta) \times [-1,1]$. A similar estimate holds for $\mathcal{SH}_\psi^{(v)} f(a,s,t)$.

Choose $\phi \in C^\infty(\mathbb{R}^2)$ to be supported in $B(t_0,\nu)$ with $\nu \ll \delta$ and let $\eta = \frac{\delta}{2}$. Let $\hat{g}_0(\xi) = (\phi\, P(f))^\wedge(\xi)$, where $P(f) = \int_{\mathbb{R}} \langle f, T_b W\rangle\, T_b W\, db$ and W is the window function defined by (3.6). Further, define $\hat{g}_i$, $i = 1,\dots,4$ as

$$\hat{g}_i(\xi) = \chi_{C_1}(\xi) \int_{\mathcal{Q}_i} \hat{\psi}_{ast}(\xi)\, \mathcal{SH}_\psi g(a,s,t)\, d\mu(a,s,t), \quad i = 1,2,$$

$$\hat{g}_{i+2}(\xi) = \chi_{C_2}(\xi) \int_{\mathcal{Q}_i} \hat{\psi}_{ast}(\xi)\, \mathcal{SH}_\psi^{(v)} g(a,s,t)\, d\mu(a,s,t), \quad i = 1,2,$$

where C_1, C_2 are defined after equation (3.6), $d\mu(a,s,t) = \frac{da}{a^3}\, ds\, dt$, and $\mathcal{Q}_1 = [0,1] \times [-2,2] \times B(t_0,\eta)$ and $\mathcal{Q}_2 = [0,1] \times [-2,2] \times B(t_0,\eta)^c$. Now set $g = \phi f$ and consider the decomposition

$$\widehat{\phi f}(\xi) = \hat{g}_0(\xi) + \hat{g}_1(\xi) + \hat{g}_2(\xi) + \hat{g}_3(\xi) + \hat{g}_4(\xi).$$

The term $\hat{g}_0(\xi)$ decays rapidly as $|\xi| \to \infty$ since ϕ and $P(f)$ are in C^∞. The term $\hat{g}_2(\xi)$ decays rapidly as $|\xi| \to \infty$ by Lemma 5.4. In addition, by Lemma 5.5, $\hat{g}_1(\xi)$ decays rapidly as $|\xi| \to \infty$ provided that $\mathcal{SH}_\psi g$ decays rapidly as $a \to 0$ uniformly over $(t,s) \in B(t_0,\eta) \times [-2,2]$. In the sequel, we will only analyze the terms $\hat{g}_i$ for $i = 1,2$; the cases $i = 3,4$ are similar.

We claim that $\mathcal{SH}_\psi g$ indeed decays rapidly as $a \to 0$ uniformly over $B(t_0,\eta) \times [-2,2]$. In order to prove this, we decompose f as $f = P(f) + P_{C_1} f + P_{C_2} f$, where $(P_{C_1} f)^\wedge = \hat{f}\,\chi_{C_1}$ and $(P_{C_2} f)^\wedge = \hat{f}\,\chi_{C_2}$. It is clear that $\mathcal{SH}_\psi(\phi P(f))$ decays rapidly by the smoothness of ϕ and $P(f)$. Next, we examine the term $P_{C_1} f$. The analysis of $P_{C_2} f$ is very similar and will be omitted. We use the decomposition $P_{C_1} f = f_1 + f_2$, where

$$f_i(x) = \int_{\mathcal{Q}_i} \psi_{ast}(x)\, \mathcal{SH}_\psi f(a,s,t)\, d\mu(a,s,t), \quad i = 1,2.$$

Let us start by considering the term corresponding to f_1. First we observe that

$$\mathcal{SH}_\psi(\phi f_1)(a,s,t) = \langle \phi\, f_1, \psi_{ast}\rangle = \int_{\mathcal{Q}_1} \langle \phi\, \psi_{ast}, \psi_{a's't'}\rangle\, \mathcal{SH}_\psi f(a',s',t')\, d\mu(a',s',t').$$

We will now decompose $\mathcal{Q}_1 = \mathcal{Q}_{10} \cup \mathcal{Q}_{11} \cup \mathcal{Q}_{12}$, corresponding to $a' > \delta$, $a' \le \delta < \sqrt{a'}$ and $\sqrt{a'} \le \delta$, respectively. In case $\sqrt{a}, \sqrt{a'} \le \delta$, by Lemma 5.7, we obtain

$$|\langle \phi\, \psi_{ast}, \psi_{a's't'}\rangle| \le C_K \left(1 + \frac{a_1}{a_0}\right)^{-K} \left(1 + \frac{\|(t-t')\|^2}{a_1}\right)^{-K}. \tag{5.11}$$

This implies that, for $m > 4$ and $K \geq m-1$,

$$\int_0^\delta \left(1+\frac{a_1}{a_0}\right)^{-K} (a')^m \,\frac{da'}{(a')^3} \leq C_{m,K}\, a^{m-2}, \qquad 0<a<\delta. \tag{5.12}$$

In fact, for $a' = a_0 \leq a = a_1$,

$$\int_0^a \left(1+\frac{a}{a'}\right)^{-K} (a')^m \,\frac{da'}{(a')^3} = a^{m-2}\int_0^1 (1+x)^{-K}\,dx = C_K\, a^{m-2}.$$

For $a = a_0 \leq a' = a_1$,

$$\begin{aligned}\int_a^\delta \left(1+\frac{a'}{a}\right)^{-K} (a')^m \,\frac{da'}{(a')^3} &= a^{m-2}\int_1^{\delta/a} x^{m-3}\,(1+x)^{-K}\,dx \\ &\leq a^{m-2}\int_1^\infty x^{m-3}\,(1+x)^{-K}\,dx = C_{K,m}\,a^{m-2}.\end{aligned}$$

Thus, (5.12) follows from the last two estimates. Using (5.12) it follows that

$$\begin{aligned}&\int_{\mathcal{Q}_{12}} \langle \phi\,\psi_{ast}, \psi_{a's't'}\rangle\, \mathcal{SH}_\psi f(a',s',t')\,d\mu(a',s',t') \\ &\quad\leq C' \int_{-2}^2 \int_{B(t_0,\eta)} \int_0^\delta \left(1+\frac{a_1}{a_0}\right)^{-K} (a')^m\,\frac{da'}{(a')^3}\,dt'\,ds' \\ &\quad\leq C_m\; a^{m-2},\end{aligned}$$

for all $m > 4$. Using the other cases of Lemma 5.7, one can show similar estimates for the integrals over the sets $\mathcal{Q}_{10}$ and $\mathcal{Q}_{11}$. This proves that $\mathcal{SH}_{\phi f_1}(a,s,t)$ decays rapidly for $a \to 0$ uniformly over $B(t_0,\eta)\times[-2,2]$.

Let us now consider the term corresponding to f_2:

$$\mathcal{SH}_\psi(\phi f_2)(a,s,t) = \langle \phi\, f_2, \psi_{ast}\rangle = \int_{\mathcal{Q}_1} \langle \phi\,\psi_{ast}, \psi_{a's't'}\rangle\, \mathcal{SH}_\psi f(a',s',t')\,d\mu(a',s',t').$$

We will decompose $\mathcal{Q}_2 = \mathcal{Q}_{21}\cup\mathcal{Q}_{22}$, corresponding to $\|(t-t')\| > \eta$ and $\|(t-t')\| \leq \eta$, respectively. Observe that, for $\|(t-t')\| > \eta$ and $K > 1$,

$$\int_{B(t_0,\eta)^c} \left(1+\frac{\|(t-t')\|^2}{a_1}\right)^{-K} dt' \leq \int_\eta^\infty \left(1+\frac{r^2}{a_1}\right)^{-K} r\,dr \leq C'\,a_1 \left(1+\frac{\eta}{a_1}\right)^{-K+2}.$$

Further notice that, on the region $\mathcal{Q}_{21}$, the function $\mathcal{SH}_\psi f(a',s',t')$ is bounded by $C'\,(a')^{3/4}$ since f is bounded. Thus

$$\begin{aligned}&\int_{\mathcal{Q}_{21}} \langle \phi\,\psi_{ast}, \psi_{a's't'}\rangle\, \mathcal{SH}_\psi f(a',s',t')\,d\mu(a',s',t') \\ &\leq C' \int_{-2}^2\int_0^\delta \int_{B(t_0,\eta)^c} \left(1+\frac{\|(t-t')\|^2}{a_1}\right)^{-K} dt' \left(1+\frac{a_1}{a_0}\right)^{-K} (a')^{3/4}\,\frac{da'}{(a')^3}\,ds' \\ &\leq C' \int_0^\eta a_1\,\left(1+\frac{\eta}{a_1}\right)^{-K+2} \left(1+\frac{a_1}{a_0}\right)^{-K} \frac{da'}{(a')^{9/4}},\end{aligned}$$

and this is of rapid decay, as $a \to 0$, uniformly over $\mathcal{Q}_{21}$. As for the region $\mathcal{Q}_{22}$, if $t \in B(t_0,\eta)$ and $\|(t-t')\| > \eta$, then $t' \in B(t_0,\delta)$ and thus the function $\mathcal{SH}_\psi f$ decays rapidly, for $a \to 0$, over this region. Repeating the analysis from the case $\mathcal{Q}_{12}$, we can prove that $\int_{\mathcal{Q}_{22}} \langle \phi\,\psi_{ast}, \psi_{a's't'}\rangle\, \Psi_f(a',s',t')\,d\mu(a',s',t')$ is of rapid decay, as $a \to 0$, uniformly over $\mathcal{Q}_{22}$. Combining these observations, we conclude that $\mathcal{SH}_\psi(\phi f_2)(a,s,t)$ decays rapidly as $a \to 0$ uniformly over $B(t_0,\eta) \times [-2,2]$.

It follows that $\mathcal{SH}_\psi g(a,s,t)$ decays rapidly as $a \to 0$ uniformly for all $(s,t) \in B(t_0,\eta) \times [-2,2]$ and, thus, by Lemma 5.5, $\hat{g}_1(\xi)$ decays rapidly as $|\xi| \to \infty$. We can now conclude that $\hat{g}$ decays rapidly as $|\xi| \to \infty$, hence completing the proof of (i).

For part (ii), we only sketch the idea of the proof, since it is very similar to part (i). Let $(t_0,s_0) \in \mathcal{D}$. We consider separately the cases $|s_0| \le 1$ and $|s_0| \ge 1$. In the first case, for all $t \in B(t_0,\delta)$ and $s \in B(s_0,\Delta)$, we have that $|\mathcal{SH}_\psi f(a,s,t)| = O(a^k)$, as $a \to 0$, for all $k \in \mathbb{N}$ with the $O(\cdot)$–term uniform over $(t,s) \in B(t_0,\delta) \times B(s_0,\Delta)$. Choose $\phi \in L^2(\mathbb{R}^2)$ which is supported in a ball $B(t_0,\nu)$ with $\nu \ll \delta$ and let $\eta = \frac{\delta}{2}$. Then the proof proceeds as in part (i), replacing $B(t_0,\delta) \times [-2,2]$ with $B(t_0,\delta) \times B(s_0,\Delta)$. Also, for the estimates involving inner products of ψ_{ast} and $\psi_{a's't'}$ we will now use Lemma 5.7 including the directionally sensitive term. For example, when $\sqrt{a}, \sqrt{a'} \le \delta$, by Lemma 5.7 we will use the estimate

$$|\langle \phi\, \psi_{ast}, \psi_{a's't'} \rangle| \le C_K \left(1 + \frac{a_1}{a_0}\right)^{-K} \left(1 + \frac{|s-s'|^2}{a_1}\right)^{-K} \left(1 + \frac{\|(t-t')\|^2}{a_1}\right)^{-K},$$

rather than (5.11). We can proceed similarly for the other estimates. The proof for the case $|s_0| \ge 1$ is exactly the same, with the transform $\mathcal{SH}_\psi^{(v)} f(a,s,t)$ replacing $\mathcal{SH}_\psi f(a,s,t)$. □

6. Extensions and generalizations of the Continuous Shearlet Transform

As mentioned above, there are several variants and generalizations of the continuous shearlets introduced in Section 3. In fact, using the theory of affine systems, we can obtain several other examples of continuous shearlets depending on the three variables: scale, shear and location.

For example, we can generalize our construction by considering the case where Λ is given by (3.1) and M is of the form

$$M_\delta = \begin{pmatrix} a & -a^\delta\, s \\ 0 & a^\delta \end{pmatrix} = B\, A_\delta, \quad a \in I, s \in S,$$

where $B = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$, $A_\delta = \begin{pmatrix} a & 0 \\ 0 & a^\delta \end{pmatrix}$ and $0 \le \delta \le 1$ is fixed. If $\delta = \frac{1}{2}$, we obtain the Continuous Shearlet Transform as defined in Section 3. In general, for other choices of δ, A_δ will provide different kinds of anisotropic scaling (up to the case $\delta = 1$, where the dilation is isotropic). Using a construction similar to the one given in Section 3 for the continuous shearlet systems, for each $0 \le \delta \le 1$ we can construct well localized systems of the form

$$\{\psi_{ast} = T_t\, D_{M_\delta}\, \psi : (M_\delta, t) \in \Lambda\},$$

where $\hat{\psi}_{ast}$ is supported on the set:

$$\operatorname{supp} \hat{\psi}_{ast} \subset \{(\xi_1,\xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}],\ |\tfrac{\xi_2}{\xi_1} - s| \le a^{1-\delta}\}.$$

This provides a new family of transforms $\mathcal{SH}_\psi^\delta f(a,s,t) = \langle f, \psi_{ast} \rangle$, for various values of δ. It turns out that, provided $0 \le \delta < 1$, the behavior of the transforms $\mathcal{SH}_\psi^\delta f(a,s,t)$ is very similar to the Continuous Shearlet Transform in dealing with pointwise and linear singularities. More precisely, one can repeat the analysis in Sections 4.1, 4.2 and 4.3 using $\mathcal{SH}_\psi^\delta f(a,s,t)$ (notice $\delta \neq 1$). Also in the case of curvilinear singularities (see Section 4.4), the behavior of the transforms $\mathcal{SH}_\psi^\delta f(a,s,t)$ is similar to the Continuous Shearlet Transform, provided $0 < \delta < 1$. However, our proof of Theorem 5.1 makes use of the fact that $\delta = \frac{1}{2}$; that is, we need to use the Continuous Shearlet Transform.

Another variant of affine systems generated by Λ, given by (3.1), is obtained by reversing the order of the shear and dilation matrices, namely, by letting

$$M_\delta = \begin{pmatrix} a & -a\,s \\ 0 & a^\delta \end{pmatrix} = A_\delta\, B, \quad a \in I, s \in S,$$

where A_δ, B are defined as above, and $0 \leq \delta \leq 1$ is fixed. Also in this case, we can construct variants of the continuous shearlet systems. Using the same ideas as above, we obtain a system of functions ψ_{ast} with support

$$\operatorname{supp} \hat{\psi}_{ast} \subset \{(\xi_1,\xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}],\ |a^{1-\delta}\tfrac{\xi_2}{\xi_1} - s| \leq a^{1-\delta}\}.$$

It turns out (as one can see from the support condition) that the transform associated with these systems is not even able to "locate" the linear singularities, in the sense described in Section 4.2.

The frequency support of the continuous shearlets is symmetric with respect to the origin. Therefore the Continuous Shearlet Transform is unable to distinguish the orientation associated with the angle θ from the angle $\theta + \pi$ (cf. Footnote 2). In order to be able to distinguish these two directions, we can modify our construction as follows. Let $\psi \in L^2(\mathbb{R}^2)$ be defined as in Section 3, except that $\operatorname{supp} \hat{\psi}_1 \subset [\frac{1}{2}, 2]$. That is, we have a one-sided version of the shearlet $\hat{\psi}$ illustrated in Figure 1. It is then clear that, if we consider the affine system generated by ψ under the action of Λ be given by (3.1) and (3.2), this cannot provide a reproducing system for all of $L^2(\mathbb{R}^2)$. In fact, ψ is a wavelet for the subspace $L^2(H)^\vee$, where $H = \{(\xi_1,\xi_2) \in \mathbb{R}^2 : \xi_1 \geq 0\}$. In order to obtain a wavelet for the space $L^2(\mathbb{R}^2)$, we can modify the set Λ as follows. Let Λ' be given by (3.1), where $G \subset GL_2(\mathbb{R})$ is the set of matrices:

$$G = \left\{ M = \begin{pmatrix} \ell a & -\ell\sqrt{a}\,s \\ 0 & \ell\sqrt{a} \end{pmatrix}, \quad a \in I,\ s \in S,\ \ell = -1, 1 \right\}, \tag{6.1}$$

where $I \subset \mathbb{R}^+$, $S \subset \mathbb{R}$. Then the function ψ is a continuous wavelet for $L^2(\mathbb{R}^2)$ with respect to Λ'. These modified versions of the continuous shearlets depend on four variables a, s, t, ℓ and have frequency support:

$$\operatorname{supp} \hat{\psi}_{ast\ell} \subset \begin{cases} \{(\xi_1,\xi_2) : \xi_1 \in [-\frac{2}{a}, -\frac{1}{2a}],\ |\frac{\xi_2}{\xi_1} - s| \leq \sqrt{a}\}, & \text{if } \ell = -1, \\ \{(\xi_1,\xi_2) : \xi_1 \in [\frac{1}{2a}, \frac{2}{a}],\ |\frac{\xi_2}{\xi_1} - s| \leq \sqrt{a}\}, & \text{if } \ell = 1. \end{cases}$$

We remark that these modified versions of the continuous shearlets are in fact complex functions, whereas the shearlets we use throughout this paper are real functions.

Finally, there exists a natural way to construct continuous shearlets also in dimensions larger than 2. We refer to [18] for a discussion about the generalizations of shear matrices to higher dimensions.

Appendix A. Additional computations

Proof of Theorem 2.1. Suppose that (2.4) holds. Then, by applying Parseval and Plancherel formulas, for any $f \in L^2(\mathbb{R}^n)$,

$$\begin{aligned}
&\int_{\mathbb{R}^n} \int_G |\langle f, T_t\, D_M\, \psi\rangle|^2 \; d\lambda(M)\, dt \\
&= \int_{\mathbb{R}^n} \int_G \left| \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{\psi}(M^t \xi)}\, e^{2\pi i \xi t} d\xi \right|^2 |\det M|\, d\lambda(M)\, dt \\
&= \int_{\mathbb{R}^n} \int_G \left| \left(\hat{f}\, \overline{\hat{\psi}(M^t \cdot)} \right)^{\vee} (t) \right|^2 |\det M|\, d\lambda(M)\, dt \\
&= \int_G \int_{\mathbb{R}^n} \left| \left(\hat{f}\, \overline{\hat{\psi}(M^t \cdot)} \right)^{\vee} (t) \right|^2 dt\, |\det M|\, d\lambda(M) \\
&= \int_G \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2\, |\hat{\psi}(M^t \xi)|^2\, |\det M|\, d\xi\, d\lambda(M) \\
&= \int_G |\hat{f}(\xi)|^2\, \Delta(\psi)(\xi)\, d\xi = \|f\|^2.
\end{aligned}$$

Equation (2.3) follows from the above equality by polarization.

Conversely, suppose that

$$\int_{\mathbb{R}^n} \int_G |\langle f, T_t\, D_M\, \psi\rangle|^2 \; d\lambda(M)\, dt = \|f\|^2$$

for all $f \in L^2(\mathbb{R}^n)$. Let ξ_0 be a point of differentiability of $\Delta(\psi)(\xi)$ and let $\hat{f}(\xi) = |B(\xi_0, r)|^{-1/2}\, \chi_{B(\xi_0,r)}(\xi)$, where $B(\xi_0, r)$ is a ball centered at ξ_0 of radius r. By reversing the chain of equalities above we conclude that

$$\frac{1}{|B(\xi_0, r)|} \int_{B(\xi_0,r)} \Delta(\psi)(\xi)\, d\xi = 1$$

for all $r > 0$. Letting $r \to 0$, we obtain that $\Delta(\psi)(\xi_0) = 1$, and, since almost each $\xi \in \mathbb{R}^n$ is a point of differentiability, (2.4) holds. $\square$

The proof easily extends to functions $f \in L^2(V)^{\vee}$. In fact, it suffices to replace $\hat{f}$ with $\hat{f}\chi_V$ in the argument above.

Proof of Equality (3.7). Using Plancherel and Parseval formulas, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^2} |\langle f, T_t\, W\rangle|^2 \; dt &= \int_{\mathbb{R}^2} \left| \int_{\mathbb{R}^2} \hat{f}(\xi)\, \overline{\widehat{W}(\xi)}\, e^{2\pi i \xi t}\, d\xi \right|^2 dt \\
&= \int_{\mathbb{R}^2} \left| \left(\hat{f}\, \overline{\widehat{W}} \right)^{\vee} (t) \right|^2 dt \\
&= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2\, |\widehat{W}(\xi)|^2 \; d\xi. \qquad \text{(A.1)}
\end{aligned}$$

Using a similar computation,

$$\begin{aligned}
&\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 |\langle P_{C_1} f, \psi_{ast}\rangle|^2 \, \frac{da}{a^3}\, ds\, dt \\
&= \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 \left|\int_{\mathbb{R}^2} \hat f(\xi)\, \chi_{C_1}(\xi)\, \overline{\hat\psi(M_{as}^t \xi)}\, e^{2\pi i \xi t}\, d\xi\right|^2 \frac{da}{a^{3/2}}\, ds\, dt \\
&= \int_{-2}^{2}\int_0^1 \int_{\mathbb{R}^2} \left|\left(\hat f\, \chi_{C_1}\, \overline{\hat\psi(M_{as}^t \cdot)}\right)^\vee (t)\right|^2 dt\, \frac{da}{a^{3/2}}\, ds
\end{aligned}$$

$$= \int_{-2}^{2}\int_0^1 \int_{\mathbb{R}^2} |\hat f(\xi)|^2\, \chi_{C_1}(\xi)\, |\hat\psi(M_{as}^t \xi)|^2\, d\xi\, \frac{da}{a^{3/2}}\, ds. \tag{A.2}$$

As in the proof of Proposition 3.2, for $\xi \in C_1$, it follows that

$$\int_{-2}^{2}\int_0^1 |\hat\psi(M_{as}^t \xi)|^2 \frac{da}{a^{3/2}}\, ds = \int_0^1 |\hat\psi_1(a\xi_1)|^2 \frac{da}{a}.$$

Thus, using the last equality, (A.2) yields

$$\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 |\langle P_{C_1} f, \psi_{ast}\rangle|^2 \, \frac{da}{a^3}\, ds\, dt = \int_{\mathbb{R}^2} |\hat f(\xi)|^2 \chi_{C_1}(\xi) \int_0^1 |\hat\psi_1(a\xi_1)|^2 \frac{da}{a}\, d\xi. \tag{A.3}$$

Similarly, we can conclude that

$$\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 \left|\langle P_{C_2} f, \psi_{ast}^{(v)}\rangle\right|^2 \, \frac{da}{a^3}\, ds\, dt = \int_{\mathbb{R}^2} |\hat f(\xi)|^2 \chi_{C_2}(\xi) \int_0^1 |\hat\psi_1(a\xi_2)|^2 \frac{da}{a}\, d\xi. \tag{A.4}$$

Thus, combining (A.1), (A.3) and (A.4) and using (3.6), it follows that

$$\begin{aligned}
&\int_{\mathbb{R}^2} |\langle f, T_t W\rangle|^2 \, dt + \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 |\langle P_{C_1} f, \psi_{ast}\rangle|^2 \, \frac{da}{a^3}\, ds\, dt \\
&+ \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 \left|\langle P_{C_2} f, \psi_{ast}^{(v)}\rangle\right|^2 \, \frac{da}{a^3}\, ds\, dt \\
&= \int_{\mathbb{R}^2} |\hat f(\xi)|^2 \left(|\widehat W(\xi)|^2 + \chi_{C_1}(\xi)\int_0^1 |\hat\psi_1(a\xi_1)|^2 \frac{da}{a} + \chi_{C_2}(\xi)\int_0^1 |\hat\psi_1(a\xi_2)|^2 \frac{da}{a}\right) d\xi \\
&= \|f\|^2.
\end{aligned}$$

□

Acknowledgments

The research for this paper was performed while the first author was visiting the Departments of Mathematics at Washington University in St. Louis and the Georgia Institute of Technology. This author thanks these departments for their hospitality and support during these visits. We are also indebted to F. Bornemann, E. Candès, L. Demanet, K. Guo, W. Lim, G. Weiss, and E. Wilson for helpful discussions.

References

1. J. Bros and D. Iagolnitzer, *Support essentiel et structure analytique des distributions*, Séminaire Goulaouic-Lions-Schwartz, exp. no. 19 (1975-1976).

2. A. Calderón, *Intermediate spaces and interpolation. The complex method*, Stud. Math. **24** (1964), 113–190. MR0167830 (29:5097)
3. E. J. Candès and L. Demanet, *The curvelet representation of wave propagators is optimally sparse*, Comm. Pure Appl. Math. **58** (2005), 1472–1528. MR2165380 (2006f:35165)
4. E. J. Candès and D. L. Donoho, *Ridgelets: A key to higher–dimensional intermittency?*, Phil. Trans. Royal Soc. London A **357** (1999), 2495–2509. MR1721227 (2000g:42047)
5. E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266. MR2012649 (2004k:42052)
6. E. J. Candès and D. L. Donoho, *Continuous curvelet transform: I. Resolution of the wavefront set*, Appl. Comput. Harmon. Anal. **19** (2005), 162–197. MR2163077 (2006d:42058a)
7. E. J. Candès and D. L. Donoho, *Continuous curvelet transform: II. Discretization and frames*, Appl. Comput. Harmon. Anal. **19** (2005), 198–222. MR2163078 (2006d:42058b)
8. P. G. Casazza, *The art of frame theory*, Taiwanese J. Math. **4** (2000), 129–201. MR1757401 (2001f:42046)
9. O. Christensen, *An introduction to frames and Riesz bases*, Birkhäuser, Boston, 2003. MR1946982 (2003k:42001)
10. A. Córdoba and C. Fefferman, *Wave packets and Fourier integral operators*, Comm. Partial Diff. Eq. **3** (1978), 979–1005. MR507783 (80a:35117)
11. S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke, *The uncertainty principle associated with the continuous shearlet transform,* Int. J. Wavelets Multiresolut. Inf. Process. **6** (2008), 157–181.
12. S. Dahlke, G. Kutyniok, G. Steidl and G. Teschke, *Shearlet coorbit spaces and associated Banach frames*, preprint (2007).
13. G. Easley, D. Labate, and W. Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal. **25** (2008), 25–46. MR2419703
14. A. Grossmann, J. Morlet, and T. Paul, *Transforms associated to square integrable group representations I: General Results*, J. Math. Phys. **26** (1985), 2473–2479. MR803788 (86k:22013)
15. K. Guo, G. Kutyniok, and D. Labate, *Sparse multidimensional representations using anisotropic dilation and shear operators,* in: Wavelets and Splines, G. Chen and M. Lai (eds.), Nashboro Press, Nashville, TN (2006), 189–201. MR2233452 (2007c:42050)
16. K. Guo and D. Labate, *Optimally sparse multidimensional representation using shearlets*, SIAM J. Math. Anal., **39** (2007), 298–318. MR2318387
17. K. Guo, W. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations*, Electron. Res. Announc. Amer. Math. Soc. **10** (2004), 78–87. MR2075899 (2005e:42119)
18. K. Guo, W. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations and their MRA properties*, Appl. Comput. Harmon. Anal. **20** (2006), 220–236. MR2207836 (2006j:42056)
19. M. Holschneider, *Wavelets. Analysis tool*, Oxford University Press, Oxford, 1995. MR1367088 (97b:42051)
20. L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Springer-Verlag, Berlin, 2003. MR1996773
21. G. Kutyniok and T. Sauer, *Adaptive directional subdivision schemes and shearlet multiresolution analysis*, preprint, 2007.
22. D. Labate, W. Lim, G. Kutyniok, and G. Weiss, *Sparse multidimensional representation using shearlets*, Wavelets XI (San Diego, CA, 2005), 254–262, SPIE Proc. **5914**, SPIE, Bellingham, WA, 2005.
23. R. S. Laugesen, N. Weaver, G. Weiss, and E. Wilson, *A characterization of the higher dimensional groups associated with continuous wavelets*, J. Geom. Anal. **12** (2001), 89–102. MR1881293 (2002m:42042)
24. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998. MR1614527 (99m:94012)
25. Y. Meyer, *Wavelets and Operators*, Cambridge Stud. Adv. Math. vol. 37, Cambridge Univ. Press, Cambridge, UK, 1992. MR1228209 (94f:42001)
26. H. F. Smith, *A Hardy space for Fourier integral operators*, J. Geom. Anal. **8**, 629–653. MR1724210 (2001f:35459)
27. C. D. Sogge, *Fourier Integrals in Classical Analysis,* Cambridge University Press, Cambridge, 1993. MR1205579 (94c:35178)

28. E. M. Stein, *Harmonic Analysis: Real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, NJ, 1993. MR1232192 (95c:42002)
29. G. Weiss and E. Wilson, *The mathematical theory of wavelets*, Proceeding of the NATO–ASI Meeting. Harmonic Analysis 2000 – A Celebration. Kluwer Publisher, 2001. MR1858791 (2002h:42078)

Department of Statistics, Stanford University, Stanford, California 94305
E-mail address: kutyniok@stanford.edu

Department of Mathematics, North Carolina State University, Campus Box 8205, Raleigh, North Carolina 27695
E-mail address: dlabate@unity.ncsu.edu