

The Shearlet Transform and Lizorkin Spaces

F. Bartolucci and S. Pilipovic and N. Teofanov

Research Report No. 2020-44
June 2020

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

The Shearlet Transform and Lizorkin Spaces

Francesca Bartolucci ^{*} Stevan Pilipović [†] Nenad Teofanov [†]

Abstract

We prove a continuity result for the shearlet transform when restricted to the space of smooth and rapidly decreasing functions with all vanishing moments. We define the dual shearlet transform, called here the shearlet synthesis operator, and we prove its continuity on the space of smooth and rapidly decreasing functions over $\mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^\times$. Then, we use these continuity results to extend the shearlet transform to the space of Lizorkin distributions, and we prove its consistency with the classical definition for test functions. Our proofs are based on the strict relation between the shearlet, the wavelet and the affine Radon transforms proved in an earlier paper ([ACHA, 2019]) by one of the authors.

Key words. shearlet transform; wavelet transform; Radon transform; Ridgelet transform; Lizorkin spaces

1 Introduction

Among the large reservoir of directional multiscale representations which have been introduced over the years, the shearlet representation has gained considerable attention for its capability to resolve the wavefront set of distributions, providing both the location and the geometry of the singularity set of signals. Indeed, when we shift from one-dimensional to multidimensional signals, it is not just of interest to locate singularities in space but also to describe how they are distributed. This additional information is expressed by the notion of wavefront set introduced by Hörmander in [14]. In [16] the authors show that the decay rate of the shearlet coefficients $\mathcal{S}_\psi f(b, s, a)$ of a signal f with respect to suitable shearlets ψ characterizes the wavefront set of f . Precisely, they show that for any signal $f \in L^2(\mathbb{R}^2)$ the shearlet coefficients $\mathcal{S}_\psi f(b, s, a)$ exhibit fast asymptotic decay as $a \rightarrow 0$ except when the pair $(b, (\xi_1, \xi_2)) \in \mathbb{R}^2 \times \mathbb{R}^2$, with $\xi_2/\xi_1 = s$, belongs to the wavefront set of f . Later this result has been generalized in [10] where it is shown that the same result holds true under much weaker assumptions on the admissible vectors by means of a new approach based on an adaptation of the Radon transform to the shearlet structure, the affine Radon transform.

^{*}Department of Mathematics, ETH Zurich, Raemistrasse 101, 8092 Zurich, Switzerland (francesca.bartolucci@sam.math.ethz.ch).

[†]Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia (stevan.pilipovic@gmail.com, nenad.teofanov@dmi.uns.ac.rs).

On the other hand, while the classical wavelet transform is widely exploited in signal analysis for describing pointwise smoothness of univariate functions (we refer to [15, 18] as classical references), it has proved not flexible enough to capture the geometry of the singularity set when we shift from one-dimensional signals to multidimensional signals. We refer to e.g. [20] for a modification of the wavelet transform which overcomes these difficulties.

In some sense, shearlets behave for high-dimensional signals as wavelets do for one-dimensional signals and the link between these two transforms has been clarified in [2] where it is shown that the shearlet transform is the composition of the affine Radon transform with a one-dimensional wavelet transform, followed by a convolution operator with a scale-dependent filter.

There are at least two classical approaches to extend integral transforms to generalized function spaces. The coorbit space theory introduced by Feichtinger and Gröchenig in [6, 7] applies when the integral transform is the voice transform associated to a square-integrable representation of a locally compact group, and this is the case of the shearlet transform. We refer to [5] for an extension of the shearlet transform based on the coorbit space theory. The second way to proceed is the duality approach introduced by Schwartz in the 50's. A classical example is the extension of the Fourier transform to the space of tempered distributions. In this paper we extend the shearlet transform to distributions following the approach of Schwartz.

Our work arises from the lack of a complete distributional framework for the shearlet transform in the literature and from the link between the shearlet transform with the Radon and the wavelet transforms, whose distribution theory is deeply investigated and well known subject in applied mathematics. We refer respectively to [13, 21] and to [11, 12] for the extension of the wavelet transform and the Radon transform to various generalized function spaces via a duality approach. The Lizorkin space plays a crucial role in the development of a distributional framework for these two classical transforms and it turns out to be a natural domain for the shearlet transform too. We recall that the Lizorkin space $\mathcal{S}_0(\mathbb{R}^d)$ consists of smooth and rapidly decreasing functions with vanishing moments of any order. Moreover, in [19] the authors show that the domain of the ridgelet transform can be enlarged to its dual space $\mathcal{S}'_0(\mathbb{R}^2)$, known as the space of Lizorkin distributions. Their proofs widely exploit the intimate connection between the Radon, the ridgelet and the wavelet transforms, which also yields a relation formula between the shearlet transform and the ridgelet transform as we show in the Appendix in Proposition 11. This has in part inspired our work and we adapt several ideas of [19] to our context.

Our main results are continuity theorems for the shearlet transform and its dual transform, called the shearlet synthesis operator, on various test function spaces. Precisely, we prove that the shearlet transform $\mathcal{S}_\psi: \mathcal{S}_0(\mathbb{R}^2) \rightarrow \mathcal{S}(\mathbb{S})$ and its dual transform $\mathcal{S}_\psi^t: \mathcal{S}(\mathbb{S}) \rightarrow \mathcal{S}(\mathbb{R}^2)$ are continuous, where $\mathcal{S}(\mathbb{S})$ is a certain space of highly localized functions (see Subsection 2.1 for the definition of $\mathcal{S}(\mathbb{S})$). Our continuity theorems hold for suitable choice of the admissible vector ψ in the space $\mathcal{S}_0(\mathbb{R}^2)$ (see Section 3) and this is not a surprising condition. Indeed, as pointed out in wavelet analysis [18], shearlet analysis [10] and in the study of the Taylorlet transform [8], vanishing moments are crucial in order to measure the local regularity and to detect anisotropic structures of a signal. Then, we use these continuity results to extend the shearlet transform to the space of Lizorkin distributions following the approach in [19]. We show that the shearlet transform can be extended as a continuous map from $\mathcal{S}'_0(\mathbb{R}^2)$ into $\mathcal{S}'(\mathbb{S})$, where $\mathcal{S}'(\mathbb{S})$ is the space of distributions of slow growth on $\mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^\times$. Observe that many

important Schwartz distribution spaces, such as $\mathcal{E}'(\mathbb{R}^d)$, $\mathcal{O}'_C(\mathbb{R}^d)$, $L^p(\mathbb{R}^d)$ and $\mathcal{D}'_{L^1}(\mathbb{R}^d)$ are embedded into the the space of Lizorkin distributions $\mathcal{S}'_0(\mathbb{R}^d)$ (see e.g. [19]).

When considering possible applications of our approach, we notice that the rectified linear units (ReLUs), which are important examples of unbounded activation functions in the context of deep learning neural networks, belong to the space of Lizorkin distributions, see [23] for details.

The chapter is organized as follows. In Section 2 we introduce the spaces that occur in our analysis. Then, we recall the definition and the basic properties of the wavelet transform and the Radon transform in polar and affine coordinates. Section 3 is devoted to an introduction of the shearlet transform and to recall one of the main results in [2]. Moreover, in Theorem 6, we give a sketch of the proof of the continuity of the shearlet transform \mathcal{S}_ψ on $\mathcal{S}_0(\mathbb{R}^2)$. In Section 4 we introduce and study the shearlet synthesis operator. In particular, in Theorem 7 we prove its continuity on $\mathcal{S}(\mathbb{S})$. The importance of this dual transform follows by the fact that it can be used to define the extension of the shearlet transform to the space of Lizorkin distributions $\mathcal{S}'_0(\mathbb{R}^2)$ in a natural way, as we show in Section 5. We conclude our analysis with Theorem 4 which proves that our definition of the shearlet transform of distributions extends the ones considered so far, see e.g. [17, 10], and it is consistent with those for test functions (see Definition 4). Moreover, Theorem 4 shows that our duality approach is equivalent to the one based on the coorbit space theory presented in [5].

1.1 Notation

We briefly introduce the notation. We set $\mathbb{N} = \{0, 1, 2, \dots\}$, \mathbb{Z}_+ denotes the set of positive integers, $\mathbb{R}_+ = (0, +\infty)$, $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ and $\mathbb{H}^{d+1} = \mathbb{R}^d \times \mathbb{R}^\times$, $d \in \mathbb{Z}_+$. We also use the notation $\mathbb{H}^{(m_1, \dots, m_d, 1)} = \mathbb{R}^{m_1} \times \dots \times \mathbb{R}^{m_d} \times \mathbb{R}^\times$, $m_j \in \mathbb{Z}_+$, $j = 1, \dots, d$. When $x, y \in \mathbb{R}^d$ and $m \in \mathbb{N}^d$, $|x|$ denotes the Euclidean norm, $x \cdot y$ their scalar product, $\langle x \rangle = (1 + |x|^2)^{1/2}$, $xy = x_1y_1 + x_2y_2 + \dots + x_dy_d$, $x^m = x_1^{m_1} \dots x_d^{m_d}$ and $\partial^m = \partial_x^m = \partial_{x_1}^{m_1} \dots \partial_{x_d}^{m_d}$. We write also $\varphi^{(m)} = \partial^m \varphi$, $m \in \mathbb{N}^d$. By a slight abuse of notation, the length of a multi-index $m \in \mathbb{N}^d$ is denoted by $|m| = m_1 + \dots + m_d$ and the meaning of $|\cdot|$ shall be clear from the context. We write $A \lesssim B$ when $A \leq C \cdot B$ for some positive constant C .

For any $p \in [1, +\infty]$ we denote by $L^p(\mathbb{R}^d)$ the Banach space of functions $f: \mathbb{R}^d \rightarrow \mathbb{C}$ that are p -integrable with respect to the Lebesgue measure dx and, if $p = 2$, the corresponding scalar product and norm are $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. The Fourier transform is denoted by \mathcal{F} both on $L^2(\mathbb{R}^d)$ and on $L^1(\mathbb{R}^d)$, where it is defined by

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i \xi \cdot x} dx, \quad f \in L^1(\mathbb{R}^d).$$

If G is a locally compact group, we denote by $L^2(G)$ the Hilbert space of square-integrable functions with respect to a left Haar measure on G , and $C(G)$ denotes the space of continuous functions on G . If $A \in M_d(\mathbb{R})$, the vector space of square $d \times d$ matrices with real entries, tA denotes its transpose and we denote the (real) general linear group of size $d \times d$ by $\text{GL}(d, \mathbb{R})$. Finally, for every $b \in \mathbb{R}^d$, the translation operator acts on a function $f: \mathbb{R}^d \rightarrow \mathbb{C}$ as $T_b f(x) = f(x - b)$ and the dilation operator $D_a: L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)$ is defined by $D_a f(x) = |a|^{-\frac{1}{2}} f(x/a)$ for every $a \in \mathbb{R}^\times$.

The dual pairing between a test function space \mathcal{A} and its dual space of distributions \mathcal{A}' is denoted by $(\cdot, \cdot) = \mathcal{A}'(\cdot, \cdot)_{\mathcal{A}}$ and we provide all distribution spaces with the strong dual topologies.

The Schwartz space of rapidly decreasing smooth test functions is denoted by $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ denotes its dual space of tempered distributions. For the seminorms on $\mathcal{S}(\mathbb{R}^d)$, we make the choice

$$\rho_\nu(\varphi) = \sup_{x \in \mathbb{R}^d, |m| \leq \nu} \langle x \rangle^\nu |\partial^m \varphi(x)|,$$

for every $\nu \in \mathbb{N}$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

2 Preliminaries

In this section we first introduce the Lizorkin space of test functions, an important subspace of $\mathcal{S}(\mathbb{R}^d)$ which plays a crucial role in our analysis. Afterwards, we recall the definition and the main properties of the wavelet transform and the Radon transform in polar and affine coordinates in order to recall in Section 3 part of the results contained in [2].

2.1 The spaces

In this subsection we introduce the spaces that occur in this chapter and we state some auxiliary results which we widely exploit in the proofs of our main results (see Lemma 1 and 2 below). In particular, the Lizorkin space $\mathcal{S}_0(\mathbb{R}^d)$ will play a crucial role in our analysis. It consists of rapidly decreasing functions with vanishing moments of any order. Precisely,

$$\mathcal{S}_0(\mathbb{R}^d) = \{\varphi \in \mathcal{S}(\mathbb{R}^d) : \mu_m(\varphi) = 0, \forall m \in \mathbb{N}^d\},$$

where $\mu_m(\varphi) = \int_{\mathbb{R}^d} x^m \varphi(x) dx$, $m \in \mathbb{N}^d$. The Lizorkin space $\mathcal{S}_0(\mathbb{R}^d)$ is a closed subspace of $\mathcal{S}(\mathbb{R}^d)$ equipped with the relative topology inhered from $\mathcal{S}(\mathbb{R}^d)$ and its dual space of Lizorkin distributions $\mathcal{S}'_0(\mathbb{R}^d)$ is canonically isomorphic to the quotient of $\mathcal{S}'(\mathbb{R}^d)$ by the space of polynomials (cf. [13, 19]).

We are also interested in the Fourier Lizorkin space $\hat{\mathcal{S}}_0(\mathbb{R}^d)$ which consists of rapidly decreasing functions that vanish in zero together with all their partial derivatives, i.e.

$$\hat{\mathcal{S}}_0(\mathbb{R}^d) = \{\varphi \in \mathcal{S}(\mathbb{R}^d) : \partial^m \varphi(0) = 0, \forall m \in \mathbb{N}^d\},$$

which is a closed subspace of $\mathcal{S}(\mathbb{R}^d)$ too and we endow it with the relative topology inhered from $\mathcal{S}(\mathbb{R}^d)$.

We observe that, since $\mathcal{S}_0(\mathbb{R}^d)$ and $\hat{\mathcal{S}}_0(\mathbb{R}^d)$ are closed subspaces of the nuclear space $\mathcal{S}(\mathbb{R}^d)$, they are nuclear as well. We denote by $X \hat{\otimes} Y$ the topological tensor product space obtained as the completion of $X \otimes Y$ in the inductive tensor product topology ε or the projective tensor product topology π , see [24] for details. Then, we have the following result.

Lemma 1. *The spaces $\mathcal{S}_0(\mathbb{R}^d)$ and $\hat{\mathcal{S}}_0(\mathbb{R}^d)$ are closed under translations, dilations, differentiations and multiplications by a polynomial. Moreover, the Fourier transform is an isomorphism between $\mathcal{S}_0(\mathbb{R}^d)$ and $\hat{\mathcal{S}}_0(\mathbb{R}^d)$ and we have the following canonical isomorphisms:*

$$\begin{aligned} \mathcal{S}_0(\mathbb{R}^d) &\cong \mathcal{S}_0(\mathbb{R}^{d_1}) \hat{\otimes} \mathcal{S}_0(\mathbb{R}^{d_2}), \\ \hat{\mathcal{S}}_0(\mathbb{R}^d) &\cong \hat{\mathcal{S}}_0(\mathbb{R}^{d_1}) \hat{\otimes} \hat{\mathcal{S}}_0(\mathbb{R}^{d_2}), \end{aligned}$$

where $d = d_1 + d_2 \in \mathbb{Z}_+$, and $\hat{\otimes}$ denotes the completion with respect to the ε -topology or the π -topology.

Proof. The proof is based on classical arguments and we omit it (cf. [24, Theorem 51.6] for the canonical isomorphisms). \square

The next Lemma is a reformulation of [18, Theorem 6.2].

Lemma 2 ([1]). *Let $f \in \mathcal{S}_0(\mathbb{R}^d)$. Then, for any given $m \in \mathbb{N}^d$ there exists $g \in \mathcal{S}_0(\mathbb{R}^d)$ such that*

$$\mathcal{F}f(\xi) = \xi^m \mathcal{F}g(\xi), \quad \xi \in \mathbb{R}^d,$$

and vice versa.

Proof. We start proving the above statement for $d = 1$ and $m = 1$. Let $f \in \mathcal{S}_0(\mathbb{R})$ and consider

$$g(x) = \int_{-\infty}^x f(t)dt = - \int_x^{+\infty} f(t)dt,$$

where in the second equality we use the fact that $f \in \mathcal{S}_0(\mathbb{R})$. For every $k \in \mathbb{N}$ and $x > 0$

$$\begin{aligned} \langle x \rangle^k |g(x)| &= \left| \int_x^{+\infty} (1+x^2)^{\frac{k}{2}} f(t)dt \right| \leq \int_x^{+\infty} (1+t^2)^{\frac{k}{2}} |f(t)|dt \\ &\leq \rho_{2k+4}(f) \int_{-\infty}^{+\infty} (1+t^2)^{\frac{k}{2}} \frac{1}{(1+t^2)^{k+2}} dt < +\infty. \end{aligned}$$

Analogously, for $x < 0$ it holds

$$\begin{aligned} \langle x \rangle^k |g(x)| &= \left| \int_{-\infty}^x (1+x^2)^{\frac{k}{2}} f(t)dt \right| \leq \int_{-\infty}^x (1+t^2)^{\frac{k}{2}} |f(t)|dt \\ &\leq \rho_{2k+4}(f) \int_{-\infty}^{+\infty} (1+t^2)^{\frac{k}{2}} \frac{1}{(1+t^2)^{k+2}} dt < +\infty. \end{aligned}$$

Thus, g is a well defined function and $\sup_{x \in \mathbb{R}} \langle x \rangle^k |g(x)| < +\infty$ for every $k \in \mathbb{N}$. Moreover, $g'(x) = f(x)$, so that $\sup_{x \in \mathbb{R}} \langle x \rangle^k |g^{(l)}(x)| < +\infty$ for every $k \in \mathbb{N}$ and $l \geq 1$. Therefore, $g \in \mathcal{S}(\mathbb{R})$. Furthermore, for any $n \in \mathbb{N}$, we have that

$$\int_{-\infty}^{+\infty} x^n g(x)dx = - \int_{-\infty}^{+\infty} x^{n+1} g'(x)dx = - \int_{-\infty}^{+\infty} x^{n+1} f(x)dx = 0.$$

Hence, $g \in \mathcal{S}_0(\mathbb{R})$ and by the definition of g we have

$$\mathcal{F}f(\xi) = \mathcal{F}g'(\xi) = (2\pi i)\xi \mathcal{F}g(\xi), \quad \xi \in \mathbb{R}.$$

The opposite direction is obviously true since the space $\mathcal{S}_0(\mathbb{R})$ is closed under multiplication by a polynomial and this concludes the proof for $d = 1$ and $m = 1$. The analogous statement holds true for $m > 1$ by iterating the above proof m -times. The case $d > 1$ follows by analogous computations. \square

By Lemma 2, if $f \in \mathcal{S}_0(\mathbb{R}^2)$, then for any given $k, l \in \mathbb{N}$ there exists $g \in \mathcal{S}_0(\mathbb{R}^2)$ such that

$$\mathcal{F}f(\xi_1, \xi_2) = \xi_1^k \xi_2^l \mathcal{F}g(\xi_1, \xi_2), \quad (\xi_1, \xi_2) \in \mathbb{R}^2.$$

Moreover, it is worth observing that by Lemmas 1 and 2, $f \in \mathcal{S}_0(\mathbb{R}^d)$ if and only if it satisfies the directional vanishing moments:

$$\int_{\mathbb{R}} x_j^m f(x_1, x_2, \dots, x_d) dx_j = 0, \quad \forall m \in \mathbb{N}, \quad j = 1, \dots, d.$$

The space $\mathcal{S}(\mathbb{H}^{(d,d-1,1)})$ of highly localized functions (see also [13]) consists of the functions $\Phi \in C^\infty(\mathbb{H}^{(d,d-1,1)})$ such that the seminorms

$$\rho_{k_1, k_2, l, m}^{\alpha_1, \alpha_2, \beta, \gamma}(\Phi) = \sup_{(b, s, a) \in \mathbb{H}^{(d, d-1, 1)}} \langle b_1 \rangle^{k_1} \langle \tilde{b} \rangle^{k_2} \langle s \rangle^l \left(|a|^m + \frac{1}{|a|^m} \right) \left| \partial_a^\gamma \partial_s^\beta \partial_{\tilde{b}}^{\alpha_2} \partial_{b_1}^{\alpha_1} \Phi(b, s, a) \right|$$

are finite for all $k_1, m, \alpha_1, \gamma \in \mathbb{N}$, $k_2, l, \alpha_2, \beta \in \mathbb{N}^{d-1}$ and where $b = (b_1, \tilde{b}) \in \mathbb{R} \times \mathbb{R}^{d-1}$. In particular, when $d = 2$, we denote $\mathbb{S} := \mathbb{H}^{(2, 2-1, 1)} = \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^\times$ and $\mathcal{S}(\mathbb{S})$ consists of the functions $\Phi \in C^\infty(\mathbb{S})$ such that the seminorms

$$p_{k_1, k_2, l, m}^{\alpha_1, \alpha_2, \beta, \gamma}(\Phi) = \sup_{((b_1, b_2), s, a) \in \mathbb{S}} \langle b_1 \rangle^{k_1} \langle b_2 \rangle^{k_2} \langle s \rangle^l \left(|a|^m + \frac{1}{|a|^m} \right) \left| \partial_a^\gamma \partial_s^\beta \partial_{b_2}^{\alpha_2} \partial_{b_1}^{\alpha_1} \Phi((b_1, b_2), s, a) \right| \quad (1)$$

are finite for all $k_1, k_2, l, m, \alpha_1, \alpha_2, \beta, \gamma \in \mathbb{N}$. The topology of $\mathcal{S}(\mathbb{S})$ is defined by means of the seminorms (1). Its dual $\mathcal{S}'(\mathbb{S})$ will play a crucial role in the definition of the shearlet transform of Lizorkin distributions since it contains the range of this transform. We fix $d\mu(b, s, a) = |a|^{-3} db ds da$ as the standard measure on \mathbb{S} , where db , ds and da are the Lebesgue measures on \mathbb{R}^2 , \mathbb{R} and \mathbb{R}^\times , respectively. If F is a function of at most polynomial growth on \mathbb{S} , i.e., if there exist $C, \nu_1, \nu_2, \nu_3 > 0$ such that

$$|F(b, s, a)| \leq C \langle b \rangle^{\nu_1} \langle s \rangle^{\nu_2} \left(|a|^{\nu_3} + \frac{1}{|a|^{\nu_3}} \right), \quad (b, s, a) \in \mathbb{S},$$

then we identify F with an element of $\mathcal{S}'(\mathbb{S})$ by means of the equality

$$(F, \Phi) = \int_{\mathbb{R}^\times} \int_{\mathbb{R}} \int_{\mathbb{R}^2} F(b, s, a) \Phi(b, s, a) \frac{db ds da}{|a|^3}, \quad (2)$$

for every $\Phi \in \mathcal{S}(\mathbb{S})$.

2.2 The wavelet transform

The one-dimensional affine group \mathbb{W} is the semidirect product $\mathbb{R} \rtimes \mathbb{R}^\times$ with group operation

$$(b, a)(b', a') = (b + ab', aa')$$

and left Haar measure $|a|^{-2} db da$. It acts on $L^2(\mathbb{R})$ by means of the square-integrable representation

$$W_{b,a} f(x) = |a|^{-\frac{1}{2}} f\left(\frac{x-b}{a}\right),$$

or, equivalently, in the frequency domain

$$\mathcal{F}W_{b,a} f(\xi) = |a|^{\frac{1}{2}} e^{-2\pi i b \xi} \mathcal{F}f(a\xi). \quad (3)$$

The wavelet transform is then $\mathcal{W}_\psi f(b, a) = \langle f, W_{b,a} \psi \rangle$, which is a multiple of an isometry from $L^2(\mathbb{R})$ into $L^2(\mathbb{W}, |a|^{-2} db da)$ provided that $\psi \in L^2(\mathbb{R})$ satisfies the admissibility condition, namely the Calderón equation,

$$0 < \int_{\mathbb{R}} \frac{|\mathcal{F}\psi(\xi)|^2}{|\xi|} d\xi < +\infty \quad (4)$$

and, in such a case, ψ is called a one-dimensional wavelet. We refer to [13, 21] for the extension of the wavelet transform to Lizorkin distributions.

2.3 The affine Radon transform

The Radon transform of a signal f is a function on the affine projective space $\mathbb{P}^1 \times \mathbb{R} = \{\Gamma \mid \Gamma \text{ line of } \mathbb{R}^2\}$ whose value at a line is the integral of f along that line. It is usually defined by parametrizing the lines by pairs $(\theta, q) \in [-\pi, \pi) \times \mathbb{R}$ as

$$\Gamma_{\theta, q} = \{(x, y) \in \mathbb{R}^2 \mid \cos \theta x + \sin \theta y = q\}$$

and we refer to it as the polar Radon transform, see [11]. Precisely, for every $f \in L^1(\mathbb{R}^2)$, the polar Radon transform of f is the map $\mathcal{R}^{\text{pol}} f: [-\pi, \pi) \times \mathbb{R} \rightarrow \mathbb{C}$ defined by

$$\mathcal{R}^{\text{pol}} f(\theta, q) = \int_{\mathbb{R}} f(q \cos \theta - y \sin \theta, q \sin \theta + y \cos \theta) dy, \quad (5)$$

where the equality (5) holds for almost every $(\theta, q) \in [-\pi, \pi) \times \mathbb{R}$.

We can also label the normal vector to a line by affine coordinates, that is

$$\Gamma_{v, t} = \{(x, y) \in \mathbb{R}^2 \mid x + vy = t\},$$

where the correspondence is $v = \tan \theta$ and $t = q / \cos \theta$. Horizontal lines can not be represented by this parametrization, but they constitute a negligible set with respect to the natural measure on $\mathbb{P}^1 \times \mathbb{R}$. The affine Radon transform of any $f \in L^1(\mathbb{R}^2)$ is the function $\mathcal{R}^{\text{aff}} f: \mathbb{R}^2 \rightarrow \mathbb{C}$ defined by

$$\mathcal{R}^{\text{aff}} f(v, t) = \int_{\mathbb{R}} f(t - vy, y) dy, \quad \text{a.e. } (v, t) \in \mathbb{R}^2,$$

and it is related to the polar Radon transform by

$$\mathcal{R}^{\text{aff}} f(v, t) = \frac{1}{\sqrt{1+v^2}} \mathcal{R}^{\text{pol}} f(\arctan v, \frac{t}{\sqrt{1+v^2}}). \quad (6)$$

We refer to [2] for the proof.

The choice of the affine parametrization is particularly well-adapted to the mathematical structure of the shearlet transform (see also [10] and [2]).

The next result is a formulation of the Fourier slice theorem written for the affine Radon transform. The function f to which \mathcal{R}^{aff} is applied is taken in $L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$.

Proposition 3 ([2, Proposition 6]). *Define $\psi: \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \rightarrow \mathbb{R}^2$ by $\psi(v, \tau) = (\tau, \tau v)$. For every $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ there exists a negligible set $E \subseteq \mathbb{R}$ such that for all $v \notin E$ the function $\mathcal{R}^{\text{aff}} f(v, \cdot)$ is in $L^2(\mathbb{R})$ and satisfies*

$$\mathcal{R}^{\text{aff}} f(v, \cdot) = \mathcal{F}^{-1}[\mathcal{F} f \circ \psi(v, \cdot)]. \quad (7)$$

We refer to [2] for the proof of Proposition 3 and we refer to [11] as a classical reference for the Radon transform theory.

3 The Shearlet transform

In this section we introduce the shearlet transform. Then, we recall the intertwining result proved in [2], see Theorem 5 below, and we give an idea of how it can be exploited to derive continuity properties of the shearlet transform when restricted to the Lizorkin

space of test functions, see Theorem 6. We refer to [9] as a classical reference for the theory of group representations of locally compact groups and to [17] for a complete overview of the shearlet analysis.

The standard shearlet group is the semidirect product $G = \mathbb{R}^2 \rtimes K$, where $K = \{S_s A_a \in \text{GL}(2, \mathbb{R}) : s \in \mathbb{R}, a \in \mathbb{R}^\times\}$ with

$$S_s = \begin{bmatrix} 1 & -s \\ 0 & 1 \end{bmatrix}, \quad A_a = a \begin{bmatrix} 1 & 0 \\ 0 & |a|^{-1/2} \end{bmatrix}.$$

We can identify the element $S_s A_a$ with the pair (s, a) and write (b, s, a) for the elements in G . With this identification the product law amounts to

$$(b, s, a)(b', s', a') = (b + S_s A_a b', s + |a|^{1/2} s', aa').$$

A left Haar measure of G is

$$d\mu(b, s, a) = |a|^{-3} db ds da,$$

with db , ds and da the Lebesgue measures on \mathbb{R}^2 , \mathbb{R} and \mathbb{R}^\times , respectively. The group G acts on $L^2(\mathbb{R}^2)$ via the square-integrable representation

$$S_{b,s,a} f(x) = |a|^{-3/4} f(A_a^{-1} S_s^{-1}(x - b)),$$

or, equivalently, in the frequency domain

$$\mathcal{F} S_{b,s,a} f(\xi) = |a|^{3/4} e^{-2\pi i b \xi} \mathcal{F} f(A_a {}^t S_s \xi). \quad (8)$$

Definition 4. Fix $\psi \in L^2(\mathbb{R}^2)$. The shearlet transform associated to ψ is the map $\mathcal{S}_\psi : L^2(\mathbb{R}^2) \rightarrow C(G) \cap L^\infty(G, d\mu)$ defined by

$$\mathcal{S}_\psi f(b, s, a) = \langle f, S_{b,s,a} \psi \rangle = |a|^{-3/4} \int_{\mathbb{R}^2} f(x) \overline{\psi(A_a^{-1} S_s^{-1}(x - b))} dx.$$

It is well-known that the shearlet transform \mathcal{S}_ψ is a non-trivial multiple of an isometry from $L^2(\mathbb{R}^2)$ into $L^2(G)$ provided that $\psi \in L^2(\mathbb{R}^2)$ satisfies the admissibility condition

$$0 < C_\psi = \int_{\mathbb{R}^2} \frac{|\mathcal{F}\psi(\xi)|^2}{|\xi_1|^2} d\xi < +\infty, \quad (9)$$

where $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, or equivalently

$$\int_{\mathbb{R}^\times} \int_{\mathbb{R}} |\mathcal{F}\psi(A_a {}^t S_s \xi)|^2 ds \frac{da}{|a|^{3/2}} = C_\psi, \quad \text{for a.e. } \xi \in \mathbb{R}^2 / \{0\},$$

see e.g. [4]. Furthermore, in such case, we have the reconstruction formula

$$f = \frac{1}{C_\psi} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \mathcal{S}_\psi f(b, s, a) S_{b,s,a} \psi \frac{db ds da}{|a|^3}, \quad (10)$$

where the integral converges in the weak sense.

We now recall part of the results in [2]. We fix $\psi \in L^2(\mathbb{R}^2)$ of the form

$$\mathcal{F}\psi(\xi_1, \xi_2) = \mathcal{F}\psi_1(\xi_1) \mathcal{F}\psi_2 \left(\frac{\xi_2}{\xi_1} \right), \quad (\xi_1, \xi_2) \in \mathbb{R}^2, \xi_1 \neq 0, \quad (11)$$

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

$$0 < \int_{\mathbb{R}} \frac{|\mathcal{F}\psi_1(\tau)|^2}{|\tau|} d\tau < +\infty, \quad \int_{\mathbb{R}} |\tau|^2 |\mathcal{F}\psi_1(\tau)|^2 d\tau < +\infty \quad (12)$$

and $\psi_2 \in L^2(\mathbb{R})$. Then, ψ satisfies the admissible condition (9) and the function $\chi_1 \in L^2(\mathbb{R})$ defined by

$$\mathcal{F}\chi_1(\tau) = |\tau| \mathcal{F}\psi_1(\tau) \quad (13)$$

is a one-dimensional wavelet, i.e. it satisfies (4). We are now ready to state one of the central results in [2] which shows that the shearlet transform is the composition of the affine Radon transform with a one-dimensional wavelet, followed by a convolution with a scale-dependent filter.

Theorem 5 ([2, Corollary 12]). *For any $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ and $((b_1, b_2), s, a) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^\times$,*

$$\mathcal{S}_\psi f((b_1, b_2), s, a) = |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\text{aff}} f(v, \cdot))(b_1 + vb_2, a) \overline{\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)} dv, \quad (14)$$

where $\phi_2 = \mathcal{F}\psi_2$.

In (14) the wavelet transform is one-dimensional and acts on the variable t . We refer to [2] for the proof.

By direct computation, applying the Plancherel theorem, the Fourier slice theorem (7) and equation (3), we obtain the useful formula

$$\begin{aligned} \mathcal{S}_\psi f((b_1, b_2), s, a) &= |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\text{aff}} f(v, \cdot))(b_1 + vb_2, a) \overline{\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)} dv \\ &= |a|^{-\frac{1}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}\mathcal{R}^{\text{aff}} f(v, \cdot)(\tau) \overline{\mathcal{F}\chi_1(a\tau)} e^{2\pi i\tau(b_1 + vb_2)} d\tau \overline{\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)} dv \\ &= |a|^{-\frac{1}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}f(\tau, \tau v) \overline{\mathcal{F}\chi_1(a\tau)} e^{2\pi i\tau(b_1 + vb_2)} \overline{\phi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)} d\tau dv, \end{aligned} \quad (15)$$

for every $((b_1, b_2), s, a) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^\times$. Alternatively, we can obtain formula (15) as a direct consequence of equations (8) and (11) combined with the Plancherel theorem and a change of variable in affine coordinates

$$\begin{aligned} \mathcal{S}_\psi f(b, s, a) &= \langle \mathcal{F}f, \mathcal{F}S_{b,s,a}\psi \rangle \\ &= |a|^{\frac{3}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}f(\xi_1, \xi_2) e^{2\pi i(b_1\xi_1 + b_2\xi_2)} \overline{\mathcal{F}\psi(a\xi_1, a|a|^{-\frac{1}{2}}(\xi_2 - s\xi_1))} d\xi_1 d\xi_2 \\ &= |a|^{\frac{3}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}f(\tau, \tau v) e^{2\pi i\tau(b_1 + vb_2)} \overline{\mathcal{F}\psi(a\tau, a|a|^{-\frac{1}{2}}\tau(v-s))} |\tau| d\tau dv \\ &= |a|^{\frac{3}{4}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathcal{F}f(\tau, \tau v) e^{2\pi i\tau(b_1 + vb_2)} |\tau| \overline{\mathcal{F}\psi_1(a\tau)} \overline{\mathcal{F}\psi_2\left(\frac{v-s}{|a|^{\frac{1}{2}}}\right)} d\tau dv, \end{aligned}$$

which is exactly formula (15) since by definition $\mathcal{F}\chi_1 = |\cdot| \mathcal{F}\psi_1$ and $\phi_2 = \mathcal{F}\psi_2$.

We are now ready to state our first main result. From now on, everytime we consider an admissible vector ψ , we assume that it is of the form (11) with $\chi_1 \in \mathcal{S}_0(\mathbb{R})$ and $\phi_2 \in \mathcal{S}(\mathbb{R})$. We observe that if $\chi_1 \in \mathcal{S}_0(\mathbb{R})$, then ψ_1 defined by (13)

as $\mathcal{F}\psi_1(\tau) = |\tau|^{-1}\mathcal{F}\chi_1(\tau)$ satisfies (12). Furthermore, under these assumptions, $\mathcal{F}\psi$ extends to a function belonging to the Fourier Lizorkin space $\hat{\mathcal{S}}_0(\mathbb{R}^2)$ and, with slight abuse of notation, ψ denotes both the admissible vector defined by (11) and its Schwartz extension over \mathbb{R}^2 .

Theorem 6 ([1]). *The shearlet transform \mathcal{S}_ψ is a continuous mapping from $\mathcal{S}_0(\mathbb{R}^2)$ into $\mathcal{S}(\mathbb{S})$.*

Proof. We give a sketch of the proof and we refer to [1] for the details. The goal is to prove that for every $f \in \mathcal{S}_0(\mathbb{R}^2)$, given $k_1, k_2, l, m, \alpha_1, \alpha_2, \beta, \gamma \in \mathbb{N}$, there exist $\nu \in \mathbb{N}$ such that

$$\rho_{k_1, k_2, l, m}^{\alpha_1, \alpha_2, \beta, \gamma}(\mathcal{S}_\psi f) \lesssim \rho_\nu(f).$$

The first step in the proof is to show that, without loss of generality, we can assume $\alpha_1 = \alpha_2 = \beta = \gamma = k_1 = k_2 = l = 0$. Then, it is enough to prove that for every $m \in \mathbb{N}$

$$\rho_{0, 0, 0, m}^{0, 0, 0, 0}(\mathcal{S}_\psi f) \lesssim \rho_\nu(f),$$

for some $\nu \in \mathbb{N}$. Our approach mimics the one followed in [19] to prove the continuity of the ridgelet transform on the Lizorkin space of test functions. Formula (15) and Lemma 2 play a crucial role throughout all the proof. In particular, equation (15) has the same role of formula (9) in [19]. \square

4 The shearlet synthesis operator

In this section we introduce the shearlet synthesis operator and we prove its continuity on the space $\mathcal{S}(\mathbb{S})$ of highly localized functions, see Theorem 7 below.

The reconstruction formula (10) suggests to define a linear operator \mathcal{S}_ψ^t which maps functions over $\mathbb{S} = \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^\times$ to functions over the Euclidean plane \mathbb{R}^2 . Given $\psi \in \mathcal{S}(\mathbb{R}^2)$, we define the shearlet synthesis operator \mathcal{S}_ψ^t by

$$\mathcal{S}_\psi^t \Phi(x) = \int_{\mathbb{R}^\times} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \Phi(b, s, a) S_{b, s, a} \psi(x) \frac{db ds da}{|a|^3}, \quad x \in \mathbb{R}^2, \quad (16)$$

for any function Φ for which the integral converges. For example, the integral in (16) is absolutely convergent if $\Phi \in \mathcal{S}(\mathbb{S})$. Furthermore, if $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ and $\Phi \in \mathcal{S}(\mathbb{S})$, then by Fubini theorem we have that

$$\begin{aligned} \int_{\mathbb{R}^2} f(x) \mathcal{S}_\psi^t \Phi(x) dx &= \int_{\mathbb{R}^2} f(x) \int_{\mathbb{R}^\times} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \Phi(b, s, a) \overline{S_{b, s, a} \psi(x)} \frac{db ds da}{|a|^3} dx \\ &= \int_{\mathbb{R}^\times} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \Phi(b, s, a) \int_{\mathbb{R}^2} f(x) \overline{S_{b, s, a} \psi(x)} dx \frac{db ds da}{|a|^3} \\ &= \int_{\mathbb{R}^\times} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \mathcal{S}_\psi f(b, s, a) \Phi(b, s, a) \frac{db ds da}{|a|^3}. \end{aligned} \quad (17)$$

In Theorem 7 we will show that the shearlet synthesis operator \mathcal{S}_ψ^t is a continuous operator from $\mathcal{S}(\mathbb{S})$ into $\mathcal{S}(\mathbb{R}^2)$. Then, since $L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ naturally embeds into $\mathcal{S}'(\mathbb{R}^2)$, by the identification (2), we may write (17) as

$$(f, \mathcal{S}_\psi^t \Phi) = (\mathcal{S}_\psi f, \Phi).$$

This duality relation will motivate our definition of the distributional shearlet transform in Section 5.

We recall that we consider admissible vectors ψ of the form (11) with χ_1 defined by (13) in $\mathcal{S}_0(\mathbb{R})$ and $\phi_2 = \mathcal{F}\psi_2 \in \mathcal{S}(\mathbb{R})$.

Theorem 7 ([1]). *The shearlet synthesis operator \mathcal{S}_ψ^t is a bounded operator from $\mathcal{S}(\mathbb{S})$ into $\mathcal{S}_0(\mathbb{R}^2)$.*

Proof. We start proving the continuity. We need to show that for every $\Phi \in \mathcal{S}(\mathbb{S})$ and $\nu \in \mathbb{N}$, there exist $k_1, k_2, l, m, \alpha_1, \alpha_2, \beta, \gamma \in \mathbb{N}$ and a positive constant C such that

$$\rho_\nu(\mathcal{S}_\psi^t \Phi) \leq C \rho_{k_1, k_2, l, m}^{\alpha_1, \alpha_2, \beta, \gamma}(\Phi).$$

We will use the fact that the families $\hat{\rho}_\nu(\chi) = \rho_\nu(\mathcal{F}\chi)$ and $\hat{\rho}_{k_1, k_2, l, m}^{\alpha_1, \alpha_2, \beta, \gamma}(\Phi) = \rho_{k_1, k_2, l, m}^{\alpha_1, \alpha_2, \beta, \gamma}(\mathcal{F}\Phi)$, where $\mathcal{F}\Phi$ denotes the Fourier transform of Φ with respect to the variable b , are bases of seminorms for the topologies of $\mathcal{S}_0(\mathbb{R})$ and $\mathcal{S}(\mathbb{S})$, respectively (cf. [19]). Furthermore, by Plancherel theorem, Fubini theorem, equation (8) and by the expression of the admissible vector, we obtain the following formula for the shearlet synthesis operator

$$\begin{aligned} & \mathcal{S}_\psi^t \Phi(x) \\ &= \int_{\mathbb{R} \times \mathbb{R}} \int_{\mathbb{R}} |a|^{\frac{3}{4}} \int_{\mathbb{R}^2} \mathcal{F}\Phi(\xi, s, a) e^{2\pi i x \cdot \xi} \overline{\mathcal{F}\psi(-a\xi_1, a|a|^{-\frac{1}{2}}(-\xi_2 + s\xi_1))} d\xi_1 d\xi_2 \frac{ds da}{|a|^3} \\ &= \int_{\mathbb{R}^2} e^{2\pi i x \cdot \xi} \int_{\mathbb{R} \times \mathbb{R}} |a|^{\frac{3}{4}} \mathcal{F}\Phi(\xi, s, a) \overline{\mathcal{F}\psi(-a\xi_1, a|a|^{-\frac{1}{2}}(-\xi_2 + s\xi_1))} \frac{ds da}{|a|^3} d\xi_1 d\xi_2 \\ &= \int_{\mathbb{R}^2} e^{2\pi i x \cdot \xi} \int_{\mathbb{R} \times \mathbb{R}} |a|^{\frac{3}{4}} \mathcal{F}\Phi(\xi, s, a) \overline{\mathcal{F}\psi_1(-a\xi_1) \mathcal{F}\psi_2\left(\frac{\xi_2/\xi_1 - s}{|a|^{\frac{1}{2}}}\right)} \frac{ds da}{|a|^3} d\xi_1 d\xi_2 \\ &= \int_{\mathbb{R}^2} e^{2\pi i x \cdot \xi} \int_{\mathbb{R} \times \mathbb{R}} |a|^{-\frac{1}{4}} \mathcal{F}\Phi(\xi, s, a) \frac{\overline{\mathcal{F}\chi_1(-a\xi_1)}}{|\xi_1|} \overline{\phi_2\left(\frac{\xi_2/\xi_1 - s}{|a|^{\frac{1}{2}}}\right)} \frac{ds da}{|a|^3} d\xi_1 d\xi_2. \quad (18) \end{aligned}$$

By formula (18), for any given $\alpha \in \mathbb{Z}_+$ we have that

$$\begin{aligned} & |\partial_{x_1}^\alpha (\mathcal{S}_\psi^t \Phi)(x_1, x_2)| \\ & \lesssim \int_{\mathbb{R}^2} \int_{\mathbb{R} \times \mathbb{R}} \int_{\mathbb{R}} |a|^{-\frac{1}{4}} |\mathcal{F}\Phi(\xi, s, a)| \frac{|\mathcal{F}\chi_1(-a\xi_1)|}{|\xi_1|^{1-\alpha}} \left| \phi_2\left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}}\right) \right| \frac{ds da}{|a|^3} d\xi_1 d\xi_2 \\ & \lesssim \int_{\mathbb{R}^2} \int_{|a| < \epsilon} \int_{\mathbb{R}} \frac{(1 + |\xi|^2)^{N/2}}{(1 + |\xi|^2)^{N/2}} (1 + s^2) |a|^{-\frac{13}{4}} |\xi_1|^{\alpha-1} |\mathcal{F}\Phi(\xi, s, a)| |\mathcal{F}\chi_1(-a\xi_1)| \\ & \quad \left| \phi_2\left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}}\right) \right| \frac{ds da}{(1 + s^2)} d\xi_1 d\xi_2 \\ & + \int_{\mathbb{R}^2} \int_{|a| > \epsilon} \int_{\mathbb{R}} \frac{(1 + |\xi|^2)^{N/2}}{(1 + |\xi|^2)^{N/2}} (1 + s^2) |a|^{-\frac{13}{4}} |\xi_1|^{\alpha-1} |\mathcal{F}\Phi(\xi, s, a)| |\mathcal{F}\chi_1(-a\xi_1)| \\ & \quad \left| \phi_2\left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}}\right) \right| \frac{ds da}{(1 + s^2)} d\xi_1 d\xi_2 \\ & \lesssim \rho_{N+\alpha-1, N, 2, \frac{13}{4}}^{0,0,0,0}(\mathcal{F}\Phi) + \rho_{N+\alpha-1, N, 2, 0}^{0,0,0,0}(\mathcal{F}\Phi), \end{aligned}$$

where $N \in \mathbb{N}$, $N > 2$ and $\epsilon > 0$. The terms of the form $|\partial_{x_2}^\beta (\mathcal{S}_\psi^t \Phi)(x_1, x_2)|$, $\beta \in \mathbb{Z}_+$, can be estimated in a similar fashion.

Next we consider multiplications by x_1^k , $k \in \mathbb{Z}_+$. By formula (18), we have that

$$\begin{aligned}
& |x_1^k (S_\psi^t \Phi)(x_1, x_2)| \\
&= \left| \int_{\mathbb{R}^2} x_1^k e^{2\pi i x \cdot \xi} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-1/4} \mathcal{F}\Phi(\xi, s, a) |\xi_1|^{-1} \times \right. \\
&\quad \left. \times \mathcal{F}\chi_1(-a\xi_1) \phi_2 \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right) \frac{ds da}{|a|^3} d\xi_1 d\xi_2 \right| \\
&= \left| \int_{\mathbb{R}^2} (2\pi i)^{-k} e^{2\pi i x \cdot \xi} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-1/4} \partial_{\xi_1}^k [\mathcal{F}\Phi(\xi, s, a) |\xi_1|^{-1} \times \right. \\
&\quad \left. \times \mathcal{F}\chi_1(-a\xi_1) \phi_2 \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right) \right] \frac{ds da}{|a|^3} d\xi_1 d\xi_2 \right| \\
&\lesssim \int_{\mathbb{R}^2} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-1/4} |\partial_{\xi_1}^k [\mathcal{F}\Phi(\xi, s, a) |\xi_1|^{-1} \times \\
&\quad \times \mathcal{F}\chi_1(-a\xi_1) \phi_2 \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right)]| \frac{ds da}{|a|^3} d\xi_1 d\xi_2,
\end{aligned}$$

which is less than or equal to a finite sum of addends of the form

$$\begin{aligned}
& \int_{\mathbb{R}^2} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-\frac{1}{4}+k_3-\frac{k_4}{2}} |\partial_{\xi_1}^{k_1} \mathcal{F}\Phi(\xi, s, a)| |\xi_1|^{-k_2-k_4} |(\mathcal{F}\chi_1)^{(k_3)}(-a\xi_1)| \times \\
& \times |\xi_2|^{k_4} \left| \phi_2^{(k_4)} \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right) \right| \frac{ds da}{|a|^3} d\xi_1 d\xi_2 \\
&= \int_{\mathbb{R}^2} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-\frac{13}{4}+k_3-\frac{k_4}{2}} (1+|\xi|^2)^{N/2} (1+s^2) |\partial_{\xi_1}^{k_1} \mathcal{F}\Phi(\xi, s, a)| |\xi_1|^{-k_2-k_4} \times \\
& \times |(\mathcal{F}\chi_1)^{(k_3)}(-a\xi_1)| |\xi_2|^{k_4} \left| \phi_2^{(k_4)} \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right) \right| \frac{ds da}{(1+s^2)(1+|\xi|^2)^{N/2}} d\xi_1 d\xi_2,
\end{aligned}$$

where $k_1, k_2, k_3, k_4 \in \mathbb{N}$ are less than k .

Since $\mathcal{S}_0(\mathbb{R})$ is closed under multiplications by a polynomial, by Lemma 2 it follows that for any given $k_3, m \in \mathbb{N}$ there exists $g \in \mathcal{S}_0(\mathbb{R})$ such that

$$(\mathcal{F}\chi_1)^{(k_3)}(-a\xi_1) = -a^m \xi_1^m \mathcal{F}g(-a\xi_1), \quad \xi_1 \in \mathbb{R}, a \in \mathbb{R}^\times,$$

and we can continue the chain of inequalities with terms of the form

$$\begin{aligned}
& \int_{\mathbb{R}^2} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-\frac{13}{4}+k_3-\frac{k_4}{2}+m} (1+|\xi|^2)^{N/2} (1+s^2) |\partial_{\xi_1}^{k_1} \mathcal{F}\Phi(\xi, s, a)| |\xi_1|^{-k_2-k_4+m} \\
& |\mathcal{F}g(-a\xi_1)| |\xi_2|^{k_4} \left| \phi_2^{(k_4)} \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right) \right| \frac{ds da}{(1+s^2)(1+|\xi|^2)^{N/2}} d\xi_1 d\xi_2.
\end{aligned}$$

Finally, choosing $N \in \mathbb{N}$, $N > 2$, $m \geq k_2 + k_4$ and splitting the integral over \mathbb{R}^\times into integrals over $|a| < \epsilon$ and $|a| > \epsilon$, with $\epsilon > 0$, we obtain that

$$\begin{aligned}
& \int_{\mathbb{R}^2} \int_{\mathbb{R}^\times} \int_{\mathbb{R}} |a|^{-\frac{13}{4}+k_3-\frac{k_4}{2}+m} (1+|\xi|^2)^{N/2} (1+s^2) |\partial_{\xi_1}^{k_1} \mathcal{F}\Phi(\xi, s, a)| |\xi_1|^{-k_2-k_4+m} \\
& |\mathcal{F}g(-a\xi_1)| |\xi_2|^{k_4} \left| \phi_2^{(k_4)} \left(\frac{\xi_2/\xi_1 - s}{|a|^{1/2}} \right) \right| \frac{ds da}{(1+s^2)(1+|\xi|^2)^{N/2}} d\xi_1 d\xi_2 \\
& \lesssim [\rho_{N+m-k_2-k_4, N+k_4, 2, |\frac{13}{4}-k_3+\frac{k_4}{2}-m|}^{k_1, 0, 0, 0} (\mathcal{F}\Phi) + \rho_{N+m-k_2-k_4, N+k_4, 2, |-k_3+\frac{k_4}{2}-m|}^{k_1, 0, 0, 0} (\mathcal{F}\Phi)],
\end{aligned}$$

which is dominated by a single seminorm. We can treat $|x_2^k(S_\psi^t\Phi)(x_1, x_2)|$, $k \in \mathbb{Z}_+$, in the same manner and we conclude that the shearlet synthesis \mathcal{S}_ψ^t is a continuous map from $\mathcal{S}(\mathbb{S})$ into $\mathcal{S}(\mathbb{R}^2)$. Finally, it remains to prove that $\mathcal{S}_\psi^t\Phi \in \mathcal{S}_0(\mathbb{R}^2)$. The idea is to prove the equivalent condition

$$\lim_{\xi \rightarrow 0} \frac{\mathcal{F}\mathcal{S}_\psi^t\Phi(\xi)}{|\xi|^k} = 0,$$

for every $k \in \mathbb{N}$, see [13, Lemma 6.0.4]. We refer to [1] for the details. \square

5 The shearlet transform on $\mathcal{S}'_0(\mathbb{R}^2)$

In this last section, we extend the definition of the shearlet transform to the space of Lizorkin distributions and we show that our definition extends the ones introduced so far. In particular, we prove its consistency with the classical definition for test functions.

We recall that we consider admissible vectors ψ of the form (11) with χ_1 defined by (13) in $\mathcal{S}_0(\mathbb{R})$ and $\phi_2 = \mathcal{F}\psi_2 \in \mathcal{S}(\mathbb{R})$.

Definition 8 ([1]). We define the shearlet transform of $f \in \mathcal{S}'_0(\mathbb{R}^2)$ with respect to ψ as follows

$$(\mathcal{S}_\psi f, \Phi) = (f, \mathcal{S}_\psi^t\Phi), \quad \Phi \in \mathcal{S}(\mathbb{S}).$$

The consistency of Definition 8 is guaranteed by Theorem 7. Furthermore, it follows straightforwardly that the shearlet transform of $f \in \mathcal{S}'_0(\mathbb{R}^2)$ is a well defined distribution in $\mathcal{S}'(\mathbb{S})$.

Proposition 9 ([1]). *The shearlet transform \mathcal{S}_ψ given by Definition 8 is a continuous and linear map from $\mathcal{S}'_0(\mathbb{R}^2)$ into $\mathcal{S}'(\mathbb{S})$.*

The next theorem shows that Definition 8 is in fact consistent with the definition for test functions (Definition 4) and that it generalizes the extension considered in [16, 10] where the shearlet transform of a tempered distribution f with respect to an admissible vector $\psi \in \mathcal{S}(\mathbb{R}^2)$ is given by the function

$$\mathcal{S}_\psi f(b, s, a) = \mathcal{S}'(\mathbb{R}^2)(f, \mathcal{S}_{b,s,a}\psi)_{\mathcal{S}(\mathbb{R}^2)},$$

for every $(b, s, a) \in \mathbb{S}$. Following the coorbit space approach, given a suitable test function space usually denoted by $\mathcal{H}_{1,w}$, where w is a weight function, and its anti-dual $\mathcal{H}_{1,w}^\sim$, the extended shearlet transform of $f \in \mathcal{H}_{1,w}^\sim$ with respect to $\psi \in \mathcal{H}_{1,w}$ is defined by

$$\mathcal{S}_\psi f(b, s, a) = \mathcal{H}_{1,w}^\sim(f, \mathcal{S}_{b,s,a}\psi)_{\mathcal{H}_{1,w}},$$

for every $(b, s, a) \in \mathbb{S}$, [5]. Theorem 10 shows the equivalence of our duality approach with the coorbit space one. Precisely, Theorem 10 states that the the shearlet transform of any Lizorkin distribution is given by the function defined as

$$(b, s, a) \mapsto \mathcal{S}'_0(\mathbb{R}^2)(f, \mathcal{S}_{b,s,a}\psi)_{\mathcal{S}_0(\mathbb{R}^2)},$$

for every $(b, s, a) \in \mathbb{S}$.

Theorem 10 ([1]). *Let $f \in \mathcal{S}'_0(\mathbb{R}^2)$. The shearlet transform of f is given by the function*

$$(b, s, a) \mapsto \mathcal{S}'_0(\mathbb{R}^2)(f, S_{b,s,a}\psi)_{\mathcal{S}'_0(\mathbb{R}^2)},$$

that is,

$$(\mathcal{S}_\psi f, \Phi) = \int_{\mathbb{S}} (f, S_{b,s,a}\psi) \Phi(b, s, a) d\mu(b, s, a), \quad \Phi \in \mathcal{S}(\mathbb{S}).$$

Proof. Consider $f \in \mathcal{S}'_0(\mathbb{R}^2)$. Since the space of Lizorkin distributions $\mathcal{S}'_0(\mathbb{R}^2)$ is canonically isomorphic to the quotient of $\mathcal{S}'(\mathbb{R}^2)$ by the space of polynomials, by Schwartz' structural theorem [22, Théorém VI], we can write $f = g^{(\alpha)} + p$, where g is a continuous slowly growing function, $\alpha \in \mathbb{N}^2$ and p is a polynomial. Then, for any $\Phi \in \mathcal{S}(\mathbb{S})$

$$\begin{aligned} (g^{(\alpha)}, \mathcal{S}_\psi^t \Phi) &= (-1)^{|\alpha|} (g, \mathcal{S}_\psi^t \Phi^{(\alpha)}) = (-1)^{|\alpha|} \int_{\mathbb{R}^2} g(x) \mathcal{S}_\psi^t \Phi^{(\alpha)}(x) dx \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^2} g(x) \int_{\mathbb{S}} \Phi(b, s, a) (S_{b,s,a}\psi)^{(\alpha)}(x) d\mu(b, s, a) dx \\ &= \int_{\mathbb{S}} \Phi(b, s, a) (-1)^{|\alpha|} \int_{\mathbb{R}^2} g(x) (S_{b,s,a}\psi)^{(\alpha)}(x) dx d\mu(b, s, a) \\ &= \int_{\mathbb{S}} \Phi(b, s, a) (-1)^{|\alpha|} (g, (S_{b,s,a}\psi)^{(\alpha)}) d\mu(b, s, a) \\ &= \int_{\mathbb{S}} \Phi(b, s, a) (g^{(\alpha)}, S_{b,s,a}\psi) d\mu(b, s, a). \end{aligned}$$

Analogously, we have that

$$(p, \mathcal{S}_\psi^t \Phi) = \int_{\mathbb{S}} \Phi(b, s, a) (p, S_{b,s,a}\psi) d\mu(b, s, a).$$

Therefore, we obtain

$$\begin{aligned} (f, \mathcal{S}_\psi^t \Phi) &= (g^{(\alpha)} + p, \mathcal{S}_\psi^t \Phi) = \int_{\mathbb{S}} \Phi(b, s, a) (g^{(\alpha)} + p, S_{b,s,a}\psi) d\mu(b, s, a) \\ &= \int_{\mathbb{S}} \Phi(b, s, a) (f, S_{b,s,a}\psi) d\mu(b, s, a), \end{aligned}$$

which concludes the proof. \square

Appendix

As mentioned in the introduction, both the ridgelet transform and the shearlet transform are related to the wavelet and the Radon transforms and, as we now show, they are related as well. The Appendix is devoted to prove this connection, which has in part inspired our work. We start briefly recalling the ridgelet transform and we refer to [3] as a classical reference.

We fix $\psi \in \mathcal{S}(\mathbb{R})$ and we define for every $(\theta, b, a) \in [-\pi, \pi) \times \mathbb{R} \times \mathbb{R}_+$ the function $R_{\theta,b,a}\psi$ as

$$R_{\theta,b,a}\psi(x) = \frac{1}{a} \psi\left(\frac{x \cdot n(\theta) - b}{a}\right), \quad x \in \mathbb{R}^2,$$

where $n(\theta) = (\cos \theta, \sin \theta)$. Then, the ridgelet transform of $f \in L^1(\mathbb{R}^2)$ with respect to ψ is given by

$$\mathcal{R}_\psi f(\theta, b, a) = \int_{\mathbb{R}^2} f(x) \overline{R_{\theta, b, a} \psi(x)} dx, \quad (\theta, b, a) \in [-\pi, \pi) \times \mathbb{R} \times \mathbb{R}_+.$$

The ridgelet transform is related to the wavelet transform and the polar Radon transform by the following formula

$$\mathcal{R}_\psi f(\theta, b, a) = \mathcal{W}_\psi(\mathcal{R}^{\text{pol}} f(\theta, \cdot))(b, a), \quad (19)$$

for every $(\theta, b, a) \in [-\pi, \pi) \times \mathbb{R} \times \mathbb{R}_+$ and where the wavelet transform is one-dimensional and acts on the variable q . By equations (19) and (6), we obtain a relation formula between the ridgelet and the shearlet transform. We consider an admissible vectors ψ of the form (11) satisfying conditions (12) and with χ_1 defined by (13) belonging to $\mathcal{S}(\mathbb{R})$.

Proposition 11. *For any $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ and $((b_1, b_2), s, a) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}_+$,*

$$\begin{aligned} & \mathcal{S}_\psi f((b_1, b_2), s, a) \\ &= |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \frac{1}{\sqrt[4]{1+v^2}} \mathcal{R}_{\chi_1} f \left(\arctan v, \frac{b_1 + vb_2}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}} \right) \overline{\phi_2 \left(\frac{v-s}{|a|^{1/2}} \right)} dv. \end{aligned}$$

Proof. By Theorem 5, for any $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ and $((b_1, b_2), s, a) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}_+$, the shearlet transform has the following expression

$$\mathcal{S}_\psi f((b_1, b_2), s, a) = |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\text{aff}} f(v, \cdot))(b_1 + vb_2, a) \overline{\phi_2 \left(\frac{v-s}{|a|^{1/2}} \right)} dv. \quad (20)$$

By equation (6), for every $v \in \mathbb{R}$, $(b_1, b_2) \in \mathbb{R}^2$ and $a \in \mathbb{R}_+$, we compute

$$\begin{aligned} & \mathcal{W}_{\chi_1}(\mathcal{R}^{\text{aff}} f(v, \cdot))(b_1 + vb_2, a) \\ &= \frac{1}{\sqrt{1+v^2}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\text{pol}} f(\arctan v, \frac{\cdot}{\sqrt{1+v^2}}))(b_1 + vb_2, a) \\ &= \frac{1}{\sqrt[4]{1+v^2}} \mathcal{W}_{\chi_1}(D_{\sqrt{1+v^2}} \mathcal{R}^{\text{pol}} f(\arctan v, \cdot))(b_1 + vb_2, a) \\ &= \frac{1}{\sqrt[4]{1+v^2}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\text{pol}} f(\arctan v, \cdot)) \left(\frac{b_1 + vb_2}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}} \right) \\ &= \frac{1}{\sqrt[4]{1+v^2}} \mathcal{R}_{\chi_1} f \left(\arctan v, \frac{b_1 + vb_2}{\sqrt{1+v^2}}, \frac{a}{\sqrt{1+v^2}} \right). \end{aligned} \quad (21)$$

Replacing formula (21) in (20) we obtain the desired relation. \square

Acknowledgement

F. Bartolucci is part of the Computational Harmonic Analysis & Machine Learning unit of the Machine Learning Genoa Center (MalGa). S. Pilipović and N. Teofanov were supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through Project 174024.

References

- [1] F. BARTOLUCCI. Radon transforms: Unitarization, Inversion and Wavefront sets [Ph.D. thesis]. <http://hdl.handle.net/11567/997903>.
- [2] F. Bartolucci, F. De Mari, E. De Vito, F. Odone. The Radon transform intertwines wavelets and shearlets. *Applied and Computational Harmonic Analysis* **47** (2019), no. 3, 822–847.
- [3] E. J. CANDÈS, D. L. DONOHO. Ridgelets: A key to higher-dimensional intermittency? *Philos. Trans. R. Soc. A* **357** (1999), no. 1760, 2495–2509.
- [4] S. DAHLKE, G. KUTYNIOK, P. MAASS, C. SAGIV, H. STARK, AND G. TESCHKE, The uncertainty principle associated with the continuous shearlet transform, *International Journal of Wavelets, Multiresolution and Information Processing* **6** (2008), no. 2, 157–181.
- [5] S. DAHLKE, G. KUTYNIOK, G. STEIDL, AND G. TESCHKE, Shearlet coorbit spaces and associated Banach spaces, *Applied and Computational Harmonic Analysis* **27** (2009), no. 2, 195–214.
- [6] H.G. FEICHTINGER, K. GRÖCHENIG, Banach spaces related to integrable group representations and their atomic decompositions, I, *Journal of Functional Analysis*, **86** (1989), no. 2, 307–340.
- [7] H.G. FEICHTINGER, K. GRÖCHENIG, Banach Spaces Related to Integrable Group Representations and Their Atomic Decompositions. Part II, *Monatshefte für Mathematik*, **108** (1989), no. 2-3, 129–148.
- [8] T. FINK, U. KAHLER. A Space-Based Method for the Generation of a Schwartz Function with Infinitely Many Vanishing Moments of Higher Order with Applications in Image Processing, *Complex Analysis and Operator Theory* **13** (2019), no. 3, 985–1010.
- [9] G. B. FOLLAND, *A course in abstract harmonic analysis*, Textbooks in Mathematics, CRC Press, Boca Raton, FL, 2nd ed., 2016.
- [10] P. GROHS, Continuous shearlet frames and resolution of the wavefront set, *Monatshefte für Mathematik* **164** (2011), no. 4, 393–426.
- [11] S. HELGASON, *The Radon transform*, vol. 5 of Progress in Mathematics, Birkhäuser Boston, Inc., Boston, MA, 2nd ed., 1999.
- [12] A. HERTLE. Continuity of the Radon transform and its inverse on Euclidean spaces. *Math Z.* **184** (1983), 165–192.
- [13] M. HOLSCHNEIDER. *Wavelets. An analysis tool*. The Clarendon Press, Oxford University Press, New York (1995).
- [14] L. HÖRMANDER. *The analysis of linear partial differential operators. I*. Grundlehren der Mathematischen Wissenschaften 256, Springer-Verlag, Berlin (1983).
- [15] S. JAFFARD. Exposants de Hölder en des points donnés et coefficients d’ondelettes. *Comptes rendus de l’Académie des Sciences Series I Mathematics* **308** (1989), no. 4, 79–81.
- [16] G. KUTYNIOK, D. LABATE. Resolution of the wavefront set using continuous shearlets. *Trans. Amer. Math. Soc.* **361** (2009), no. 5, 2719–2754.

- [17] G. KUTYNIOK, D. LABATE. *Shearlets*. Appl. Numer. Harmon. Anal. Birkhäuser/Springer, New York (2012).
- [18] S. MALLAT. *A wavelet tour of signal processing, The sparse way*. Elsevier, Academic Press, Amsterdam (2009).
- [19] S. KOSTADINOVA, S. PILIPOVIĆ, K. SANEVA, J. VINDAS. The ridgelet transform of distributions. *Integral Transforms Spec. Funct.* **25** (2014), no. 5, 344–358.
- [20] S. PILIPOVIĆ, M. VULETIĆ. Characterization of wave front sets by wavelet transforms. *Tohoku Math. J.* **58** (2006), no. 3, 369–391.
- [21] S. PILIPOVIĆ, D. RAKIĆ, N. TEOFANOV, J. VINDAS. The wavelet transforms in Gelfand-Shilov spaces, *Collectanea Mathematica* **67** (2016), no. 3, 443–460.
- [22] L. SCHWARTZ. *Théorie des distributions. Tome II*. Actualités Sci. Ind., no. 1122 Publ. Inst. Math. Univ. Strasbourg 10. Hermann & Cie., Paris (1951).
- [23] S. SONODA, N. MURATA. Neural network with unbounded activation functions is universal approximator. *Applied and Computational Harmonic Analysis* **43** (2017), no. 2, 233–268.
- [24] F. TRÈVES. *Topological vector spaces, distributions and kernels*. Academic Press, New York-London (1967).