RADON TRANSFORM INTERTWINES SHEARLETS AND WAVELETS

F. BARTOLUCCI, F. DE MARI, E. DE VITO, AND F. ODONE

ABSTRACT. We prove that the unitary affine Radon transform intertwines the quasi-regular representation of a class of semidirect products, built by shearlet dilation groups and translations, and the tensor product of a standard wavelet representation with a wavelet-like representation. This yields a formula for shearlet coefficients that involves only integral transforms applied to the affine Radon transform of the signal, thereby opening new perspectives in the inversion of the Radon transform.

1. INTRODUCTION

The use of wavelets in signal analysis and computer vision has proved almost optimal for one-dimensional signals in many ways, and the mathematics behind classical wavelets has reached a high degree of elaboration. In higher dimensions, however, the picture is less clear and this partially explains the huge class of representations that has been introduced over the years to handle high dimensional problems, such as directional wavelets [1], ridgelets [2], curvelets [3], wavelets with composite dilations [4], contourlets [5], shearlets [6], reproducing groups of the symplectic group [7], Gabor ridge functions [7] and mocklets [8] – to name a few.

Among them, shearlets stand out because of their ability to efficiently capture anisotropic features, to provide optimal sparse representations, to detect singularities and to be stable against noise, see [9] for an overview and a complete list of references. From the purely mathematical perspective, their construction is based on the well-established theory of square-integrable representations [10], just as wavelets are, and because of this many powerful mathematical tools are available. As far as applications are concerned, their effectiveness has been tested primarily in image processing, where many efficient algorithms have been designed using them (see [9, 11] and the website `http://www.shearlab.org/` for further details and references).

Thus, in some sense, shearlets behave for high-dimensional signals as wavelets do for 1D-signal, and it is therefore natural to try to understand if the many strong connections are a consequence of some general mathematical principle.

The purpose of this paper is to address this issue, and to give a partial answer, showing that the link between the shearlet transform and wavelets is the unitary Radon transform in affine coordinates, because it actually intertwines the shearlet representation with a tensor product of two wavelet representations. This fact can be exploited to show that by carefully choosing the mother shearlet it is possible to obtain the classical shearlet coefficients as a sequence of operations performed on the Radon transform of the signal, namely a one-dimensional wavelet transform, with respect to the "sliding" coordinate

---

Key words and phrases. shearlets; wavelets; Radon transform.

that parametrizes all the hyperplanes parallel to a given one (for the two dimensional case see Fig. 1), followed by a convolution operator with a scale-dependent filter in the variables of the hyperplane. As the shearlet transform admits an inversion formula, it is in principle possible to invert the Radon transform of a given signal by means of it and the aforementioned operations.

For two-dimensional signals, our results, which have been announced in [12], can be described as we now explain. In order to formulate them precisely, we recall the definition of the three main ingredients, namely wavelets, shearlets and the Radon transform. The wavelet group is $\mathbb{R} \rtimes \mathbb{R}^\times$ with law $(b, a)(b', a') = (b + ab', aa')$. The square integrable wavelet representation W acts on $L^2(\mathbb{R})$ by

$$W_{b,a}\psi(x) = |a|^{-1/2}\psi\big(\frac{x - b}{a}\big)$$

and the wavelet transform, defined by $\mathcal{W}_\psi f(b, a) = \langle f, \psi_{b,a}\rangle$, is a multiple of an isometry from $L^2(\mathbb{R})$ to $L^2(\mathbb{R} \rtimes \mathbb{R}^\times)$ provided that ψ satisfies the Calderón condition, see (22) below. Next, denote by $\mathbb{S}$ the (parabolic) shearlet group, namely $\mathbb{R}^2 \rtimes (\mathbb{R} \rtimes \mathbb{R}^\times)$ with multiplication

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

where

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

and where the vectors are understood as column vectors. The group $\mathbb{S}$ acts on $L^2(\mathbb{R}^2)$ via the shearlet representation, namely

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

The shearlet transform is then $\mathcal{S}_\psi f(b, s, a) = \langle f, S_{b,s,a}\psi\rangle$, and is a multiple of an isometry provided that an admissibility condition on ψ is satisfied [13, 14], see (17) below. Finally, the Radon transform in affine coordinates of a signal $f \in L^1(\mathbb{R}^2)$ is the function $\mathcal{R}^{\text{aff}}f : \mathbb{R}^2 \to \mathbb{C}$ defined by

$$\mathcal{R}^{\text{aff}}f(v, t) = \int_\mathbb{R} f(t - vy, y)\,\mathrm{d}y, \qquad (v, t) \in \mathbb{R}^2.$$

An important fact is that it is possible to define a version of $\mathcal{R}^{\text{aff}}$ as a unitary map on $L^2(\mathbb{R}^2)$. First, it is necessary to compose it with the Riesz-type operator $\mathcal{I}$ that we now describe. Its natural domain is the dense subspace of $L^2(\mathbb{R}^2)$

$$\mathcal{D} = \Big\{g \in L^2(\mathbb{R}^2) : \int_{\mathbb{R}^2} |\xi_2|\,|\widehat{g}(\xi_1, \xi_2)|^2 \mathrm{d}\xi_1 \mathrm{d}\xi_2 < +\infty\Big\},$$

where $\widehat{g}$ denotes the Fourier transform of g. The densely defined, self-adjoint unbounded operator $\mathcal{I} : \mathcal{D} \to L^2(\mathbb{R}^2)$ is defined by

$$\widehat{(\mathcal{I}g)}(\xi_1, \xi_2) = |\xi_2|^{\frac{1}{2}}\,\widehat{g}(\xi_1, \xi_2), \qquad (\xi_1, \xi_2) \in \mathbb{R}^2,$$

i.e. a Fourier multiplier in the second variable. It is not hard to show that for all f in the dense subspace of $L^2(\mathbb{R}^2)$

$$\mathcal{A} = \Big\{f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^2) : \int_{\mathbb{R}^2} \frac{|\widehat{f}(\xi_1, \xi_2)|^2}{|\xi_1|}\mathrm{d}\xi_1 \mathrm{d}\xi_2 < +\infty\Big\},$$

the Radon transform $\mathcal{R}^{\mathrm{aff}} f$ belongs to $\mathcal{D}$ and that the map $f \longmapsto \mathcal{I}\mathcal{R}^{\mathrm{aff}} f$ from $\mathcal{A}$ to $L^2(\mathbb{R}^2)$ extends to a unitary map, denoted by $\mathcal{Q}$, from $L^2(\mathbb{R}^2)$ onto itself.

In the two dimensional case our main formula reads now

$$\mathcal{Q}\, S_{b,s,a} f = \left(W_{s,|a|^{1/2}} \otimes \mathrm{I}\right) W_{(1,\mathbf{v})\cdot b,a}\, \mathcal{Q} f \tag{1}$$

where the meaning of the dummy variable $\mathbf{v}$ is

$$W_{(1,\mathbf{v})\cdot b,a} g(v,t) = |a|^{-\frac{1}{2}}\, g\left(v, \frac{t - (1,v) \cdot b}{a}\right).$$

Our second most important result is the formula

$$\mathcal{S}_\psi f(x,y,s,a) = |a|^{-\frac{3}{4}} \int_{\mathbb{R}} \mathcal{W}_{\chi_1}\left(\mathcal{R}^{\mathrm{aff}} f(v,\bullet)\right)(x+vy,a) \overline{\phi_2\left(\frac{v-s}{|a|^{1/2}}\right)}\, \mathrm{d}v, \tag{2}$$

provided that $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and ψ is of the form

$$\widehat{\psi}(\xi_1,\xi_2) = \widehat{\psi_1}(\xi_1)\widehat{\psi_2}\left(\frac{\xi_2}{\xi_1}\right).$$

The 1D-wavelet χ_1 and the 1D-filter ϕ_2 are related to the shearlet admissible vector ψ by the following relations

$$\widehat{\chi_1}(\xi_1) = |\xi_1|\widehat{\psi_1}(\xi_1) \tag{3}$$

$$\phi_2(\xi_2/\xi_1) = \widehat{\psi_2}(\xi_2/\xi_1). \tag{4}$$

The first equality shows that $2\pi\chi_1 = H\psi_1'$ is the Hilbert transform H of the weak derivative of ψ_1.

Equation (2) shows that for any signal f in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ the shearlet coefficients can be computed by means of three *classical* trasforms. Indeed, in order to obtain $\mathcal{S}_\psi f(x,y,s,a)$ one can:

a) compute the Radon transform $\mathcal{R}^{\mathrm{aff}} f(v,t)$ of the original signal f;
b) apply the wavelet transform with respect to the variable t

$$G(v,b,a) = \mathcal{W}_{\chi_1}\left(\mathcal{R}^{\mathrm{aff}} f(v,\cdot)\right)(b,a), \tag{5}$$

    where χ_1 is given by (3);
c) convolve the result with the scale-dependent filter

$$\Phi_a(v) = \overline{\phi_2\left(-\frac{v}{|a|^{1/2}}\right)},$$

    where ϕ_2 is given by (4) and the convolution is computed with respect to the variable v, that is

$$\mathcal{S}_\psi f(x,y,s,a) = \left(G(\bullet, x + \bullet\, y, a) * \Phi_a\right)(s).$$

Finally, since S is a square-integrable representation, there is a reconstruction formula, namely

$$f = \int_{\mathbb{S}} \mathcal{S}_\psi f(x,y,s,a)\, S_{x,y,s,a}\psi\, \frac{\mathrm{d}x\mathrm{d}y\mathrm{d}s\mathrm{d}a}{|a|^3}, \tag{6}$$

where the integral converges in the weak sense. Note that $\mathcal{S}_\psi f$ depends on f only through its Radon transform $\mathcal{R}^{\mathrm{aff}} f$, see (2). The above equation allows to reconstruct an unknown signal f from its Radon transform by computing the shearlet coefficients by means of (6).

The fact that the Radon transform does play a prominent role in this circle of ideas is not new. Indeed, it is known that ridgelets are constructed via wavelet analysis in the Radon domain [15], Gabor frames are defined as the directionally-sensitive Radon transforms [7], discrete shearlet frames are used to invert the Radon transform [16] and the Radon transform is at the root of the proof that shearlets are able to detect the wavefront set of a 2D signal [17].

Our contribution is to clarify this relation from the point of view of non-commutative harmonic analysis. We are actually able to prove a rather general result, Theorem 3, which generalizes (1) and holds for the class of groups that were introduced by Führ [18, 19] and that are known as *shearlet dilation groups*.

The paper is organized as follows. In Section 2 we present in full detail all the various ingredients, namely the groups, the representations, the Radon transform and the unitary extensions that need to be defined. In Section 3 we state and prove the main results.

2. Preliminaries

2.1. **Notation.** We briefly introduce the notation. We set $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$. The Euclidean norm of a vector $v \in \mathbb{R}^d$ is denoted by $|v|$ and its scalar product with $w \in \mathbb{R}^d$ by $v \cdot w$. For any $p \in [1, +\infty]$ we denote by $L^p(\mathbb{R}^d)$ the Banach space of functions $f : \mathbb{R}^d \to \mathbb{C}$, which are p-integrable with respect to the Lebesgue measure $\mathrm{d}x$ and, if $p = 2$, the corresponding scalar product and norm are $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. The Fourier trasform is denoted by $\mathcal{F}$ both on $L^2(\mathbb{R}^d)$ and on $L^1(\mathbb{R}^d)$, where it is defined by

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

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. We denote the (real) general linear group of size $d \times d$ by $\mathrm{GL}(d, \mathbb{R})$ and by $\mathrm{T}(d, \mathbb{R})$ the closed subgroup of unipotent upper triangular matrices.

If H is a closed subgroup of $\mathrm{GL}(d, \mathbb{R})$, the semidirect product $G = \mathbb{R}^d \rtimes H$ is the product $\mathbb{R}^d \times H$ with group operation

$$(b_1, h_1)(b_2, h_2) = (b_1 + h_1[b_2], h_1 h_2),$$

where $b_1, b_2 \in \mathbb{R}^d$, $h_1, h_2 \in H$ and where $h[b]$ is the natural linear action of the matrix h on the column vector b.

2.2. **Shearlet dilation groups.** In this section we introduce the groups in which we are interested. This family includes the groups introduced by Führ in [18, 19], and called generalized shearlet dilation groups for the purpose of generalizing the standard shearlet group introduced in [6, 20].

Definition 1. A *shearlet dilation group* $H < \mathrm{GL}(d, \mathbb{R})$ is a subgroup of the form $H = SD$, where

(i) S is a Lie subgroup of $\mathrm{T}(d,\mathbb{R})$ consisting of matrices of the form

$$\begin{bmatrix} 1 & -{}^{t}s \\ 0 & B(s) \end{bmatrix}$$

with $s \in \mathbb{R}^{d-1}$ and $B : \mathbb{R}^{d-1} \to \mathrm{T}(d-1,\mathbb{R})$ a smooth map;

(ii) D is the one-parameter subgroup of $\mathrm{GL}(d,\mathbb{R})$ consisting of the diagonal matrices

$$a \operatorname{diag}(1, |a|^{\lambda_1}, \ldots, |a|^{\lambda_{d-1}}) = a \begin{bmatrix} 1 & 0 \\ 0 & \Lambda(a) \end{bmatrix} \tag{7}$$

as a ranges in $\mathbb{R}^{\times}$. Here $(\lambda_1, \ldots, \lambda_{d-1})$ is a fixed vector in $\mathbb{R}^{d-1}$.

The group S is called the shearing sugroup of H and D is called the diagonal complement or scaling subgroup of H.

Several observations are in order. First of all, if one requires the shearing subgroup S to be Abelian, then one obtains the class introduced by Führ, with a slightly more general definition. This has inspired Definition 1.

Since the map B is continuous, S is automatically connected, and hence by Theorem 3.6.2 in [21], it is closed and simply connected. By construction the elements of H are of the form

$$h_{s,a} = h_{s,1}h_{0,a} = a \begin{bmatrix} 1 & -{}^{t}s\Lambda(a) \\ 0 & B(s)\Lambda(a) \end{bmatrix}. \tag{8}$$

Furthermore, since the diagonal matrices of $\mathrm{GL}(d,\mathbb{R})$ normalize $\mathrm{T}(d,\mathbb{R})$, then H is the semidirect product of S and D.

Finally, the assumption that S is a subgroup normalized by D forces the maps B and Λ to satisfy some equalities. Indeed, since

$$\begin{bmatrix} 1 & -{}^{t}u \\ 0 & B(u) \end{bmatrix} \begin{bmatrix} 1 & -{}^{t}v \\ 0 & B(v) \end{bmatrix} = \begin{bmatrix} 1 & -{}^{t}(v + {}^{t}B(v)u) \\ 0 & B(u)B(v) \end{bmatrix},$$

then S is a group if and only if

$$B(0) = \mathrm{I}_{d-1} \tag{9}$$

$$B(u)B(v) = B(v + {}^{t}B(v)u) \tag{10}$$

$$B(u)^{-1} = B(-{}^{t}B(u)^{-1}u) \tag{11}$$

for every $u, v \in \mathbb{R}^{d-1}$. Since

$$\begin{bmatrix} 1 & 0 \\ 0 & \Lambda(a) \end{bmatrix} \begin{bmatrix} 1 & -{}^{t}s \\ 0 & B(s) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \Lambda(a)^{-1} \end{bmatrix} = \begin{bmatrix} 1 & {}^{t}(\Lambda(a)^{-1}s) \\ 0 & \Lambda(a)B(s)\Lambda(a)^{-1} \end{bmatrix}$$

the compatibility of D with S is equivalent to asking for the following condition to hold for all $a \neq 0$ and all $s \in \mathbb{R}^{d-1}$:

$$\Lambda(a)B(s)\Lambda(a)^{-1} = B(\Lambda(a)^{-1}s). \tag{12}$$

It follows that H is diffeomorphic as a manifold to $\mathbb{R}^{d-1} \times \mathbb{R}^{\times}$, so that we can identify the element $h_{s,a}$ with the pair (s, a). With this identification the product law amounts to

$$(s, a)(s', a') = \left(\Lambda(a)^{-1}s' + {}^{t}B(\Lambda(a)^{-1}s')s, aa' \right). \tag{13}$$

We stress that, in general, S is not isomorphic as a Lie group to the additive Abelian group $\mathbb{R}^{d-1}$, unless S is the standard shearlet group introduced in [20], see the examples below.

Remark. It should be clear that a slightly larger class would be obtained by allowing for diagonal matrices of the form

$$\operatorname{sign}(a) \operatorname{diag}(|a|^{\mu_0}, |a|^{\mu_1}, \ldots, |a|^{\mu_{d-1}}), \qquad a \in \mathbb{R}^{\times}.$$

The case $\mu_0 = 0$, however, is uninteresting because any shearlet dilation group corresponding to this choice never admits admissible vectors [19, 22]. But then a simple change of variables permits to assume $\mu_0 = 1$, as we did, and to set $\lambda_j = \mu_j - 1$.

Remark. In [19] the authors introduce the notion of shearlet dilation group by means of structural properties and then prove that in the case when S is Abelian they can be parametrized as in Definition 1.

We now give three examples. If S is Abelian a full characterization is provided in [19], see also [22] for a connection with a suitable class of subgroups of the symplectic group.

Example (The standard shearlet group). A possible choice for B is the map $B(s) = \mathrm{I}_{d-1}$, which satisfies all the above properties. In this case, $s \mapsto h_{s,1}$ defines a group isomorphism between $\mathbb{R}^{d-1}$ and the Abelian group S.

Clearly, any choice of the weights $\lambda_1, \ldots, \lambda_{d-1}$ is compatible with (12). In particular, if we choose as D the group of matrices

$$A_a = a \begin{bmatrix} 1 & 0 \\ 0 & |a|^{\gamma-1} \mathrm{I}_{d-1} \end{bmatrix} \qquad \Longleftrightarrow \qquad \Lambda(a) = |a|^{\gamma-1} \mathrm{I}_{d-1} \qquad a \in \mathbb{R}^{\times},$$

where $\gamma \in \mathbb{R}$ is a fixed parameter, then we obtain the d-dimensional shearlet group, usually denoted $\mathbb{S}^{\gamma}$, and, often, the parameter γ is chosen to be $1/d$ [20, 23].

Example (The Toeplitz shearlet group). Another important example arises when $B(s)$ is the Toeplitz matrix

$$B(s) = T(\hat{s}) = \begin{bmatrix} 1 & -s_1 & -s_2 & \ldots & -s_{d-2} \\ 0 & 1 & -s_1 & -s_2 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & -s_1 \\ 0 & \ldots & \ldots & 0 & 1 \end{bmatrix}, \tag{14}$$

where $\hat{s} = {}^{t}(s_1, \ldots, s_{d-2})$. It is easy to see that $T(\hat{u})T(\hat{v}) = T(\hat{u}\sharp\hat{v})$ where

$$(\hat{u}\sharp\hat{v})_i := u_i + v_i + \sum_{j+k=i} v_j u_k, \quad i = 1, \ldots d-2$$

and that consequently all the equalities in (10) hold. This case corresponds to Toeplitz shearlet groups (see [24]).

Not all dilation matrices as in (7) are compatible with (12). In [19] it is shown that

$$\lambda_k = k\lambda_1, \qquad k = 2, \ldots, d-1 \tag{15}$$

for any fixed λ_1.

Example (A non-Abelian shearlet dilation group). The matrices

$$g(u_1, u_2, u_3) = \begin{bmatrix} 1 & -u_1 & -u_2 & -u_3 \\ 0 & 1 & -u_1 & -u_2 - \frac{1}{2}u_1^2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

as $u = (u_1, u_2, u_3)$ ranges in $\mathbb{R}^3$ give rise to a non-Abelian shearlet group S. Indeed, it is easily checked that

$$g(u_1, u_2, u_3)g(v_1, v_2, v_3) = g(u_1 + v_1, u_2 + v_2 - u_1 v_1, u_3 + v_3 - u_1(v_2 + \frac{1}{2}v_1^2)),$$

a product which is not Abelian in the third coordinate. Evidently,

$$B(u) = \begin{bmatrix} 1 & -u_1 & -u_2 - \frac{1}{2}u_1^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is a smooth function of u. The group S is isomorphic to the standard Heisenberg group, as is most clearly seen at the level of Lie algebra. Indeed, the Lie algebra of S is given by the matrices

$$X(q, p, t) = \begin{bmatrix} 0 & q & p & t \\ 0 & 0 & q & p \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

because $X^3(q, p, t) = 0$ and hence

$$\exp(X(q, p, t)) = I_4 + X(q, p, t) + \frac{1}{2}X^2(q, p, t)$$

$$= \begin{bmatrix} 1 & q & p + \frac{1}{2}q^2 & t + \frac{1}{2}qp \\ 0 & 1 & q & p \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= g(-q, -(p + \frac{1}{2}q^2), -(t + \frac{1}{2}qp)).$$

Further,

$$[X(q, p, t), X(q', p', t')] = X(0, 0, qp' - pq')$$

exhibits the Lie algebra of S as the three dimensional Heisenberg Lie algebra. A straightforward calculation shows that for any choice of $\lambda \in \mathbb{R}$ the diagonal matrices

$$\Lambda(a) = \begin{bmatrix} |a|^\lambda & & \\ & |a|^{2\lambda} & \\ & & |a|^{3\lambda} \end{bmatrix}$$

normalize $B(u)$ because $\Lambda(a)B(u)\Lambda(a)^{-1} = B(\Lambda(a)^{-1}u)$. Conversely, these are easily seen to be the only rank-one dilations that normalize the matrices $B(u)$. In conclusion, the group D consisting of the matrices

$$a \begin{bmatrix} 1 & \\ & \Lambda(a) \end{bmatrix}$$

together with S give rise to the non-Abelian shearlet dilation group $H = SD$. It is worth observing that the dilations in D are not the standard dilations of the Heisenberg

group. Indeed, the Lie algebra of D consists of the diagonal matrices $A_\lambda(\tau) = \mathrm{diag}(\tau, (\lambda + 1)\tau, (2\lambda + 1)\tau, (3\lambda + 1)\tau)$ and

$$[A_\lambda(\tau), X(q, p, t)] = X(-\lambda\tau q, -2\lambda\tau p, -3\lambda\tau t)$$

shows that these homogeneous dilations are not the standard dilations of the Heisenberg Lie algebra (see [25], p. 620).

2.3. The shearlet representation and admissible vectors.
From now on we fix a group $G = \mathbb{R}^d \rtimes H$ where H is a shearlet group as in Definition 1 and we parametrize its elements as (b, s, a). By (13) we get that a left Haar measure of H is

$$\mathrm{d}h = |a|^{\lambda_D - 1}\mathrm{d}s\mathrm{d}a$$

where $\lambda_D = \lambda_1 + \ldots + \lambda_{d-1}$ and $\mathrm{d}s$, $\mathrm{d}a$ are the Lebesgue measures of $\mathbb{R}^{d-1}$ and $\mathbb{R}^\times$. As a consequence, a left Haar measure on G is

$$\mathrm{d}g = \mathrm{d}b\,\frac{\mathrm{d}h}{|\det h_{s,a}|} = |a|^{-(d+1)}\mathrm{d}b\mathrm{d}s\mathrm{d}a$$

where $\mathrm{d}b$ is the Lebesgue measure on $\mathbb{R}^d$ and the last equality holds true since

$$|\det h_{s,a}| = |a|^{d + \lambda_D}.$$

The quasi-regular representation of G on $L^2(\mathbb{R}^d)$ is

$$S_{b,s,a}f(x) = |a|^{-\frac{d + \lambda_D}{2}}f(h_{s,a}^{-1}(x - b)). \tag{16}$$

The next result generalizes Theorem 4.12 in [19] to the case when S is not Abelian.

Theorem 2. *The representation S is square-integrable and its admissible vectors ψ are the elements of $L^2(\mathbb{R}^d)$ satisfying*

$$0 < C_\psi = \int_{\mathbb{R}^d} \frac{|\mathcal{F}\psi(\xi)|^2}{|\xi_1|^d}\mathrm{d}\xi < +\infty, \tag{17}$$

where $\xi = (\xi_1, \xi') \in \mathbb{R} \times \mathbb{R}^{d-1}$.

We recall that a unitary representation π of G acting on a Hilbert space $\mathcal{H}$ is *square integrable* if it is irreducible and if there exists a (non-zero) element $\psi \in \mathcal{H}$, called *admissible vector*, such that the associated *voice transform, i.e.* the linear map $f \mapsto \langle f, \pi(b, h)\psi \rangle$, takes values in $L^2(G)$ and in such case it is a multiple of an isometry, denoted by $\mathcal{W}_\psi : \mathcal{H} \to L^2(G)$.

Proof of Theorem 2. The proof is an immediate consequence of the following result due to Führ, see [26] and the references therein. The quasi-regular representation of $\mathbb{R}^d \rtimes H$ is square integrable if and only if there exists a vector $\xi_0 \in \mathbb{R}^d$ such that

 (i) the dual orbit $\mathcal{O}_{\xi_0} = \{^t h\xi_0 \in \mathbb{R}^d : h \in H\}$ is open and it is of full measure,
 (ii) the stabilizer $H_{\xi_0} = \{h \in H : {}^t h\xi_0 = \xi_0\}$ is compact,

where saying that $\mathcal{O}_{\xi_0}$ has full measure means that its complement has Lebesgue measure zero. In such case, a vector ψ is admissible if and only if

$$\int_H |\mathcal{F}\psi(^t h\xi_0)|^2\mathrm{d}h < +\infty. \tag{18}$$

In our setting, with the choice $\xi_0 = (1, 0, \ldots, 0)$ we have that

$$^t h_{s,a}\xi_0 = {}^t h_{0,a}\, {}^t h_{s,1} \begin{bmatrix} 1 \\ 0 \\ \ldots \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ \Lambda(a)s \end{bmatrix}$$

so that $\mathcal{O}_{\xi_0} = \mathbb{R}^\times \times \mathbb{R}^{d-1}$, which is of full measure, and H_{ξ_0} is trivial. Hence S is square-integrable.

To compute the admissible vectors, notice that by (18)

$$\int_H |\mathcal{F}\psi(^t h\xi_0)|^2 \mathrm{d}h = \int_{\mathbb{R}^{d-1}\times\mathbb{R}^\times} |\mathcal{F}\psi(a\Lambda(a)s, a)|^2 |a|^{\lambda_D - 1} \mathrm{d}s\mathrm{d}a$$

$$= \int_{\mathbb{R}^{d-1}\times\mathbb{R}^\times} \frac{|\mathcal{F}\psi(\xi_1, \xi')|^2}{|\xi_1|^d} \mathrm{d}\xi_1\mathrm{d}\xi'$$

with the change of variables $a = \xi_1$ and $s = \Lambda(\xi_1)^{-1}\xi'/\xi_1$. $\qquad\square$

Theorem 2 states the surprising fact that the admissibility condition is the same for all generalized shearlet dilation groups. A canonical choice is to assume that

$$\mathcal{F}\psi(\xi_1, \xi') = \mathcal{F}\psi_1(\xi_1)\mathcal{F}\psi_2(\xi'/\xi_1) \tag{19}$$

where $\psi_1 \in L^2(\mathbb{R})$ satisfies

$$\int_{\mathbb{R}^\times} \frac{|\mathcal{F}\psi_1(\xi_1)|^2}{|\xi_1|}\mathrm{d}\xi_1 < +\infty. \tag{20}$$

and $\psi_2 \in L^2(\mathbb{R}^{d-1})$. However, other choices are available and, in particular, it is possible to build shearlets with compact support in space [27]. We finally recall that, since the representation S is square-integrable, we have the weakly-convergent reproducing formula [28]

$$f = \frac{1}{C_\psi}\int_G \mathcal{S}_\psi f(b, s, a)\, S_{b,s,a}\psi\, \frac{\mathrm{d}b\,\mathrm{d}s\,\mathrm{d}a}{|a|^{d+1}}. \tag{21}$$

2.4. **Wavelet Transform.** We recall that the one-dimensional affine group $\mathbb{W}$ is $\mathbb{R} \rtimes \mathbb{R}^\times$ endowed with the product

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

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

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

The corresponding wavelet transform is $\mathcal{W}_\psi : L^2(\mathbb{R}) \to L^2(\mathbb{W})$, given by

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

which is a multiple of an isometry provided that $\psi \in L^2(\mathbb{R})$ satisfies the admissibility condition, namely the Calderón equation,

$$0 < \int_\mathbb{R} \frac{|\mathcal{F}\psi(\xi)|^2}{|\xi|}\mathrm{d}\xi < +\infty \tag{22}$$

and, in such a case, ψ is called an admissible wavelet.

2.5. **The quasi-regular representation of H.** Consider now the shearlet dilation group H, with S and D its shearing and dilation subgroups, respectively (see Definition 1). As mentioned above, we identify H with $\mathbb{R}^{d-1} \times \mathbb{R}^\times$ as manifolds and sometimes denote by (s, a) the element $h_{s,a}$ of H. Recall that by (13) the product law is then

$$(s, a)(s', a') = (\Lambda(a)^{-1}s' + {}^tB(\Lambda(a)^{-1}s')s, aa').$$

Observe that H acts naturally on $\mathbb{R}^d$ and its (right) dual action is

$$
{}^th_{s,a} \begin{bmatrix} v_1 \\ v \end{bmatrix} = a \begin{bmatrix} v_1 \\ \Lambda(a)({}^tB(s)v - s) \end{bmatrix}.
$$

This implies that H acts naturally on $\mathbb{P}^{d-1} = (\mathbb{R}^d \setminus \{0\})/\sim$ as well. By identifying $\mathbb{R}^{d-1}$ with $\{(1, v) : v \in \mathbb{R}^{d-1}\}/\sim$ we get that H acts on $\mathbb{R}^{d-1}$ as

$$
{}^th_{s,a}.v = \Lambda(a)({}^tB(s)v - s).
$$

Hence we can define the quasi-regular representation of H acting on $L^2(\mathbb{R}^{d-1})$ by means of

$$
V_{s,a}f(v) = |a|^{\frac{\lambda_D}{2}} f(\Lambda(a)({}^tB(s)v - s)),
$$

where we recall that $\lambda_D = \lambda_1 + \cdots + \lambda_{d-1}$. In general, V is not irreducible, but we can always define the voice transform associated to a fixed vector $\psi \in L^2(\mathbb{R}^{d-1})$, namely the mapping $\mathcal{V}_\psi : L^2(\mathbb{R}^{d-1}) \longrightarrow C(H)$ defined by

$$
\mathcal{V}_\psi f(s, a) = \langle f, V_{s,a}\psi \rangle_2,
$$

where $C(H)$ is the space of continuous functions on H.

Example (The standard shearlet group, continued). For the classical shearlet group $\mathbb{S}^\gamma$ the shearlet representation on $L^2(\mathbb{R}^d)$ becomes

$$
S^\gamma_{b,s,a}f(x) = |a|^{-\frac{1+\gamma(d-1)}{2}} f(A_a^{-1}S_s^{-1}(x - b)), \tag{23}
$$

whereas the group H is the affine group $\mathbb{R}^{d-1} \rtimes \mathbb{R}^\times$ in dimension $d - 1$ and V is the corresponding wavelet representation

$$
V_{s,a}f(v) = |a|^{\frac{(d-1)(\gamma-1)}{2}} f\left(\frac{v - s}{|a|^{1-\gamma}}\right), \tag{24}
$$

which is not irreducible unless $d = 2$. Furthermore, the voice transform can be written as convolution operator

$$
\mathcal{V}_\psi f(s, a) = |a|^{\frac{(d-1)(\gamma-1)}{2}} \int_{\mathbb{R}^{d-1}} f(v)\overline{\psi\left(\frac{v - s}{|a|^{1-\gamma}}\right)} dv = f * \Psi_a(s)
$$

where

$$
\Psi_a(v) = |a|^{\frac{(d-1)(\gamma-1)}{2}} \overline{\psi\left(-\frac{v}{|a|^{1-\gamma}}\right)}.
$$

2.6. **The affine Radon transform.** In this section we recall the definition and the main properties of the Radon transform. Then we introduce the particular restriction of the Radon transform in which we are interested, the so-called *affine Radon transform*, obtained by parametrizing the space of hyperplanes by affine coordinates.

We first define the Radon transform on $L^1(\mathbb{R}^d)$ by following the approach in [29], see also [30] as a classical reference. Given $f \in L^1(\mathbb{R}^d)$ its Radon transform is the function $\mathcal{R}f : (\mathbb{R}^d \setminus \{0\}) \times \mathbb{R} \to \mathbb{C}$ defined by

$$\mathcal{R}f(n,t) = \frac{1}{|n|} \int_{n \cdot x = t} f(x) \, dm(x), \tag{25}$$

where m is the Euclidean measure on the hyperplane

$$(n : t) := \{x \in \mathbb{R}^d : n \cdot x = t\} \tag{26}$$

and the equality (25) holds for almost all $(n,t) \in (\mathbb{R}^d \setminus \{0\}) \times \mathbb{R}$. We add some comments. Definition (25) makes sense since, given $n \in \mathbb{R}^d \setminus \{0\}$ Fubini theorem gives that

$$\int_{\mathbb{R}^d} |f(x)| dx = \int_{\mathbb{R}} \left(\int_{n \cdot x = t} |f(x)| dm(x) \right) dt < +\infty,$$

so that for almost all $t \in \mathbb{R}$ the integral $\int_{n \cdot x = t} |f(x)| dm(x)$ is finite and $\mathcal{R}(n,t)$ is well defined.

Furthermore, each pair $(n,t) \in (\mathbb{R}^d \setminus \{0\}) \times \mathbb{R}$ defines the hyperplane $(n : t)$ by means of (26). Clearly, the correspondence between parameters (n,t) and hyperplanes is not bijective. Indeed (n',t') and (n,t) determine the same hyperplane if and only if there exists $\lambda \in \mathbb{R}^\times$ such that $n' = \lambda n$ and $t' = \lambda t$ and this equivalence relation motivates the notation $(n : t)$ for the hyperplane in (26). Because of the factor $1/|n|$ in (25), $\mathcal{R}f$ is a positively homogenous function of degree -1, *i.e.* for all $\lambda \in \mathbb{R}^\times$

$$\mathcal{R}f(\lambda n, \lambda t) = |\lambda|^{-1} \mathcal{R}f(n,t). \tag{27}$$

This means that $\mathcal{R}f$ is completely defined by choosing a representative (n,t) for each hyperplane $(n : t)$, *i.e.* by choosing a suitable system of coordinates on the affine Grassmannian

$$\{\text{hyperplanes of } \mathbb{R}^d\} \simeq \mathbb{P}^{d-1} \times \mathbb{R}.$$

The canonical choice [30] is given by parametrizing $\mathbb{P}^{d-1}$ with its two-fold covering S^{d-1}, where S^{d-1} is the unit sphere in $\mathbb{R}^d$.

We are interested in another restriction of the Radon transform. For all $v \in \mathbb{R}^{d-1}$ set ${}^t n(v) = (1, {}^t v)$.

Definition 3. Given $f \in L^1(\mathbb{R}^d)$, the *affine Radon transform of f* is the function $\mathcal{R}^{\text{aff}} f : \mathbb{R}^{d-1} \times \mathbb{R} \to \mathbb{C}$ given by

$$\mathcal{R}^{\text{aff}} f(v,t) = \mathcal{R}f(n(v), t)$$

$$= \frac{1}{\sqrt{1 + |v|^2}} \int_{n(v) \cdot x = t} f(x) dm(x) = \int_{\mathbb{R}^{d-1}} f(t - v \cdot y, y) dy. \tag{28}$$

Remark. The transform $\mathcal{R}^{\text{aff}}$ is obtained from $\mathcal{R}$ by parametrizing the projective space $\mathbb{P}^{d-1}$ with affine coordinates. Indeed, the map $(v,t) \mapsto (n(v) : t)$ is a diffeomorphism of
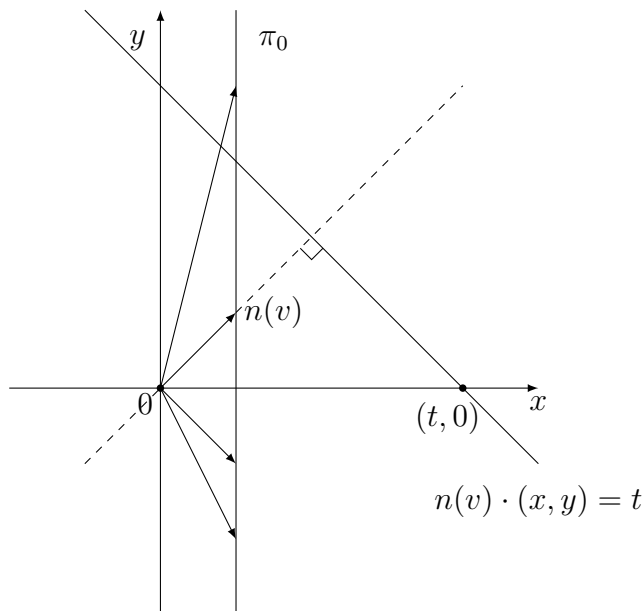
FIGURE 1. space of hyperplanes parametrized by affine coordinates (2-dimensional case)

$\mathbb{R}^{d-1} \times \mathbb{R}$ onto the open subset

$$U_0 = \{(n : t) : \exists \lambda \in \mathbb{R}^\times \text{ s.t. } \lambda n \in \pi_0\},$$

where $\pi_0 = \{n(v) : v \in \mathbb{R}^{d-1}\}$. The complement of U_0 is the set of *horizontal hyperplanes*, those for which the normal vector has the first component equal to zero (see Figure 1 for the 2-dimensional case). The set of pairs (v, t) such that $(n(v) : t) \notin U_0$ is negligible, so that $\mathcal{R}^{\mathrm{aff}} f$ completely defines $\mathcal{R} f$. In A we recall the relation between the affine Radon transform and the usual Radon transform in polar coordinates.

The next proposition, whose proof can be found in [29], summarizes the behaviour of the Radon transform under affine linear actions. The translation and dilation operators act on a function $f : \mathbb{R}^d \to \mathbb{C}$ as

$$T_b f(x) = f(x - b), \qquad D_A f(x) = |\det A|^{-1} f\left(A^{-1}x\right),$$

respectively, for $b \in \mathbb{R}^d$ and $A \in \mathrm{GL}(d, \mathbb{R})$. Both operators map each $L^p(\mathbb{R}^d)$ onto itself and D_A is normalized to be an isometry on $L^1(\mathbb{R}^d)$.

Proposition 4. *Given $f \in L^1(\mathbb{R}^d)$, the following properties hold true:*

    (i) $\mathcal{R} T_b f(n, t) = \mathcal{R} f(n, t - n \cdot b)$, for all $b \in \mathbb{R}^d$;
    (ii) $\mathcal{R} D_A f(n, t) = \mathcal{R} f({}^t A n, t)$, for all $A \in GL(d, \mathbb{R})$.

We now state a crucial result in Radon transform theory in its standard version. Below we prove two variations that are taylored to our setting but are also of some independent interest.

Proposition 5 (Fourier slice theorem, 1). *For any $f \in L^1(\mathbb{R}^d)$*

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

for all $n \in \mathbb{R}^d \setminus \{0\}$ and all $\tau \in \mathbb{R}$.

Here the Fourier transform on the right hand side is in $\mathbb{R}^d$, whereas the operator $\mathcal{F}$ on the left hand side is 1-dimensional and acts on the variabile t. We repeat this slight abuse of notation in other formulas below.

In the next formulation, written for the affine Radon transform, the function f to which $\mathcal{R}^{\mathrm{aff}}$ is applied is taken in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

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

$$\mathcal{R}^{\mathrm{aff}} f(v, \cdot) = \mathcal{F}^{-1}[\mathcal{F}f \circ \psi(v, \cdot)]. \tag{29}$$

Proof. By Proposition 5 we know that for all $v \in \mathbb{R}^{d-1}$ the affine Radon transform $\mathcal{R}^{\mathrm{aff}} f(v, \cdot)$ is in $L^1(\mathbb{R})$ and satisfies

$$\mathcal{F}(\mathcal{R}^{\mathrm{aff}} f(v, \cdot))(\tau) = \mathcal{F}f \circ \psi(v, \tau), \qquad \tau \in \mathbb{R}.$$

We start by proving that the function $\tau \mapsto \mathcal{F}f \circ \psi(v, \tau)$ is in $L^2(\mathbb{R})$, that is

$$\int_{\mathbb{R}} |\mathcal{F}f \circ \psi(v, \tau)|^2 \mathrm{d}\tau < +\infty.$$

The map $\psi : \mathbb{R}^{d-1} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}^d$, defined by $\psi(v, \tau) = \tau n(v)$, is a diffeomorphism onto the open set $V = \{\xi \in \mathbb{R}^d : \xi_1 \neq 0\}$ with Jacobian $J\psi(v, \tau) = \tau^{d-1}$. By hypothesis we know that

$$\|f\|_2^2 = \int_{\mathbb{R}^d} |\mathcal{F}f(\xi)|^2 \mathrm{d}\xi = \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |\mathcal{F}f \circ \psi(v, \tau)|^2 |\tau|^{d-1} \mathrm{d}\tau \mathrm{d}v < +\infty,$$

so that there exists a negligible set $E \subseteq \mathbb{R}^{d-1}$ such that

$$C_f := \int_{\mathbb{R}} |\mathcal{F}f \circ \psi(v, \tau)|^2 |\tau|^{d-1} \mathrm{d}\tau < +\infty$$

for all $v \notin E$. Therefore, for all $v \notin E$ it holds

$$\int_{\mathbb{R}} |\mathcal{F}f \circ \psi(v, \tau)|^2 \mathrm{d}\tau = \int_{|\tau| \leq 1} |\mathcal{F}f \circ \psi(v, \tau)|^2 \mathrm{d}\tau + \int_{|\tau| > 1} \frac{|\tau|^{d-1}}{|\tau|^{d-1}} |\mathcal{F}f \circ \psi(v, \tau)|^2 \mathrm{d}\tau$$

$$\leq 2\|\mathcal{F}f\|_\infty + \int_{\mathbb{R}} |\mathcal{F}f \circ \psi(v, \tau)|^2 |\tau|^{d-1} \mathrm{d}\tau$$

$$\leq 2\|f\|_1^2 + C_f < +\infty.$$

Hence the function $t \mapsto \mathcal{R}^{\mathrm{aff}} f(v, t)$ is in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and (29) follows by the Fourier inversion formula in $L^2(\mathbb{R})$. $\qquad\square$

It is possible to extend the affine Radon transform $\mathcal{R}^{\mathrm{aff}}$ to $L^2(\mathbb{R}^d)$ as a unitary map. However, this raises some technical issues, that are addressed in the next section.

2.7. **The unitary extension.** Consider the subspace

$$\mathcal{D} = \left\{ f \in L^2(\mathbb{R}^{d-1} \times \mathbb{R}) : \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1} |\mathcal{F}f(\xi, \tau)|^2 \, \mathrm{d}\xi \mathrm{d}\tau < +\infty \right\}$$

of $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ and define the operator $\mathcal{I} : \mathcal{D} \to L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ by

$$\mathcal{F}\mathcal{I}f(\xi, \tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}f(\xi, \tau), \tag{30}$$

a Fourier multiplier with respect to the last variable. Since $\tau \mapsto |\tau|^{\frac{d-1}{2}}$ is a strictly positive (almost everywhere) Borel function on $\mathbb{R}$, the spectral theorem for unbounded operators, see Theorem VIII.6 of [31], shows that $\mathcal{D}$ is dense and that $\mathcal{I}$ is a positive self-adjoint injective operator.

Remark. The operator $\mathcal{I}$ is related to the inverse of the Riesz potential with exponent $(d-1)/2$ on $L^2(\mathbb{R})$. Indeed, if $\psi_2 \in L^2(\mathbb{R}^{d-1})$ and if $\psi_1 \in L^2(\mathbb{R})$ is such that

$$\int_{\mathbb{R}} |\tau|^{d-1} |\mathcal{F}\psi_1(\tau)|^2 \, d\tau < +\infty,$$

then $\psi_2 \otimes \psi_1 \in \mathcal{D}$, because

$$\int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1} |\mathcal{F}(\psi_2 \otimes \psi_1)(\xi, \tau)|^2 \, d\xi d\tau$$

$$= \int_{\mathbb{R}^{d-1}} |\mathcal{F}\psi_2(\xi)|^2 \, d\xi \int_{\mathbb{R}} |\tau|^{d-1} |\mathcal{F}\psi_1(\tau)|^2 \, d\tau < +\infty,$$

so that

$$\mathcal{I}(\psi_2 \otimes \psi_1) = \psi_2 \otimes \mathcal{I}_0 \psi_1,$$

where $\mathcal{I}_0$ is the inverse of the standard Riesz potential defined by

$$\mathcal{F}\mathcal{I}_0\psi_1(\tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}\psi_1(\tau). \tag{31}$$

Furthermore, $\mathcal{D}$ is invariant under translations and dilations by matrices of the form

$$A = \begin{bmatrix} A_0 & 0 \\ v & a \end{bmatrix}, \tag{32}$$

where $A_0 \in \mathrm{GL}(d-1, \mathbb{R})$, $v \in \mathbb{R}^{d-1}$, $a \in \mathbb{R}^\times$.

Lemma 7. *For all $b \in \mathbb{R}^d$ and A as in (32) it holds*

$$\mathcal{I}T_b = T_b\mathcal{I}, \qquad \mathcal{I}D_A = |a|^{-\frac{d-1}{2}} D_A\mathcal{I}. \tag{33}$$

Proof. The first of relations (33) is a consequence of the fact that $\mathcal{F}T_b f(\xi, \tau) = e^{-2\pi i b \cdot \xi} \mathcal{F}f(\xi, \tau)$ for all $f \in L^2(\mathbb{R}^{d-1} \times \mathbb{R})$. Precisely, for all $f \in \mathcal{D}$ we have that

$$\mathcal{F}\mathcal{I}T_b f(\xi, \tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}T_b f(\xi, \tau)$$

$$= |\tau|^{\frac{d-1}{2}} e^{-2\pi i b \cdot \xi} \mathcal{F}f(\xi, \tau)$$

$$= e^{-2\pi i b \cdot \xi} \mathcal{F}\mathcal{I}f(\xi, \tau)$$

$$= \mathcal{F}T_b\mathcal{I}f(\xi, \tau),$$

whence $\mathcal{I}T_b = T_b\mathcal{I}$. The second follows from $\mathcal{F}D_A f(\xi, \tau) = \mathcal{F}f(^t A(\xi, \tau))$. Indeed, for all $f \in \mathcal{D}$

$$\mathcal{F}\mathcal{I}D_A f(\xi, \tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}D_A f(\xi, \tau)$$

$$= |\tau|^{\frac{d-1}{2}} \mathcal{F}f(^t A(\xi, \tau))$$

$$= |\tau|^{\frac{d-1}{2}} \mathcal{F}f(^t A_0 \xi + \tau v, a\tau)$$

$$= |\tau|^{\frac{d-1}{2}} |a\tau|^{-\frac{d-1}{2}} \mathcal{F}\mathcal{I}f(^t A_0 \xi + \tau v, a\tau)$$

$$= |a|^{-\frac{d-1}{2}} \mathcal{F}D_A\mathcal{I}f(\xi, \tau).$$

This proves (33). $\square$

The space $\mathcal{D}$ becomes a pre-Hilbert space with respect to the scalar product

$$\langle f, g \rangle_{\mathcal{D}} = \langle \mathcal{I}f, \mathcal{I}g \rangle_2$$
$$= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1} \mathcal{F}(f(v, \cdot))(\tau) \overline{\mathcal{F}(g(v, \cdot))(\tau)} \, \mathrm{d}v \mathrm{d}\tau. \tag{34}$$

Furthermore,

$$||f||_{\mathcal{D}}^2 = \langle f, f \rangle_{\mathcal{D}} = \langle \mathcal{I}f, \mathcal{I}f \rangle_2 = ||\mathcal{I}f||_2^2,$$

for all $f \in \mathcal{D}$. Hence $\mathcal{I}$ is an isometric operator from $\mathcal{D}$, with the new scalar product (34), to $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$. Since $\mathcal{I}$ is self-adjoint and injective, $\mathrm{Ran}(\mathcal{I})$ is dense in $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$. Hence, by standard arguments, it extends uniquely to a unitary operator, denoted $\mathscr{I}$, from the completion $\mathcal{H}$ of $\mathcal{D}$ onto $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$.

To extend $\mathcal{R}^{\mathrm{aff}}$ to $L^2(\mathbb{R}^d)$ as a unitary operator, note that, by Proposition 5 with $n = n(v)$, the affine Radon transform of $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ belongs to $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ if and only if is finite the integral

$$\int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\mathcal{R}^{\mathrm{aff}} f(v, t)|^2 \mathrm{d}v \mathrm{d}t = \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |\mathcal{F}(\mathcal{R}^{\mathrm{aff}} f(v, \cdot))(\tau)|^2 \mathrm{d}\tau \mathrm{d}v$$
$$= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\mathcal{F}f(\tau, \tau v)|^2 \mathrm{d}\tau \mathrm{d}v$$
$$= \int_{\mathbb{R}^d} \frac{|\mathcal{F}f(\xi)|^2}{|\xi_1|^{d-1}} \mathrm{d}\xi,$$

where ξ_1 is the first component of the vector $\xi \in \mathbb{R}^d$. Therefore requiring that $\mathcal{R}^{\mathrm{aff}} f$ belongs to $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ is equivalent to

$$\int_{\mathbb{R}^d} \frac{|\mathcal{F}f(\xi)|^2}{|\xi_1|^{d-1}} \mathrm{d}\xi < +\infty.$$

We denote by

$$\mathcal{A} = \{f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d} \frac{|\mathcal{F}f(\xi)|^2}{|\xi_1|^{d-1}} \mathrm{d}\xi < +\infty\},$$

which is dense in $L^2(\mathbb{R}^d)$ since it contains the functions whose Fourier transform is smooth and has compact support disjoint from the hyperplane $\xi_1 = 0$. By definition of $\mathcal{A}$, $\mathcal{R}^{\mathrm{aff}} f \in L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ for all $f \in \mathcal{A}$.

We shall need a suitable formulation of the main result in Radon transform theory, namely the following version of Theorem 4.1 in [30]. For the sake of completeness we include the proof in B.

Theorem 8. *The affine Radon transform extends to a unique unitary operator from $L^2(\mathbb{R}^d)$ onto $\mathcal{H}$, denoted with $\mathscr{R}$ and, hence, $\mathcal{Q} = \mathscr{I}\mathscr{R}$ is a unitary operator from $L^2(\mathbb{R}^d)$ onto $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$.*

As mentioned above, we need yet another generalization of the Fourier slice theorem (Proposition 5). We think that it is perhaps known, but we could not locate it in the literature. The proof is given in B.

Proposition 9 (Fourier slice theorem, 3)**.** *For all* $f \in L^2(\mathbb{R}^d)$

$$\mathcal{F}(\mathcal{Q}f(v,\cdot))(\tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}f(\tau n(v)) \tag{35}$$

for almost every $(v,\tau) \in \mathbb{R}^{d-1} \times \mathbb{R}$.

3. The Intertwining Theorem and its consequences

3.1. **The main Theorem.** We recall that the group G is the semidirect product $G = \mathbb{R}^d \rtimes H$ where $H = SD$ is the shearlet dilation group, S is the shearing subgroup and D the scaling subgroup of H, as in Definition 1. Each element in G is parametrized by a triple $(b, s, a) \in \mathbb{R}^d \times \mathbb{R}^{d-1} \times \mathbb{R}^\times$ and $S_{b,s,a}$ is as in (16).

Theorem 10. *The unitary operator* $\mathcal{Q}$ *intertwines the shearlet representation with the tensor product of two unitary representations, precisely*

$$\mathcal{Q}S_{b,s,a}f(v,t) = (V_{s,a} \otimes W_{n(v)\cdot b,a})\mathcal{Q}f(v,t) \tag{36}$$

for every $f \in L^2(\mathbb{R}^d)$.

Proof. By density, it is enough to prove the equality on $\mathcal{A}$. We shall use throughout the fact that $\mathcal{R}^{\mathrm{aff}}f \in \mathcal{D}$ for every $f \in \mathcal{A}$ and that $\mathcal{D}$ is invariant under all translations and under the dilations described in (32). Since for all $(b,s,a) \in G$ it holds $(b,s,a) = (b,0,1)(0,s,1)(0,0,a)$, it is sufficient to prove the equality for each of the three factors. For $f \in \mathcal{A}$ and $b \in \mathbb{R}^d$ we have

$$
\begin{aligned}
\mathcal{R}^{\mathrm{aff}}S_{b,0,1}f(v,t) &= \mathcal{R}^{\mathrm{aff}}T_b f(v,t) \\
&= \mathcal{R}T_b f(n(v),t) \\
&= \mathcal{R}f(n(v), t - n(v)\cdot b) \\
&= \mathcal{R}^{\mathrm{aff}}f(v, t - n(v)\cdot b) \\
&= (\mathrm{I} \otimes W_{n(v)\cdot b,1})\mathcal{R}^{\mathrm{aff}}f(v,t).
\end{aligned}
$$

Since $\mathcal{I}$ commutes with translations, $\mathrm{I} \otimes W_{n(v)\cdot b,1} = T_{(0,n(v)\cdot b)}$ implies

$$\mathcal{I}\mathcal{R}^{\mathrm{aff}}S_{b,0,1}f(v,t) = \mathcal{I}(\mathrm{I} \otimes W_{n(v)\cdot b,1})\mathcal{R}^{\mathrm{aff}}f(v,t) = (\mathrm{I} \otimes W_{n(v)\cdot b,1})\mathcal{I}\mathcal{R}^{\mathrm{aff}}f(v,t).$$

For $f \in \mathcal{A}$ and $a \in \mathbb{R}^\times$ we have

$$
\begin{aligned}
\mathcal{R}^{\mathrm{aff}}S_{0,0,a}f(v,t) &= |a|^{\frac{d+\lambda_D}{2}}\mathcal{R}^{\mathrm{aff}}D_{h_{0,a}}f(v,t) \\
&= |a|^{\frac{d+\lambda_D}{2}}\mathcal{R}D_{h_{0,a}}f(n(v),t) \\
&= |a|^{\frac{d+\lambda_D}{2}}\mathcal{R}f({}^t h_{0,a}n(v),t).
\end{aligned}
$$

A direct calculation gives

$${}^t h_{0,a}n(v) = a \begin{bmatrix} 1 & 0 \\ 0 & \Lambda(a) \end{bmatrix} \begin{bmatrix} 1 \\ v \end{bmatrix} = a \begin{bmatrix} 1 \\ \Lambda(a)v \end{bmatrix} = an\left(\Lambda(a)v\right).$$

The behavior of the Radon transform under linear operations implies that

$$\mathcal{R}^{\text{aff}} S_{0,0,a} f(v,t) = |a|^{\frac{d+\lambda_D}{2}} \mathcal{R} f\left(an\left(\Lambda(a)v\right), a\frac{t}{a}\right)$$

$$= |a|^{\frac{d+\lambda_D}{2}-1} \mathcal{R}^{\text{aff}} f\left(\Lambda(a)v, \frac{t}{a}\right)$$

$$= |a|^{\frac{d-1}{2}} (V_{0,a} \otimes W_{0,a}) \mathcal{R}^{\text{aff}} f(v,t).$$

Since

$$(V_{0,a} \otimes W_{0,a}) = |a|^{\frac{3(1-\lambda_D)}{2}} D_A,$$

where the matrix A is of the form

$$A = \begin{bmatrix} \Lambda(a)^{-1} & 0 \\ 0 & a \end{bmatrix},$$

and because of the behavior of the operator $\mathcal{I}$ under dilations, we obtain

$$\mathcal{I}\mathcal{R}^{\text{aff}} S_{0,0,a} f(v,t) = \mathcal{I}|a|^{\frac{d-1}{2}} (V_{0,a} \otimes W_{0,a}) \mathcal{R}^{\text{aff}} f(v,t)$$

$$= (V_{0,a} \otimes W_{0,a}) \mathcal{I}\mathcal{R}^{\text{aff}} f(v,t).$$

Finally, let $s = {}^t(s_1, \ldots, s_{d-1}) \in \mathbb{R}^{d-1}$. Then

$$\mathcal{R}^{\text{aff}} S_{0,s,1} f(v,t) = \mathcal{R}^{\text{aff}} D_{h_{s,1}} f(v,t)$$

$$= \mathcal{R} D_{h_{s,1}} f(n(v),t) = \mathcal{R} f({}^t h_{s,1} n(v), t).$$

Since

$${}^t h_{s,1} n(v) = \begin{bmatrix} 1 & 0 \\ -s & {}^t B(s) \end{bmatrix} \begin{bmatrix} 1 \\ v \end{bmatrix} = \begin{bmatrix} 1 \\ {}^t B(s)v - s \end{bmatrix} = n\left({}^t B(s)v - s\right),$$

by Proposition 4 we obtain the following string of equalities:

$$\mathcal{R}^{\text{aff}} S_{0,s,1} f(v,t) = \mathcal{R} f\left(n\left({}^t B(s)v - s\right), t\right)$$

$$= \mathcal{R}^{\text{aff}} f\left({}^t B(s)v - s, t\right)$$

$$= (V_{s,1} \otimes \text{I}) \mathcal{R}^{\text{aff}} f(v,t).$$

Finally,

$$(V_{s,1} \otimes \text{I}) = |a|^{\frac{3(1-\lambda_D)}{2}} T_{(-({}^t B(s))^{-1} s, 0)} D_A,$$

where

$$A = \begin{bmatrix} {}^t B(s)^{-1} & 0 \\ 0 & 1 \end{bmatrix},$$

so that the behavior of $\mathcal{I}$ under dilations implies

$$\mathcal{I}\mathcal{R}^{\text{aff}} S_{0,s,1} f(v,t) = \mathcal{I}(V_{s,1} \otimes \text{I}) \mathcal{R}^{\text{aff}} f(v,t)$$

$$= (V_{s,1} \otimes \text{I}) \mathcal{I}\mathcal{R}^{\text{aff}} f(v,t).$$

Therefore, by

$$\mathcal{I}\mathcal{R}^{\text{aff}} S_{b,s,a} f = \mathcal{I}\mathcal{R}^{\text{aff}} S_{b,0,1} S_{0,s,1} S_{0,0,a} f,$$

equation (36) follows applying the relations obtained above. □

3.2. **The admissibility conditions.** In this subsection we discuss the admissibility conditions and some of their consequences.

Our objective is to obtain an expression for the shearlet transform that makes use of formula (36). To this end, we start by looking for natural conditions that guarantee that $\psi \in L^2(\mathbb{R}^d)$ is an admissible vector for the shearlet representation S, namely that it satisfies (17).

Equation (36) suggests that a good choice for the admissible vector ψ is of the form

$$\mathcal{Q}\psi = \phi_2 \otimes \phi_1$$

where $\phi_1 \in L^2(\mathbb{R})$, $\phi_2 \in L^2(\mathbb{R}^{d-1})$. If this is the case, then by (35) it follows that

$$\phi_2(v)\mathcal{F}\phi_1(\tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}\psi(\tau n(v))$$

so that $\mathcal{F}\psi$ factorizes as

$$\mathcal{F}\psi(\tau, \tau v) = \mathcal{F}\psi_1(\tau)\mathcal{F}\psi_2(v), \tag{37}$$

where we assume that $\psi_2 \in L^2(\mathbb{R}^{d-1})$ and $\psi_1 \in L^2(\mathbb{R})$. Equation (37) is the canonical choice of admissible vectors given by (19). Furthermore,

$$\mathcal{F}\phi_1(\tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}\psi_1(\tau) \qquad \phi_2(v) = \mathcal{F}\psi_2(v),$$

so that the assumption that $\psi_2 \in L^2(\mathbb{R}^{d-1})$ is automatically satisfied. Since $\phi_1 \in L^2(\mathbb{R})$, then

$$\int_{\mathbb{R}} |\tau|^{d-1} |\mathcal{F}\psi_1(\tau)|^2 \, \mathrm{d}\tau < +\infty.$$

This, together with the fact that $\psi_1 \in L^2(\mathbb{R})$, implies that ψ_1 belongs to the domain of the differential operator $\mathcal{I}_0$ (see (31)). Therefore

$$\phi_1 = \mathcal{I}_0\psi_1. \tag{38}$$

With the choice (37) the admissibility condition (17) reduces to

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

From now on we fix $\psi \in L^2(\mathbb{R}^d)$ of the form (37) with $\psi_1 \in L^2(\mathbb{R})$ satisfying

$$\int_{\mathbb{R}} |\tau|^{d-1} |\mathcal{F}\psi_1(\tau)|^2 \, \mathrm{d}\tau < +\infty, \qquad 0 < \int_{\mathbb{R}} \frac{|\mathcal{F}\psi_1(\tau)|^2}{|\tau|} \mathrm{d}\tau < +\infty, \tag{39}$$

and $\psi_2 \in L^2(\mathbb{R}^{d-1})$.

Corollary 11. *Under the assumptions* (39)*, for every* $L^2(\mathbb{R}^d)$

$$\mathcal{S}_\psi f(b, s, a) = \mathcal{V}_{\phi_2} \left(\mathcal{W}_{\phi_1}(\mathcal{Q}f(v, t))(n(v) \cdot b, a) \right)(s, a) \tag{40}$$

Proof. For all $f \in L^2(\mathbb{R}^d)$ and $(b, s, a) \in G$

$$
\begin{aligned}
\mathcal{S}_\psi f(b, s, a) &= \langle f, S_{b,s,a}\psi \rangle_2 \\
&= \langle \mathcal{Q}f, \mathcal{Q}S_{b,s,a}\psi \rangle_2 \\
&= \langle \mathcal{Q}f, (V_{s,a} \otimes W_{n(\cdot) \cdot b, a})\mathcal{Q}\psi \rangle_2 \\
&= \langle \mathcal{Q}f, (V_{s,a} \otimes W_{n(\cdot) \cdot b, a})(\phi_2 \otimes \phi_1) \rangle_2 \\
&= \langle \mathcal{Q}f, V_{s,a}\phi_2 \otimes W_{n(\cdot) \cdot b, a}\phi_1 \rangle_2 \\
&= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} \mathcal{Q}f(v, \tau)\overline{V_{s,a}\phi_2(v)W_{n(v) \cdot b, a}\phi_1(\tau)} \; \mathrm{d}v\mathrm{d}\tau \\
&= \int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} \mathcal{Q}f(v, \tau)\overline{W_{n(v) \cdot b, a}\phi_1(\tau)} \; \mathrm{d}\tau \right) \overline{V_{s,a}\phi_2(v)} \; \mathrm{d}v \\
&= \int_{\mathbb{R}^{d-1}} \mathcal{W}_{\phi_1}(\mathcal{Q}f(v, \bullet))(n(v) \cdot b, a)\overline{V_{s,a}\phi_2(v)} \; \mathrm{d}v, \tag{41}
\end{aligned}
$$

where in the last equality we have used the fact that ϕ_1 is an admissible wavelet. This is true because by (38)

$$
\begin{aligned}
\int_{\mathbb{R}} \frac{|\mathcal{F}\phi_1(\tau)|^2}{|\tau|} \; \mathrm{d}\tau &= \int_{\mathbb{R}} \frac{|\mathcal{F}\mathcal{I}_0\psi_1(\tau)|^2}{|\tau|} \; \mathrm{d}\tau \\
&\leq \int_{0 < |\tau| < 1} \frac{|\mathcal{F}\psi_1(\tau)|^2}{|\tau|} \; \mathrm{d}\tau + \int_{|\tau| \geq 1} |\tau|^{d-1}|\mathcal{F}\psi_1(\tau)|^2 \; \mathrm{d}\tau \\
&\leq \int_{\mathbb{R}} \frac{|\mathcal{F}\psi_1(\tau)|^2}{|\tau|} \; \mathrm{d}\tau + \int_{\mathbb{R}} |\tau|^{d-1}|\mathcal{F}\psi_1(\tau)|^2 \; \mathrm{d}\tau,
\end{aligned}
$$

which are both finite. $\square$

Equation (41) shows that the shearlet coefficients $\mathcal{S}_\psi f(b, s, a)$ can be computed in terms of the unitary Radon transform $\mathcal{Q}f$, which involves the pseudo-differential operator $\mathcal{I}$ and it is difficult to compute numerically. However, if $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, there is yet a different way to express the shearlet transform. To this end we need to choose ψ in such a way that $\mathcal{Q}\psi$ is in the domain of the operator $\mathcal{I}$, that is, in such a way that

$$
\int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1}|\mathcal{F}\mathcal{Q}\psi(v, \tau)|^2 \; \mathrm{d}v\mathrm{d}\tau < +\infty.
$$

Assuming this and recalling that $\mathcal{Q}\psi = \mathcal{F}\psi_2 \otimes \mathcal{I}_0\psi_1$ we obtain

$$
\begin{aligned}
&\int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1}|\mathcal{F}\mathcal{Q}\psi(v, \tau)|^2 \; \mathrm{d}v\mathrm{d}\tau \\
&= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1}|\mathcal{F}(\mathcal{F}\psi_2 \otimes \mathcal{I}_0\psi_1)(v, \tau)|^2 \; \mathrm{d}v\mathrm{d}\tau \\
&= \int_{\mathbb{R}^{d-1}} |\mathcal{F}\psi_2(v)|^2 \; \mathrm{d}v \int_{\mathbb{R}} |\tau|^{d-1}|\mathcal{F}\mathcal{I}_0\psi_1(\tau)|^2 \; \mathrm{d}\tau \\
&= ||\psi_2||_2^2 \int_{\mathbb{R}} |\tau|^{2(d-1)}|\mathcal{F}\psi_1(\tau)|^2 \; \mathrm{d}\tau.
\end{aligned}
$$

This shows that $\mathcal{Q}\psi$ is in the domain of $\mathcal{I}$ if and only if ψ_1 satisfies the additional condition

$$
\int_{\mathbb{R}} |\tau|^{2(d-1)}|\mathcal{F}\psi_1(\tau)|^2 \; \mathrm{d}\tau < +\infty. \tag{42}
$$

In this case, by (38)

$$\mathcal{I}Q\psi = \mathcal{I}(\phi_2 \otimes \phi_1) = \phi_2 \otimes \mathcal{I}_0\phi_1.$$

Corollary 12. *Under the assumptions* (39) *and* (42),

$$\mathcal{S}_\psi f(b,s,a) = |a|^{-\frac{d-1}{2}} \mathcal{V}_{\phi_2}\left(\mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{aff}} f(v,t))(n(v)\cdot b, a)\right)(s,a).$$

for all $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

Proof. For all $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and $(b,s,a) \in G$

$$\mathcal{S}_\psi f(b,s,a) = \int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} Qf(v,\tau)\overline{W_{n(v)\cdot b, a}\phi_1(\tau)}\ \mathrm{d}\tau\right)\overline{V_{s,a}\phi_2(v)}\ \mathrm{d}v$$
$$= \int_{\mathbb{R}^{d-1}} \langle Qf(v,\cdot), W_{n(v)\cdot b, a}\phi_1 \rangle_2 \overline{V_{s,a}\phi_2(v)}\ \mathrm{d}v. \tag{43}$$

Since $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, Proposition 6 and Proposition 9 imply that for almost all $v \in \mathbb{R}^{d-1}$, $\mathcal{R}^{\mathrm{aff}} f(v,\cdot)$ is in $L^2(\mathbb{R})$ and

$$\mathcal{F}(Qf(v,\cdot))(\tau) = |\tau|^{\frac{d-1}{2}} \mathcal{F}f(\tau, \tau v) = |\tau|^{\frac{d-1}{2}} \mathcal{F}\mathcal{R}^{\mathrm{aff}} f(v,\cdot)(\tau).$$

Since $\mathcal{F}(Qf(v,\cdot)) \in L^2(\mathbb{R})$ for almost all $v \in \mathbb{R}^{d-1}$, the above equality implies that $\mathcal{R}^{\mathrm{aff}} f(v,\cdot)$ is in the domain of $\mathcal{I}_0$ and, by definition of $\mathcal{I}_0$,

$$Qf(v,\cdot)(\tau) = \mathcal{I}_0\mathcal{R}^{\mathrm{aff}} f(v,\cdot).$$

By assumption ϕ_1 is in the domain of $\mathcal{I}_0$ and the same property holds true for $W_{n(v)\cdot b, a}\phi_1$. Since $\mathcal{I}_0$ is self-adjoint, we get

$$\langle Qf(v,\cdot), W_{n(v)\cdot b, a}\phi_1 \rangle_2 = \langle \mathcal{R}^{\mathrm{aff}} f(v,\cdot), \mathcal{I}_0 W_{n(v)\cdot b, a}\phi_1 \rangle_2$$
$$= |a|^{-\frac{d-1}{2}} \langle \mathcal{R}^{\mathrm{aff}} f(v,\cdot), W_{n(v)\cdot b, a}\mathcal{I}_0\phi_1 \rangle_2,$$

by taking into account that

$$\mathcal{I}_0 W_{n(\cdot)\cdot b, a} = |a|^{-\frac{d-1}{2}} W_{n(\cdot)\cdot b, a}\mathcal{I}_0.$$

Setting $\chi_1 = \mathcal{I}_0\phi_1 = \mathcal{I}_0^2\psi_1$, *i.e.*

$$\mathcal{F}\chi_1(\tau) = |\tau|\mathcal{F}\psi_1(\tau), \tag{44}$$

from (43) we finally get

$$\mathcal{S}_\psi f(b,s,a) = \tag{45}$$
$$|a|^{\frac{\lambda_D+1-d}{2}} \int_{\mathbb{R}^{d-1}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{aff}} f(v,\cdot))(n(v)\cdot b, a)\overline{\phi_2\left(\Lambda(a)({}^tB(s)v - s)\right)}\ \mathrm{d}v$$

Observe that we have used the fact that χ_1 is an admissible wavelet, too, the proof is analogous to the proof that ψ_1 is such. As for (41) we can rewrite the above formula by using the voice transform of H, *i.e.*

$$\mathcal{S}_\psi f(b,s,a) = |a|^{-\frac{d-1}{2}} \mathcal{V}_{\phi_2}\left(\mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{aff}} f(v,t))(n(v)\cdot b, a)\right)(s,a).$$

$\square$

Observe that formulas (41) and (45) can be also written in terms of the polar Radon transform using relation (51).

Equation (45) shows that for any signal $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ the shearlet coefficients can be computed by means of three *classical* transforms: first compute the affine Radon transform $\mathcal{R}^{\mathrm{aff}} f$, then apply the wavelet transform to the last variable

$$G(v, b, a) = \mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{aff}} f(v, \cdot))(n(v) \cdot b, a),$$

where χ_1 is given by (44), and, finally, "mock-convolve" with respect to the variable v

$$\mathcal{S}_\psi^\gamma f(b, s, a) = \int_{\mathbb{R}^{d-1}} G(v, b, a) \Phi_a(s - {}^t B(s) v) \, \mathrm{d}v$$

with the scale-dependent filter

$$\Phi_a(v) = \overline{\phi_2\left(-\Lambda(a)v\right)}.$$

Note that the "mok-convolution" reduces to the standard convolution in $\mathbb{R}^{d-1}$ when $B(t) = \mathrm{I}_{d-1}$.

Notice that the shearlet coefficients $\mathcal{S}_\psi^\gamma f(b, s, a)$ depend on f only through its affine Radon transform $\mathcal{R}^{\mathrm{aff}} f$. Therefore Equation (21) allows to reconstruct any unknown signal $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ from its Radon transform by computing the shearlet coefficients by means of (45). Finally, it is worth observing that this reconstruction does not involve the differential operator $\mathcal{I}$ as applied to the signal. Hence, another interesting aspect of our result is that it could open the way to new methods for inverting the Radon transform, a very important issue in applications. Indeed, this result leads to an inversion formula for the Radon transform based on the shearlet and the wavelet transforms.

Example (The standard shearlet group, continued). For the classical shearlet group $\mathbb{S}^\gamma$, (36) becomes

$$\mathcal{Q} S_{b,s,a}^\gamma f(v, t) = (V_{s,a} \otimes W_{n(v) \cdot b, a}) \mathcal{Q} f(v, t), \tag{46}$$

where $S_{b,s,a}^\gamma$ is given by (23) and $V_{s,a}$ is the wavelet representation in dimension $d - 1$ as in (24). Therefore in the case of the standard shearlet group Theorem 10 shows that the unitary operator $\mathcal{Q}$ intertwines the shearlet representation S^γ with the tensor product of two wavelet representations.

For a fixed admissible vector $\psi \in L^2(\mathbb{R}^d)$ of the form (37) with $\psi_1 \in L^2(\mathbb{R})$ satisfying (39) and $\psi_2 \in L^2(\mathbb{R}^{d-1})$, equation (41) becomes

$$\mathcal{S}_\psi^\gamma f(b, s, a)$$
$$= |a|^{\frac{(d-1)(\gamma-1)}{2}} \int_{\mathbb{R}^{d-1}} \mathcal{W}_{\phi_1}(\mathcal{Q} f(v, \bullet))(n(v) \cdot b, a) \overline{\phi_2\left(\frac{v - s}{|a|^{1-\gamma}}\right)} \, \mathrm{d}v, \tag{47}$$

for any $f \in L^2(\mathbb{R}^d)$ and $(b, s, a) \in G$. Assuming that ψ_1 satisfies the additional condition (42), for any $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and $(b, s, a) \in G$, equality (45) becomes

$$\mathcal{S}_\psi^\gamma f(b, s, a)$$
$$= |a|^{\frac{(d-1)(\gamma-2)}{2}} \int_{\mathbb{R}^{d-1}} \mathcal{W}_{\chi_1}(\mathcal{R}^{\mathrm{aff}} f(v, \bullet))(n(v) \cdot b, a) \overline{\phi_2\left(\frac{v - s}{|a|^{1-\gamma}}\right)} \, \mathrm{d}v, \tag{48}$$

where χ_1 is the admissible vector defined by (44).

For the sake of clarity we write the above equation for $d = 2$ and in terms of the Radon transform in polar coordinates

$$\mathcal{S}_{\psi}^{\gamma} f(x, y, s, a)$$
$$= |a|^{\frac{\gamma-2}{2}} \int_{\mathbb{R}} \mathcal{W}_{\chi_1}\left(\mathcal{R}^{\mathrm{pol}} f(\arctan v, \frac{\bullet}{\sqrt{1+v^2}})\right)(x + vy, a) \overline{\phi_2\left(\frac{v-s}{|a|^{1-\gamma}}\right)} \frac{\mathrm{d}v}{\sqrt{1+v^2}},$$

where $x, y, s \in \mathbb{R}$, $a \in \mathbb{R}^{\times}$ and $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^2)$. As mentioned in the introduction, $\chi_1 = \mathcal{I}_0^2 \psi_1$, see (38), so that the admissible 1D-wavelet χ_1 is proportional to the Hilbert transform H of the weak derivative of ψ_1, which is the first factor of the shearlet admissible vector $\widehat{\psi}(\xi_1, \xi_2) = \widehat{\psi_1}(\xi_1)\widehat{\psi_2}(\xi_2/\xi_1)$.

Acknowledgement

Appendix A. The polar and the affine Radon transform

As mentioned in Section 2.6, the natural restriction of the Radon transform is the *polar Radon transform* $\mathcal{R}^{\mathrm{pol}} f$, which is obtained by restricting $\mathcal{R} f$ to the closed subset $S^{d-1} \times \mathbb{R}$, where S^{d-1} is the unit sphere in $\mathbb{R}^d$. Define $\Theta^{d-1} = [0, \pi]^{d-2} \times [0, 2\pi)$. For all $\theta \in \Theta^{d-1}$ we write inductively

$${}^t\theta = (\theta_1, {}^t\hat{\theta}), \qquad \theta_1 \in [0, \pi], \ \hat{\theta} \in \Theta^{d-2}$$

and then we put

$${}^t\eta(\theta) = (\cos\theta_1, \sin\theta_1 {}^t\eta(\hat{\theta})),$$

where $\eta(\hat{\theta}) \in S^{d-2}$ corresponds to the previous inductive step. Clearly, the map $\eta : \Theta^{d-1} \to S^{d-1}$ induces a parametrization of the unit sphere in $\mathbb{R}^d$. Also, observe that the map $\Theta^{d-1} \to \mathbb{P}^{d-1}$ given by $(\theta, t) \mapsto (\eta(\theta) : t)$ is a two-fold covering of $\mathbb{P}^{d-1}$.

Definition 13. Take $f \in L^1(\mathbb{R}^d)$. The *polar Radon transform* of f is the function $\mathcal{R}^{\mathrm{pol}} f : \Theta^{d-1} \times \mathbb{R} \to \mathbb{C}$ defined by

$$\mathcal{R}^{\mathrm{pol}} f(\theta, t) = \mathcal{R} f(\eta(\theta), t) = \int_{\eta(\theta)\cdot x = t} f(x) \, \mathrm{d}m(x). \tag{49}$$

It is easy to find the relation between $\mathcal{R}^{\mathrm{pol}}$ and $\mathcal{R}^{\mathrm{aff}}$. Using the parametrization η of the unit sphere, we can write any vector $n(v)$ as $n(v) = \sqrt{1 + |v|^2}\, \eta(\theta)$. More precisely, there exists ${}^t\theta = (\theta_1, {}^t\hat{\theta}) \in \Theta^{d-1}$ such that

$$(1, {}^tv) = \sqrt{1 + |v|^2}(\cos\theta_1, \sin\theta_1 {}^t\eta(\hat{\theta})). \tag{50}$$

Equality (50) holds if and only if

$$\cos\theta_1 = \frac{1}{\sqrt{1 + |v|^2}}, \qquad \sin\theta_1 \eta(\hat{\theta}) = \frac{v}{\sqrt{1 + |v|^2}}.$$

It follows that

$$\theta_1 = \arccos\left(\frac{1}{\sqrt{1 + |v|^2}}\right) \in [0, \frac{\pi}{2}), \qquad \eta(\hat{\theta}) = \frac{v}{|v|},$$

unless $v = 0$, in which case $\eta(\hat{\theta})$ can be any vector in S^{d-2}. Then, item (i) of Proposition 4 gives that

$$\mathcal{R}^{\text{aff}} f(v, t) = \frac{1}{\sqrt{1 + |v|^2}} \mathcal{R}^{\text{pol}} f\left(\theta, \frac{t}{\sqrt{1 + |v|^2}}\right) \tag{51}$$

and this is the relation that we need.

APPENDIX B. OTHER PROOFS

Proof of Theorem 8. For the reader's convenience we adapt the proof of [30] to our context. Recall that the map $\psi : \mathbb{R}^{d-1} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}^d$, defined by $\psi(v, \tau) = \tau n(v)$, is a diffeomorphism onto the open set $V = \{\xi \in \mathbb{R}^d : \xi_1 \neq 0\}$ with Jacobian $J\Phi(v, \tau) = \tau^{d-1}$. Thus, the Plancherel theorem and the Fourier slice theorem give that for any $f \in \mathcal{A}$

$$\begin{aligned}
\|f\|_2^2 &= \int_V |\mathcal{F}f(\xi)|^2 \, d\xi \\
&= \int_{\mathbb{R}^{d-1} \times (\mathbb{R} \setminus \{0\})} |\mathcal{F}f(\tau n(v))|^2 |\tau|^{d-1} \, dvd\tau \\
&= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\mathcal{F}(\mathcal{R}^{\text{aff}} f(v, \cdot))(\tau)|^2 |\tau|^{d-1} \, dvd\tau \\
&= \|\mathcal{R}^{\text{aff}} f\|_{\mathcal{D}}^2.
\end{aligned}$$

Thus, $\mathcal{R}^{\text{aff}} f$ belongs to $\mathcal{D}$ for all $f \in \mathcal{A}$ and $\mathcal{R}^{\text{aff}}$ is an isometric operator from $\mathcal{A}$ into $\mathcal{D}$. We want to prove that $\mathcal{R}^{\text{aff}} : \mathcal{A} \to \mathcal{H}$ has dense image in $\mathcal{H}$. Since $\mathcal{H}$ is the completion of $\mathcal{D}$, it is enough to prove that $\mathcal{R}^{\text{aff}}$ has dense image in $\mathcal{D}$, that is $(\text{Ran}(\mathcal{R}^{\text{aff}}))^{\perp} = \{0\}$ in $\mathcal{D}$. Take then $\varphi \in \mathcal{D}$ such that $\langle \varphi, \mathcal{R}^{\text{aff}} f \rangle_{\mathcal{D}} = 0$ for all $f \in \mathcal{A}$. By the definition of the scalar product on $\mathcal{D}$ and the Fourier slice theorem we have that

$$\begin{aligned}
\langle \varphi, \mathcal{R}^{\text{aff}} f \rangle_{\mathcal{D}} &= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1} \mathcal{F}(\varphi(v, \cdot))(\tau) \overline{\mathcal{F}(\mathcal{R}^{\text{aff}} f(v, \cdot))(\tau)} \, dvd\tau \\
&= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1} \mathcal{F}(\varphi(v, \cdot))(\tau) \overline{\mathcal{F}f(\tau n(v))} \, dvd\tau \\
&= \int_{\mathbb{R}^d} \mathcal{F}\left(\varphi\left(\frac{\tilde{\xi}}{\xi_1}, \cdot\right)\right)(\xi_1) \overline{\mathcal{F}f(\xi)} d\xi,
\end{aligned}$$

where ${}^t\xi = (\xi_1, {}^t\tilde{\xi})$. Therefore, if $\langle \varphi, \mathcal{R}^{\text{aff}} f \rangle_{\mathcal{D}} = 0$ for all $f \in \mathcal{A}$, then

$$\mathcal{F}\left(\varphi\left(\frac{\tilde{\xi}}{\xi_1}, \cdot\right)\right)(\xi_1) = 0$$

almost everywhere. However,

$$\|\varphi\|_{\mathcal{D}} = \int_{\mathbb{R}^{d-1} \times \mathbb{R}} |\tau|^{d-1} |\mathcal{F}(\varphi(v, \cdot))(\tau)|^2 \, dvd\tau = \int_{\mathbb{R}^d} |\mathcal{F}\left(\varphi\left(\frac{\tilde{\xi}}{\xi_1}, \cdot\right)\right)(\xi_1)|^2 \, d\xi$$

and hence $\varphi = 0$ in $\mathcal{D}$. Therefore $\mathcal{R}^{\text{aff}} : \mathcal{A} \to \mathcal{H}$ has dense image in $\mathcal{H}$ and we can extend it to a unique unitary operator $\mathscr{R}$ from $L^2(\mathbb{R}^d)$ onto $\mathcal{H}$. Hence, $\mathcal{Q} = \mathscr{I}\mathscr{R}$ is a unitary operator from $L^2(\mathbb{R}^d)$ onto $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$. $\square$

Proof of Proposition 9. We start by observing that (35) is true if $f \in \mathcal{A}$, by the Fourier slice theorem and by the definition of $\mathcal{Q}$. Take now $f \in L^2(\mathbb{R}^d)$. By density there exists a sequence $(f_n)_n \in \mathcal{A}$ such that $f_n \to f$ in $L^2(\mathbb{R}^d)$. Since $\mathcal{Q}$ is unitary from $L^2(\mathbb{R}^d)$ onto $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ and $\mathrm{I} \otimes \mathcal{F}$ is unitary from $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$ into itself, where I is the identity operator, $(\mathrm{I} \otimes \mathcal{F})\mathcal{Q}f_n \to (\mathrm{I} \otimes \mathcal{F})\mathcal{Q}f$ in $L^2(\mathbb{R}^{d-1} \times \mathbb{R})$. Since $f_n \in \mathcal{A}$, for almost every $(v, \tau) \in \mathbb{R}^{d-1} \times \mathbb{R}$

$$(\mathrm{I} \otimes \mathcal{F})\mathcal{Q}f_n(v, \tau) = (\mathrm{I} \otimes \mathcal{F})\mathcal{I}\mathcal{R}^{\mathrm{aff}} f_n(v, \tau)$$
$$= |\tau|^{\frac{d-1}{2}}(\mathrm{I} \otimes \mathcal{F})\mathcal{R}^{\mathrm{aff}} f_n(v, \tau)$$
$$= |\tau|^{\frac{d-1}{2}}\mathcal{F}f_n(\tau n(v)).$$

So that, passing to a subsequence if necessary,

$$|\tau|^{\frac{d-1}{2}}\mathcal{F}f_n(\tau n(v)) \to (\mathrm{I} \otimes \mathcal{F})\mathcal{Q}f(v, \tau)$$

for almost every $(v, \tau) \in \mathbb{R}^{d-1} \times \mathbb{R}$. Therefore for almost every $(v, \tau) \in \mathbb{R}^{d-1} \times \mathbb{R}$,

$$(\mathrm{I} \otimes \mathcal{F})\mathcal{Q}f(v, \tau) = \lim_{n \to +\infty} |\tau|^{\frac{d-1}{2}}\mathcal{F}f_n(\tau n(v)) = |\tau|^{\frac{d-1}{2}}\mathcal{F}f(\tau n(v)),$$

where the last equality holds true using a subsequence if necessary. $\square$

References

[1] J.-P. Antoine, R. Murenzi, Two-dimensional directional wavelets and the scale-angle representation, Signal processing 52 (3) (1996) 259–281.

[2] E. J. Candès, D. L. Donoho, Ridgelets: A key to higher-dimensional intermittency?, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 357 (1760) (1999) 2495–2509.

[3] E. J. Candès, D. L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities, Communications on pure and applied mathematics 57 (2) (2004) 219–266.

[4] K. Guo, D. Labate, W.-Q. Lim, G. Weiss, E. Wilson, Wavelets with composite dilations, Electronic research announcements of the American Mathematical Society 10 (9) (2004) 78–87.

[5] M. N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Transactions on image processing 14 (12) (2005) 2091–2106.

[6] D. Labate, W.-Q. Lim, G. Kutyniok, G. Weiss, Sparse multidimensional representation using shearlets, in: Optics & Photonics 2005, International Society for Optics and Photonics, 2005, pp. 59140U–59140U.

[7] L. Grafakos, C. Sansing, Gabor frames and directional time–frequency analysis, Applied and Computational Harmonic Analysis 25 (1) (2008) 47–67.

[8] F. De Mari, E. De Vito, Admissible vectors for mock metaplectic representations, Applied and Computational Harmonic Analysis 34 (2) (2013) 163–200.

[9] G. Kutyniok, D. Labate, Shearlets, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, New York, 2012.

[10] M. Duflo, C. C. Moore, On the regular representation of a nonunimodular locally compact group, J. Functional Analysis 21 (2) (1976) 209–243.

[11] M. A. Duval-Poo, F. Odone, E. De Vito, Edges and corners with shearlets, IEEE Transactions on Image Processing 24 (11) (2015) 3768–3780.

[12] F. Bartolucci, F. De Mari, E. De Vito, F. Odone, Shearlets and the multi-scale radon transform, Tech. rep., Università di Genova, available `http://www.dima.unige.it/~Edevito/selected.html` (2017).

[13] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, G. Teschke, The uncertainty principle associated with the continuous shearlet transform, International Journal of Wavelets, Multiresolution and Information Processing 6 (02) (2008) 157–181.

[14] G. Kutyniok, D. Labate, Resolution of the wavefront set using continuous shearlets, Trans. Amer. Math. Soc. 361 (5) (2009) 2719–2754.

[15] J. Fadili, J.-L. Starck, Curvelets and ridgelets, in: Encyclopedia of Complexity and Systems Science, Springer, 2009, pp. 1718–1738.

[16] F. Colonna, G. Easley, K. Guo, D. Labate, Radon transform inversion using the shearlet representation, Applied and Computational Harmonic Analysis 29 (2) (2010) 232–250.

[17] P. Grohs, Continuous shearlet frames and resolution of the wavefront set, Monatshefte für Mathematik 164 (4) (2011) 393–426.

[18] H. Führ, Continuous wavelet transforms with abelian dilation groups, J. Math. Phys. 39 (8) (1998) 3974–3986.

[19] H. Führ, R. R. Tousi, Simplified vanishing moment criteria for wavelets over general dilation groups, with applications to abelian and shearlet dilation groups, Applied and Computational Harmonic Analysis(2017) to appear, available at `https://arxiv.org/abs/1407.0824`.

[20] S. Dahlke, G. Steidl, G. Teschke, The continuous shearlet transform in arbitrary space dimensions, Journal of Fourier Analysis and Applications 16 (3) (2010) 340–364.

[21] V. S. Varadarajan, Lie groups, Lie algebras, and their representations, Springer, 1984.

[22] G. Alberti, S. Dahlke, F. De Mari, E. De Vito, H. Führ, Recent progress in shearlet theory: Systematic construction of shearlet dilation groups, characterization of wavefront sets, and new embeddings, in: Recent Progress in Shearlet Theory: Systematic Construction of Shearlet Dilation Groups, Characterization of Wavefront Sets, and New Embeddings, Springer, 2017, pp. 127–162, to appear, available at `arXiv:1605.02873`.

[23] S. Dahlke, F. De Mari, P. Grohs, D. Labate, From group representations to signal analysis, in: Harmonic and applied analysis, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, 2015, pp. 1–5.

[24] S. Dahlke, F. De Mari, E. De Vito, S. Häuser, G. Steidl, G. Teschke, Different faces of the shearlet group, J. Geom. Anal. 26 (3) (2016) 1693–1729.

[25] E. M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, Vol. 43, Princeton University Press, Princeton, NJ, 1993.

[26] H. Führ, Generalized Calderón conditions and regular orbit spaces, Colloq. Math. 120 (1) (2010) 103–126.

[27] P. Kittipoom, G. Kutyniok, W.-Q. Lim, Construction of compactly supported shearlet frames, Constructive Approximation 35 (1) (2012) 21–72.

[28] H. Führ, Abstract harmonic analysis of continuous wavelet transforms, Springer, 2005.

[29] A. G. Ramm, A. I. Katsevich, The Radon transform and local tomography, CRC Press, Boca Raton, FL, 1996.

[30] S. Helgason, The Radon transform, 2nd Edition, Vol. 5 of Progress in Mathematics, Birkhäuser Boston, Inc., Boston, MA, 1999.

[31] M. Reed, B. Simon, Methods of modern mathematical physics. vol. 1. Functional analysis, Academic Press, 1980.

F. Bartolucci, Dipartimento di Matematica, Università di Genova, Via Dodecaneso 35, Genova, Italy

E-mail address: `bartolucci@dima.unige.it`

F. De Mari, Dipartimento di Matematica, Università di Genova, Via Dodecaneso 35, Genova, Italy

E-mail address: `demari@dima.unige.it`

E. De Vito ,Dipartimento di Matematica, Università di Genova, Via Dodecaneso 35, Genova, Italy

E-mail address: `devito@dima.unige.it`

F. Odone, Dibris, Università di Genova, Via Dodecaneso 35, Genova, Italy

E-mail address: `odne@unige.it`