ELECTRONIC RESEARCH ANNOUNCEMENTS
OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 00, Pages 000–000 (Xxxx XX, XXXX)
S 1079-6762(XX)0000-0

SPARSE SHEARLET REPRESENTATION OF FOURIER INTEGRAL OPERATORS

KANGHUI GUO AND DEMETRIO LABATE

Abstract. Fourier Integral Operators appear naturally in a variety of problems related to hyperbolic partial differential equations. While wavelets and other traditional time-frequency methods have been successfully employed for representing many classes of singular integral operators, these methods are not equally effective in dealing with Fourier Integral Operators. In this paper, we show that the shearlets provide a very efficient tool for the analysis of a large class of Fourier Integral Operators. The shearlets, recently introduced by the authors and their collaborators, are an affine-like system of well-localized waveform at various scales, locations and orientations. It turns out that these waveforms are particularly adapted to the action of Fourier Integral Operators. In particular, we prove that the matrix representation of a Fourier Integral Operator with respect to a Parseval frame of sherlets is sparse and well-organized. This fact confirms similar results recently obtained by Candès and Demanet, pointing out to the effectiveness of appropriately constructed directional multiscale representations in dealing with operators associated with hyperbolic problems.

1. Introduction

A *Fourier Integral Operators* is an operator of the form

$$Tf(x) = \int e^{2\pi i \Phi(x,\xi)}\, \sigma(x,\xi)\, \hat{f}(\xi)\, d\xi,$$

where the *phase function* $\Phi(x,\xi)$ and the *amplitude function* $\sigma(x,\xi)$ satisfy the following assumptions:

- $\Phi(x,\xi)$ is a real-valued function, C^∞ in (x,ξ), for $\xi \neq 0$, on the support of σ, and positive homogeneous in ξ; that is, $\Phi(x,\lambda\xi) = \lambda\, \Phi(x,\xi)$, for all $\lambda > 0$.
- σ is a standard symbol of order μ, that is, σ is in C^∞ and

$$|\partial_\xi^\alpha\, \partial_x^\beta \sigma(x,\xi)| \le C_{\alpha\beta}\, (1+|\xi|)^{\mu-|\alpha|}; \tag{1.1}$$

  in addition, it is assumed that σ has compact support in the x variable.

Fourier Integral Operators appear naturally in the solution of problems arising from partial differential equations [22]. One of the simplest example is the free space wave equation in $\mathbb{R}^n$, $n > 1$:

$$\frac{\partial^2 u}{\partial t^2}(x,t) = c^2(x)\, \triangle u(x,t), \quad t > 0,$$

$$u(x,0) = u_0(x), \quad \frac{\partial u}{\partial t}(x,0) = 0.$$

For constant speed $c(x) = c$, the solution is given by the sum of two Fourier Integral Operators with phases $\Phi(x,\xi) = x\xi \pm ct|\xi|$:

$$u(x,t) = \frac{1}{2}\left(\int e^{2\pi i(x\xi + c|\xi|t)}\, \hat{u}_0(\xi)\, d\xi + \int e^{2\pi i(x\xi - c|\xi|t)}\, \hat{u}_0(\xi)\, d\xi \right).$$

Similarly, when $c(x)$ is variable and smooth, for small times, the solution operator is still the sum of two Fourier Integral Operators with more general phases and amplitudes.

Received by the editors May 8, 2007.
1991 *Mathematics Subject Classification.* 42C15, 42C40.
Key words and phrases. Affine systems, curvelets, Fourier integral operators, shearlets, sparsity, wavelets.
The second author was supported in part by National Science Foundation grant DMS 0604561.

Another important example is the solution of symmetric systems of linear hyperbolic PDEs:

$$\frac{\partial u}{\partial t} + \sum_k A_k(x)\,\frac{\partial u}{\partial x_k} + B(x)\,u = 0, \quad u(0,x) = u_0(x),$$

where u is an m–vector, $x \in \mathbb{R}^n$ and A_k, B are symmetric matrices. As shown by Lax in a classical paper which initiated the study of this subject [16], the solution to this problem can be expressed as a sum of Fourier Integral Operator, modulo a smoothing operator.

The traditional methods of time-frequency and multiscale analysis have proven to be very effective in the study of a large class of operators, including pseudodifferential operators and Calderón–Zygmung operators (see, for example, [8, 17]), but they do not apply directly to the study of Fourier Integral Operators. In this paper we show that the shearlet representation, recently introduced by the authors and they collaborators, provides an appropriate framework for the study of Fourier integral operators. More precisely, we prove that the shearlet representation of Fourier integral operators is sparse and well organized. It is important to recall that the ability to provide sparse representations of an operator has very significant implication for both its theoretical and numerical analysis. In fact, sparse representations are useful to deduce sharp estimates [20] and to design low complexity algorithms [1].

This paper is motivated and inspired by some remarkable recent results by H. Smith [19, 20] and E. Candès and L. Demanet [2], who have shown that Fourier Integral Operators can be very efficiently analyzed by combining the methods of multiscale analysis with a notion of anisotropy and directionality. In fact, it is shown in [2] that *curvelets* provide sparse representations of Fourier Integral Operators.

The shearlets, introduced by the authors and their collaborators [12, 13, 9] provide an alternative approach to the curvelet construction [3]. As will be discussed in more detail in the paper, the shearlets are a Parseval frame of well localized functions at various scales, locations and directions. Similarly to the curvelets, they have optimal approximation properties for bivariate functions with discontinuities along curves [10], and have the ability to exactly characterize the wavefront set of distributions [14]. In addition, the shearlets have the following distinctive features:

- They form an affine-like system. As a result, they are generated from the action of translation and dilation operators on a single function (while curvelets need infinitely many generators).
- They are defined on a Cartesian grid.
- They are associated with a group representation..

These properties provide a number of advantages in different theoretical and numerical applications (see further discussion in [5, 6, 13]).

The paper is organized as follows. In Section 2 we recall the construction of the shearlets. In Section 3 we describe the shearlet representation of Fourier Integral Operators and present the main result of the paper. For reasons of space, only the main ideas of the proof will be presented. We refer to [11] for the complete proofs and constructions.

1.1. **Notation and definitions.** We adopt the convention that $x \in \mathbb{R}^n$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^n$ is a row vector, i.e., $\xi = (\xi_1, \dots, \xi_n)$. A vector x multiplying a matrix $M \in GL_n(\mathbb{R})$ on the right is understood to be a column vector, while a vector ξ multiplying M on the left is a row vector. Thus, $Mx \in \mathbb{R}^n$ and $\xi M \in \widehat{\mathbb{R}}^n$. The Fourier transform is defined as $\hat f(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \xi x}\, dx$, where $\xi \in \widehat{\mathbb{R}}^n$, and the inverse Fourier transform is $\check f(x) = \int_{\widehat{\mathbb{R}}^n} f(\xi)\, e^{2\pi i \xi x}\, d\xi$.

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

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

This is equivalent to the reproducing formula $f = \sum_i \langle f, \psi_i \rangle\, \psi_i$, for all $f \in \mathcal{H}$, where the series converges in the norm of $\mathcal{H}$. This shows that a Parseval frame provides a basis-like representation even though a Parseval frame need not be a basis in general. We refer the reader to [4] for more details about frames.

2. Shearlets

The *shearlets* are the elements of the affine-like system:

$$\mathcal{A}_{AB}(\psi) = \{\psi_{j,\ell,k}(x) = |\det A|^{j/2}\, \psi(B^\ell\, A^j x - k) : \, j, \ell \in \mathbb{Z}, k \in \mathbb{Z}^2\},$$

where

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

and ψ is defined as follows. For any $\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$, $\xi_1 \neq 0$, the function ψ is given in the frequency domain by

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

where $\hat\psi_1, \hat\psi_2 \in C^\infty(\widehat{\mathbb{R}})$, $\operatorname{supp} \hat\psi_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ and $\operatorname{supp} \hat\psi_2 \subset [-1, 1]$. We assume that

$$\sum_{j \geq 0} |\hat\psi_1(2^{-j}\omega)|^2 = 1 \quad \text{for } |\omega| \geq \frac{1}{8}, \tag{2.2}$$

and, for any $j \geq 0$,

$$\sum_{\ell=-2^j}^{2^j} |\hat\psi_2(2^j\, \omega + \ell)|^2 = 1 \quad \text{for } |\omega| \leq 1. \tag{2.3}$$

There are several examples of functions ψ_1, ψ_2 satisfying the properties described above [10].

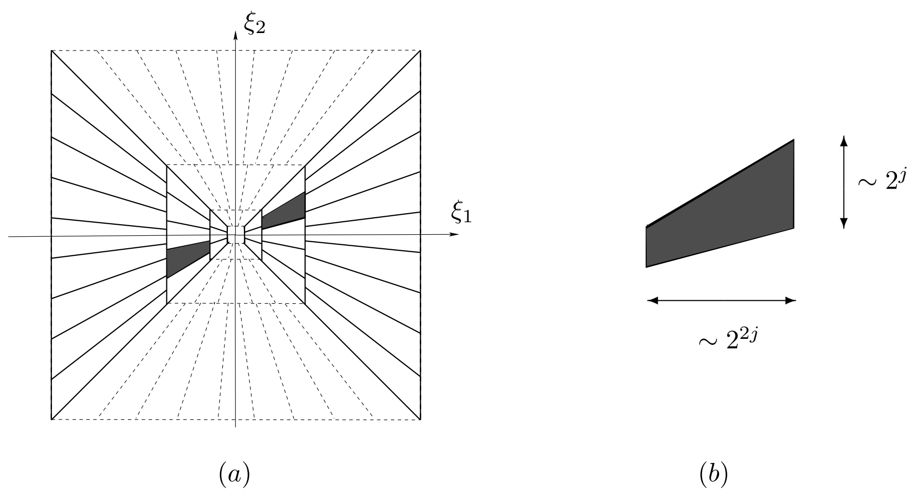

Figure 1. *(a) The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by the shearlets. (b) Frequency support of the shearlet $\psi_{j,\ell,k}$, for $\xi_1 > 0$.*

Observe that $(\xi_1, \xi_2)\, A^{-j} B^{-\ell} = (2^{-2j}\xi_1, -\ell 2^{-j}\xi_1 + 2^{-j}\xi_2)$. Thus, in the Fourier domain, the elements of the shearlet system $\psi_{j,\ell,k}$ have the form

$$\hat\psi_{j,\ell,k}(\xi) = |\det A|^{-j/2}\, \psi(\xi A^{-j} B^{-\ell})\, e^{2\pi i \xi A^{-j} B^{-\ell} k} = 3^{-3j/2}\, \hat\psi_1(2^{-2j}\, \xi_1)\, |\hat\psi_2(2^j\, \frac{\xi_2}{\xi_1} - \ell)\, e^{2\pi i \xi A^{-j} B^{-\ell} k},$$

and, as a result, the elements $\hat\psi_{j,\ell,k}$ are supported in the sets:

$$W_{j,\ell} = \{(\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], |\tfrac{\xi_2}{\xi_1} - \ell\, 2^{-j}| \leq 2^{-j}\}.$$

It follows from (2.2) and (2.3) that:

$$\sum_{j\geq 0} \sum_{\ell=-2^j}^{2^j} |\hat{\psi}(\xi\, A^{-j} B^{-\ell})|^2 = 1,$$

for $(\xi_1,\xi_2) \in \mathcal{D}_C$, where $\mathcal{D}_C = \{(\xi_1,\xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \geq \frac{1}{2}, |\frac{\xi_2}{\xi_1}| \leq 1\}$. This equation, together with the fact that $\hat{\psi}$ is supported inside $[-\frac{1}{2},\frac{1}{2}]^2$, implies that the collection of shearlets:

$$\mathcal{S}(\psi) = \{\psi_{j,\ell,k}(x) = 2^{\frac{3j}{2}}\, \psi(B^\ell A^j x - k) : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\}, \tag{2.4}$$

is a Parseval frame for $L^2(\mathcal{D}_C)^\vee = \{f \in L^2(\mathbb{R}^2) : \operatorname{supp} \hat{f} \subset \mathcal{D}_C\}$. To obtain a Parseval frame for $L^2(\mathbb{R}^2)$, one can easily construct a similar system of shearlets $\{\psi^{(v)}_{j,\ell,k}\}$ for the vertical cone $\mathcal{D}_{\widetilde{C}} = \{(\xi_1,\xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_2| \geq \frac{1}{2}, |\frac{\xi_1}{\xi_2}| \leq 1\}$. Finally, one can construct a Parseval frame (or even an orthonormal basis) for $L^2([-\frac{1}{2},\frac{1}{2}]^2)$. Then any function in $L^2(\mathbb{R}^2)$ can be expressed as a sum of 3 components $f = P_C f + P_{\widetilde{C}} f + P_0 f$, where each component corresponds to the orthogonal projection of f into one of the 3 subspaces of $L^2(\mathbb{R}^2)$ described above. The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by this system is illustrated in Figure 1(a).

As shown in Figure 1, for each j,ℓ, the set $W_{j,\ell}$ is a symmetric pair of trapezoids centered about $\pm\xi_{j,\ell}$, where $\xi_{j,\ell} = 2^{-2}(2^{2j}, \ell 2^j)$. Each trapezoid is approximately contained in a box of size $2^{2j} \times 2^j$ in the frequency domain (see Figure 1(b)), oriented along the line $\xi_2 = \ell 2^{-j} \xi_1$. Thus, the frequency support of the shearlets satisfies a *parabolic scaling*, and becomes increasingly elongated as j increases. The notion of parabolic scaling has long history in harmonic analysis, going back to the second dyadic decomposition, introduced by C. Fefferman for the study of Bochner–Riesz multipliers [7] and later used in the study of Fourier Integral Operator [18]. We recall that also curvelets satisfy a parabolic scaling. This property plays a key role in the approximation results of shearlets and their ability to deal with distributed discontinuities (see [9, 15, 10] for details).

In the following, we will use the notation $\{\psi_\mu : \mu \in \mathcal{M}\}$ and $\{\psi^{(v)}_\mu : \mu \in \mathcal{M}\}$, where $\mathcal{M} = \{\mu = (j,\ell,k) : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\}$, to denote the collection of shearlets (2.4) and the corresponding vertical shearlets, respectively.

3. Shearlet Representation of Fourier Integral Operators

The shearlets are not an orthogonal family. However, they are almost orthogonal with respect to an appropriate distance. That is, the inner product of two shearlets ψ_μ and $\psi_{\mu'}$ exhibits an almost exponential decay as a function of the *dyadic parabolic pseudo-distance.* Given a pair of indices $\mu = (j,\ell,k)$ and $\mu' = (j',\ell',k')$ in $\mathcal{M}$, the dyadic parabolic pseudo-distance $\omega(\mu,\mu')$ is defined by:

$$\omega(\mu,\mu') = 2^{|j-j'|}\left(1 + 2^{\min(j,j')}\, d(\mu,\mu')\right),$$

where

$$d(\mu,\mu') = |\ell 2^{-j} - \ell' 2^{-j'}|^2 + |k_{j,\ell} - k'_{j',\ell'}|^2 + |\langle e_\mu, k_{j,\ell} - k'_{j',\ell'}\rangle|,$$

$e_\mu = (\cos\theta_\mu, \sin\theta_\mu)$ and $\theta_\mu = \arctan(\ell 2^{-j})$. Thus, ω increases as the difference between the scales, the orientations and the locations increases. This definition is a slight variation of a similar definition introduced by Smith [19] and modified by Candès and Demanet [2].

We can now state the main theorem of this paper:

Theorem 3.1. *Let T be an Fourier Integral Operator of order 0, acting on functions on $\mathbb{R}^2$, for which the phase satisfies the nondegeneracy condition*

$$|\det \Phi_{x\xi}(x,\xi)| \geq c > 0,$$

uniformly in x and ξ, where $\Phi_{x\xi} = \nabla_x\, \nabla_\xi \Phi$. For $\mu,\mu' \in \mathcal{M}$, let $T(\mu,\mu') = \langle T\psi_\mu, \psi_{\mu'}\rangle$, where ψ_μ and $\psi_{\mu'}$ are elements of the Parseval frame of shearlets $\{\psi_\mu : \mu \in \mathcal{M}\} \cup \{\psi^{(v)}_\mu : \mu \in \mathcal{M}\}$. Then, for each $N > 0$, there is a constant $C_N > 0$ such that

$$|T(\mu,\mu')| \leq C_N\, \omega(\mu, h_{\mu'}(\mu'))^{-N}.$$

For each $\mu' \in \mathcal{M}$, the function $h_{\mu'}$ is a bijective mapping on $\mathcal{M}$, induced by the canonical transformation associated with the operator T. This function will be more precisely defined in Section 3.2. As described in the Introduction, a result similar to Theorem 3.1 is obtained by Candès and Demanet in [2]. Also notice that, using Schur's Lemma, it follows from Theorem 3.1 that, for every $0 < p \le \infty$, $T(\mu, \mu')$ is bounded from ℓ^p to ℓ^p.

To describe the main ideas of the proof, we shall start by introducing a convenient represention of T with respect to the Parseval frame of shearlets. Our construction adapts several ideas from [2].

For $j \ge 0$ and $|\ell| \le 2^j$, let

$$B_{j,\ell} = A^{-j}\, B^{-\ell}\, A^{j} = \begin{pmatrix} 1 & -\ell\, 2^{-j} \\ 0 & 1, \end{pmatrix}$$

and define the sets

$$W_j = W_{j,0} = W_{j,\ell}\, B_{j,\ell} = \{(\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}],\ |\tfrac{\xi_2}{\xi_1}| \le 2^{-j}\}.$$

Observe that $B_{j,\ell}$ is mapping $W_{j,\ell}$ into another pair of symmetric trapezoids oriented along the ξ_1 axis. For μ a fixed index triple in $\mathcal{M}$, let ψ_μ be a shearlet with location (x_μ, ξ_μ) in the phase space. Then, using change of variables $\xi = \eta\, B_{j,\ell}^{-1}$, we have:

$$T\,\psi_\mu(x) = \int e^{2\pi i \Phi(x,\xi)}\, \sigma(x,\xi)\, \hat{\psi}_\mu(\xi)\, d\xi = 2^{-3j/2} \int_{W_j} e^{2\pi i \left(\Phi(x,\eta\, B_{j,\ell}^{-1}) - \eta\, A^{-j} k\right)}\, \sigma(x, \eta\, B_{j,\ell}^{-1})\, \hat{\psi}(\eta\, A^{-j})\, d\eta.$$

It will be convenient to locally linearize the phase $\Phi(x,\xi)$ to separate the nonlinearities of ξ from those of x. This is a standard approach in the study of of Fourier Integral Operators and can be found, for example, in [21, Ch.9]. To do this, we define

$$\delta_{j,\ell}(x,\eta) = \begin{cases} \Phi(x, \eta\, B_{j,\ell}^{-1}) - \eta\, B_{j,\ell}^{-1} \cdot \nabla_\xi \Phi\left(x, (1,0)\, B_{j,\ell}^{-1}\right), & \text{for } \eta \in W_j^+; \\ \Phi(x, \eta\, B_{j,\ell}^{-1}) - \eta\, B_{j,\ell}^{-1} \cdot \nabla_\xi \Phi\left(x, (-1,0)\, B_{j,\ell}^{-1}\right), & \text{for } \eta \in W_j^-, \end{cases}$$

where

$$W_j^+ = W_j \bigcap \{(\xi_1,\xi_2) :\ \xi \ge 0\}, \quad W_j^- = W_j \bigcap \{(\xi_1,\xi_2) :\ \xi < 0\}.$$

For μ fixed, this allows us to decompose T as

$$T = T_\mu^{(2)}\, T_\mu^{(1)},$$

where

$$\begin{aligned} T_\mu^{(1)} f(x) &= \int_{W_j} e^{2\pi i \eta\, B_{j,\ell}^{-1} x}\, \beta_\mu(x,\eta)\, \hat{f}(\eta\, B_{j,\ell}^{-1})\, d\eta, \\ T_\mu^{(2)} f(x) &= f(\phi_\mu(x)), \end{aligned}$$

with $\beta_\mu(x,\eta) = e^{2\pi i \delta_{j,\ell}(\phi_\mu^{-1}(x),\eta)}\, \sigma(\phi_\mu^{-1}(x), \eta\, B_{j,\ell}^{-1})$, and $\phi_\mu(x) = \nabla_\xi \Phi\left(x, (1,0)\, B_{j,\ell}^{-1}\right)$. We can now proceed with the analysis of the operators $T_\mu^{(1)}$ and $T_\mu^{(2)}$.

3.1. **Analysis of the operator $T_\mu^{(1)}$.** The operator $T_\mu^{(1)}$ obtained from this decomposition has linear phase but is not a 'standard' pseudodifferential operator. That is, $\beta_\mu(x,\eta)$ is not a standard class symbol in general. To illustrate this point, consider the following example. Let $\Phi(\xi) = |\xi|$, for $\xi = (\xi_1, \xi_2) \neq 0$, and $\xi_\mu = 2^j\, e_\mu$, where $e_\mu = (\cos\theta_\mu, \sin\theta_\mu)$. Then

$$\nabla_\xi \Phi(\xi_\mu) = \frac{\xi_\mu}{|\xi_\mu|} = e_\mu$$

and

$$\delta_\mu(\xi) = \Phi(\xi) - \nabla_\xi \Phi(\xi_\mu)\, \xi = \Phi(\xi) - e_\mu\, \xi.$$

For $\theta_\mu = 0$, it follows that $e_\mu = (1,0)$ and

$$\delta_\mu(\xi) = |\xi| - \xi_1 = \sqrt{\xi_1^2 + \xi_2^2} - \xi_1.$$

Notice that the derivatives of $\delta_\mu(\xi)$ are homogeneous of degree 0 in ξ. Hence they exhibit no decay in ξ and, as a result, $\beta_\mu(x,\xi)$ does not satisfy (1.1) unless $\delta_\mu(x,\xi) = 0$. Also observe that $\delta_\mu(\xi)$ is unbounded in general, but is bounded for $\xi \in \operatorname{supp} \hat\psi_\mu$ (thanks to the parabolic scaling condition). This is one major reason why the shearlets (and the curvelets) are effective in dealing with the operator $T^{(1)}_\mu$.

We will show that the operator $T^{(1)}_\mu$ maps a shearlet ψ_μ into a shearlet-like function m_μ, with essentially the same phase space location. For analogy with similar notions in wavelet analysis (see the notion of vaguelettes of Coifman and Meyer [17]), the function m_μ will be referred to as a *shearlet molecule*.

Definition 3.2. *For $\mu = (j,\ell,k) \in \mathcal{M}$, the function $m_\mu(x) = 2^{3j/2}\, a_\mu(B^\ell A^j x - k)$ is an horizontal shearlet molecule with regularity R if the a_μ satisfies the following properties:*

(i) *for each $\gamma = (\gamma_1,\gamma_2) \in \mathbb{N}\times\mathbb{N}$ and each $N \ge 0$ there is a constant $C_N > 0$ independent of μ such that*

$$|\partial_x^\gamma a_\mu(x)| \le C_N\,(1+|x|)^{-N}; \tag{3.5}$$

(ii) *for each $M \le R$ and each $N \ge 0$ there is a constant $C_{N,M} > 0$ independent of μ such that*

$$|\hat a_\mu(\xi)| \le C_{N,M}\,(1+|\xi|)^{-N}\,(2^{-2j} + |\xi_1|)^M. \tag{3.6}$$

Vertical shearlet molecules are similarly defined.

The second factor in the inequality (3.6) is associated with the almost vanishing moments. By this property, the frequency support of a shearlet molecule m_μ is mostly concentrated around $|\xi| \approx 2^{2j}$. Observe that a shearlet ψ_μ is also a shearlet molecule, but a shearlet molecule has no compact support in the frequency domain, in general. Coarse scale molecules are defined as elements of the form $\{a_\mu(x-k) : k \in \mathbb{Z}^2\}$, where a_μ satisfies (3.5).

Let us examine a few implications of Definition 3.2. If $m_\mu(x)$ is an horizontal shearlet molecule with regularity R, then it follows from (3.5) that

$$|(2\pi i\xi)^\gamma\, \hat a_\mu(\xi)| \le \|\partial^\gamma a_\mu\|_{L^1} \le C_\gamma,$$

and, thus, for all $N \ge 0$ there is a constant C_N such that

$$|\hat a_\mu(\xi)| \le C_N\,(1+|\xi|)^{-N}.$$

It follows that for all $N \ge 0$ there is a constant C_N such that

$$|\hat m_\mu(\xi)| \le C_N\, 2^{-3j/2}\,(1+|\xi A^{-j} B^{-\ell}|)^{-N}. \tag{3.7}$$

On the other hand, from (3.6) it follows that for each $M \le R$ and each $N \ge 0$ there is a constant $C_{N,M} > 0$ such that

$$|\hat m_\mu(\xi)| = |\hat a_\mu(\xi\, A^{-j}\, B^{-\ell})| \le C_{N,M}\, 2^{-3j/2}\, \left\{2^{-2j}\,(1+|\xi_1|)\right\}^M\, \left(1+|\xi\, A^{-j}\, B^{-\ell}|\right)^{-N} \tag{3.8}$$

Thus, combining (3.7) and (3.8), it follows that for each $M \le R$ and each $N \ge 0$ there is a constant $C_{N,M} > 0$ such that

$$|\hat m_\mu(\xi)| \le C_{N,M}\, 2^{-3j/2}\, \min\left\{1, 2^{-2j}\,(1+|\xi_1|)\right\}^M\, \left(1+|\xi\, A^{-j}\, B^{-\ell}|\right)^{-N}.$$

We have the following result.

Theorem 3.3. *Let $\{\psi_\mu : \mu \in \mathcal{M}\}$ be a Parseval frame of shearlets. For each $\mu \in \mathcal{M}$ the operator $T^{(1)}_\mu$ maps ψ_μ into a shearlet molecule $m_\mu = T^{(1)}_\mu \psi_\mu$ with arbitrary regularity R, uniformly in μ. That is, the constant in Definition 3.2 is independent of μ. The same result holds for vertical shearlets as well.*

In addition, we have that the shearlet molecules, defined by Definition 3.2, form an almost orthogonal family with respect to the dyadic parabolic pseudo-distance ω.

Proposition 3.4. *Let m_μ and $m_{\mu'}$ be two shearlet molecules with regularity R. Let $j, j' \ge 0$. For every $N \le f(R)$, there is a constant $C_N > 0$ such that*

$$|\langle m_\mu, m_{\mu'}\rangle| \le C_N\, \omega(\mu,\mu')^{-N}.$$

The number $f(R)$ increases with R and goes to infinity as R goes to infinity.

Also in this case, the almost orthogonality extends to the vertical shearlets and to the case where one molecule is horizontal and the other one is vertical.

3.2. **Analysis of the operator $T_\mu^{(2)}$.** To analyze the operator $T_\mu^{(1)}$ in Section 3.1 we have taken advantage of the compact frequency support of the shearlets $\{\psi_\mu : \mu \in \mathcal{M}\}$. In contrast, to analyze the operator $T_\mu^{(2)}$ it will be convenient to introduce a family of shearlet-like function with compact support in the space domain. Using this analyzing family, it will be possible to introduce an atomic decomposition

$$f(x) = \sum_\mu \nu_\mu \, \rho_\mu(x),$$

for functions $f \in L^2(\mathbb{R}^2)$, where the *shearlet atoms* ρ_μ have compact support and satisfy certain regularity and vanishing moments conditions. This construction adapts several ideas from [19, 2].

We will construct a family of shearlet-like functions with compact support of the form

$$\psi_{ast}(x) = |\det A_a|^{-1/2}\, \psi(A_a^{-1}\, B_s^{-1}(x-t),$$

where

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \quad B_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix},$$

and a, s, t are continuous parameters satisfying: $0 < a \le 1$, $|s| \le 2$, $t \in \mathbb{R}$.

We say that a Schwartz function ϕ has k *vanishing moments in the x_1 direction* if ϕ can be expressed as

$$\phi(x) = \partial_{x_1}^k \tilde{\phi}(x)$$

for some other Schwartz function $\tilde{\phi}$. Observe that if ϕ has a certain number of vanishing moments in the x_1 direction, then $\hat{\phi}(\xi) = (2\pi i \xi_1)^k \hat{\tilde{\phi}}(\xi)$, and, thus, $\hat{\phi}(0, \xi_2) = 0$. This shows that $\hat{\phi}(\xi)$ is concentrated along the ξ_1 axis way from the origin. As a consequence, $\hat{\phi}(\xi B_s A_a)$ is concentrated in elongated regions (increasingly elongated as $a \to 0$), symmetric with respect to the origin, along the direction $\xi_2 = s\xi_1$.

We have the following result, which is similar to [19, Lemma 2.11].

Proposition 3.5. *Let ψ be a Schwartz obeying $\hat{\psi}(\pm 1, 0) \neq 0$ and having at least one vanishing moment in the x_1 direction. Then there is a function $q(\xi)$ such that the following formula holds*

$$q(\xi) \int_{|s|\le 2} \int_{a \le 1} a^{3/2} \left| \hat{\psi}(\xi B_s A_a) \right|^2 \frac{da}{a^3}\, ds = 1, \qquad \textit{for } \xi \in \Gamma.$$

$q(\xi)$ is a smooth function satisfying $|\partial^\alpha q(\xi)| \le C\, |\xi|^{-\frac{|\alpha|}{2}}$ on Γ, where

$$\Gamma = \left\{ (\xi_1, \xi_2) \in \mathbb{R}^2 : \ |\xi_1| \ge 1, \ \left| \frac{\xi_2}{\xi_1} \right| \le 1 \right\}. \tag{3.9}$$

There many choices of ψ satisfying the assumptions of Proposition 3.5. We will choose ψ of the form

$$\psi(x_1, x_2) = \psi_1(x_1)\, \psi_2(x_2),$$

where ψ_1, and ψ_2 are C^∞ functions with compact support, satisfying $\operatorname{supp}\psi_1, \operatorname{supp}\psi_2 \subset [0,1]$. In addition, we assume that ψ_1 has vanishing moments up to order R, that is,

$$\int_{\mathbb{R}} \psi_1(x)\, x^k\, dx = 0, \quad k = 0, 1, \dots, R.$$

We thus obtain the following reproducing formula.

Proposition 3.6. *Suppose that $\hat{f}$ vanishes outside the set Γ, given by (3.9). Then we have the reproducing formula*

$$f(x) = \int_{\mathbb{R}^2} \int_{|s|\le 2} \int_{a\le 1} \langle q(D) f, \psi_{ast} \rangle\, \psi_{ast}(x)\, \frac{da}{a^3}\, ds\, dt, \tag{3.10}$$

where $(q(D)f)^\wedge(\xi) = q(\xi)\, \hat{f}(\xi)$ and $\psi_{ast}(x) = a^{-3/4}\, \psi(A_a^{-1} B_s^{-1}(x-t))$.

The reproducing formula (3.10) can be written as an *atomic decomposition* where the integral is broken into several components associated with distinct regions. For $\mu = (j, \ell, k)$, let

$$Q_\mu = \{(a,s,t) : 2^{-2(j+1)} \le a < 2^{-2j},\ \ell 2^{-j} \le s < (\ell+1)2^{-j},\ A^{-j}B^{-\ell}t \in [k_1, k_1+1) \times [k_2, k_2+1)\}.$$

Observe that the sets Q_μ are disjoint and $\bigcup_{j\ge 0}\bigcup_{\ell=-2^{j+1}}^{2^{j+1}-1}\bigcup_{(k_a,k_2)\in\mathbb{Z}^2} Q_\mu = \{(a,s,t) : a \le 1, |s| \le 2, t \in \mathbb{R}^2\}$. Then, by breaking the integral (3.10) into components arising from different cells Q_μ we have:

$$f(x) = \sum_{j\ge 0}\sum_{\ell=-2^{j+1}}^{2^{j+1}-1}\sum_{(k_a,k_2)\in\mathbb{Z}^2} \nu_\mu\, \rho_\mu(x), \tag{3.11}$$

where

$$\rho_\mu(x) = \frac{1}{\nu_\mu}\iiint_{Q_\mu} \langle q(D)f, \psi_{ast}\rangle\, \psi_{ast}(x)\, \frac{da}{a^3}\, ds\, dt, \quad \nu_\mu = \left(\iiint_{Q_\mu} |\langle q(D)f, \psi_{ast}\rangle|^2\, \frac{da}{a^3}\, ds\, dt\right)^{1/2}.$$

We define the functions α_μ by:

$$\alpha_\mu(x) = 2^{-\frac{3}{2}j}\, \rho_\mu(A^{-j}B^{-\ell}(x+k)) = 2^{-\frac{3}{2}j}\, \rho_\mu(B_{j,\ell}A^{-j}(x+k)).$$

One can verify that they satisfy the following properties:

(i) *Compact support:* $\operatorname{supp}\alpha_\mu \subset C\,[-1,1]^2$, where C is independent of μ and f.

(ii) *Regularity:* for each $\beta = (\beta_1, \beta_2)$, there is a constant C_β independent of μ and f such that

$$\left|\partial_x^\beta \alpha_\mu(x)\right| \le C_\beta.$$

(iii) *Vanishing moments* along the x_1 direction: for all $n = 0, 1, \dots, R$,

$$\int_{\mathbb{R}} \alpha_\mu(x_1, x_2)\, x_1^n\, dx_1 = 0.$$

We will refer to all elements α_μ satisfying (i)–(iii) as *atoms* with regularity R. The corresponding functions ρ_μ, given by

$$\rho_\mu(x) = 2^{\frac{3}{2}j}\, \alpha_\mu(B^\ell A^j x - k)$$

will be referred to as *shearlet atoms* with regularity R. It is an easy exercise to verify that a shearlet atom is also a shearlet molecule, according to Definition 3.2.

We have the following result.

Theorem 3.7. *Let $\{\rho_{\mu'} : \mu' \in \mathcal{M}\}$ be a family of shearlet atoms with regularity R. For each $\mu' \in \mathcal{M}$, the operator $T_\mu^{(2)}$ maps $\rho_{\mu'}$ into a shearlet atom $m_{h_\mu(\mu')}$ with the same regularity R, uniformly over $\mu' \in \mathcal{M}$.*

As in Theorem 3.1, for each $\mu \in \mathcal{M}$, the function h_μ is a bijective mapping on $\mathcal{M}$ induced by the Canonical Transformation. To describe the function h_μ recall that the Canonical Transformation associated with the phase $\Phi(x,\xi)$ of the Fourier Integral Operator T is the map:

$$(\nabla_\xi \Phi(x,\xi), \xi) \to (x, \nabla_x \Phi(x,\xi))\,.$$

Let $k_{\mu'}$ and $\xi_{\mu'}$ be the space and frequency locations of $\rho_{\mu'}$, and define $\phi_\mu(y_{\mu,\mu'}) = k_{\mu'}$, where $\phi_\mu(x) = \nabla_\xi \Phi(x, \xi_\mu)$. Since $\Phi(x,\xi)$ is homogeneous of degree one in ξ, then $\nabla_x \Phi(x,\xi) = \xi\, \nabla_x \nabla_\xi \Phi(x,\xi)$. Using these observations we have that

$$\begin{aligned}
\nabla_\xi \Phi(k_{\mu'}, \xi_\mu) &= \phi_\mu(y_{\mu,\mu'}) = k_{\mu'},\\
\nabla_x \Phi(k_{\mu'}, \xi_\mu) &= \xi_\mu\, \nabla_x \nabla_\xi \Phi(k_{\mu'}, \xi_\mu) := \eta_{\mu,\mu'}.
\end{aligned}$$

The action of the operator $T_\mu^{(2)}$ on the phase space coordinates of the shearlet atom $\rho_{\mu'}$ corresponds to the canonical transformation:

$$(k_{\mu'}, \xi_{\mu'}) \to (y_{\mu,\mu'}, \eta_{\mu,\mu'}), \tag{3.12}$$

and the index mapping h_μ on $\mathcal{M}$ is the bijective mapping $(j_{\mu'}, \ell_{\mu'}, k_{\mu'}) \to (j_{\mu,\mu'}, \ell_{\mu,\mu'}, k_{\mu,\mu'})$ induced by the canonical transformation (3.12).

3.3. **Proof of Main Theorem.** The following properties of the psedo-distance ω follow from [2, Prop.2.2].

Proposition 3.8. *For $\mu, \mu', \mu'', \mu_0 \in \mathcal{M}$, the dyadic parabolic pseudo-distance ω satisfies the following:*

(1) *Symmetry.* $\omega(\mu, \mu') \sim \omega(\mu', \mu)$.
(2) *Triangle Inequality. There is a constant $C > 0$ such that $d(\mu, \mu') \leq C\, d(\mu, \mu'') + d(\mu'', \mu')$.*
(3) *Composition. For every $N > 0$, there is a constant $C_N > 0$ such that*

$$\sum_{\mu''} \omega(\mu, \mu'')^{-N}\, \omega(\mu'', \mu')^{-N} \leq C_N\, \omega(\mu, \mu')^{-N+1}.$$

(4) *Invariance under the bijective index mapping h_{μ_0} induced by the canonical transformation: $\omega(\mu, \mu') \sim \omega(h_{\mu_0}(\mu), h_{\mu_0}(\mu'))$, uniformly over $\mu_0 \in \mathcal{M}$.*

We can now prove Theorem 3.1.

Let ψ_{μ_0} and ψ_{μ_1} be two fixed shearlets (both horizontal, for simplicity). By Theorem 3.3, $m_{\mu_0} = T^{(1)}_{\mu_0}\, \psi_{\mu_0}$ is a shearlet molecule. Also observe that, using the atomic decomposition (3.11), we can expand ψ_{μ_1} as a superposition of shearlet atoms $\rho_{\mu'}$:

$$\psi_{\mu_1} = \sum_{\mu'} c_{\mu',\mu_1}\, \rho_{\mu'},$$

where

$$c_{\mu',\mu_1} = \left(\iiint_{Q_{\mu'}} |\langle q(D)\psi_{\mu_1}, \psi_{ast}\rangle|^2\ \frac{da}{a^3}\, ds\, dt \right)^{1/2}. \tag{3.13}$$

Therefore, using these observations and Theorem 3.7, we have that

$$\begin{aligned} \langle \psi_{\mu_1}, T\, \psi_{\mu_0}\rangle = \langle \psi_{\mu_1}, T^{(2)}_{\mu_0} T^{(1)}_{\mu_0}\, \psi_{\mu_0}\rangle &= \langle (T^{(2)}_{\mu_0})^*\, \psi_{\mu_1}, T^{(1)}_{\mu_0}\, \psi_{\mu_0}\rangle \\ &= \sum_{\mu'} c_{\mu',\mu_1}\, \langle (T^{(2)}_{\mu_0})^*\, \rho_{\mu'}, m_{\mu_0}\rangle \\ &= \sum_{\mu'} c_{\mu',\mu_1}\, \langle m_{\tilde h_{\mu_0}(\mu')}, m_{\mu_0}\rangle, \end{aligned}$$

where $m_{\tilde h(\mu')}$ is a shearlet molecule and $\tilde h = h^{-1}$ is the inverse of the bijective index mapping h.

Next observe that for every $N > 0$, there is a constant C_N such that:

$$|c_{\mu',\mu_1}| \leq C_N\, \omega(\mu', \mu_1)^{-N}.$$

This follows by discretizing the integral (3.13) and using Proposition 3.4. Finally, using Propositions 3.4 and 3.8, we have that for every $N > 0$, there is a constant C_N such that:

$$\begin{aligned} |\langle \psi_{\mu_1}, T\, \psi_{\mu_0}\rangle| &\leq \sum_{\mu'} |c_{\mu',\mu_1}|\, \left|\langle m_{\tilde h_{\mu_0}(\mu')}, m_{\mu_0}\rangle\right| \\ &\leq C_N \sum_{\mu'} \omega(\mu', \mu_1)^{-N}\, \omega(\tilde h_{\mu_0}(\mu'), \mu_0)^{-N} \\ &\leq C_N \sum_{\mu'} \omega(\tilde h_{\mu_0}(\mu'), \tilde h_{\mu_0}(\mu_1))^{-N}\, \omega(\tilde h_{\mu_0}(\mu'), \mu_0)^{-N} \\ &\leq C_N \sum_{\mu'} \omega(\tilde h_{\mu_0}(\mu_1), \tilde h_{\mu_0}(\mu'))^{-N}\, \omega(\tilde h_{\mu_0}(\mu'), \mu_0)^{-N} \\ &\leq C_N\, \omega(\tilde h_{\mu_0}(\mu_1), \mu_0)^{-N+1} \\ &\sim C_N\, \omega(\mu_1, h_{\mu_0}(\mu_0))^{-N+1}. \qquad \square \end{aligned}$$

Acknowledgments

This authors are indebted to E. Candès, L. Demanet and G. Weiss for many helpful discussions.

References

1. G. Beylkin, R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms,* Comm. Pure Appl. Math. **44** (1992), 141–183.
2. E. J. Candès, and L. Demanet, *The curvelet representation of wave propagators is optimally sparse*, Comm. Pure Appl. Math. **58** (2005), 1472–1528.
3. E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with C^2 singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.
4. O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, 2003.
5. S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, and H.-G. Stark, *The Uncertainty Principle Associated with the Continuous Shearlet Transform*, to appear in Int. J. Wavelets Multiresolut. Inf. Process., (2007).
6. G. Easley, D. Labate, and W-Q. Lim *Sparse Directional Image Representations using the Discrete Shearlet Transform*, preprint (2006).
7. C. Fefferman. *A note on spherical summation multipliers*, Israel J. Math. **15** (1973), 44-52.
8. K. Gröchenig, *Foundations of Time–Frequency Analysis*, Birkäuser, Boston (2001).
9. K. Guo, G. Kutyniok, and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators,* in: *Wavelets and Splines*, G. Chen and M. Lai (eds.), Nashboro Press, Nashville, TN (2006), 189–201.
10. K. Guo and D. Labate, *Optimally Sparse Multidimensional Representation using Shearlets*, to appear in SIAM J. Math. Anal. (2007).
11. K. Guo and D. Labate, *Representation of Fourier Integral Operators using Shearlets*, preprint (2007)
12. K. Guo, W-Q. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations*, Electr. Res. Announc. of AMS **10** (2004), 78–87.
13. K. Guo, W-Q. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilations and their MRA properties*, Appl. Comput. Harmon. Anal. **20** (2006), 220–236.
14. G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, preprint (2006).
15. D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss, *Sparse multidimensional representation using shearlets*, Wavelets XI (San Diego, CA, 2005), 254–262, SPIE Proc. **5914**, SPIE, Bellingham, WA, 2005.
16. P. Lax, *Asymptotic solutions of oscillatory initial value problems*, Duke Math. J. **24** (1957) 627–646.
17. Y. Meyer, R. Coifman *Wavelets, Calderón-Zygmund Operators and Multilinear Operators*, Cambridge Univ. Press, Cambridge, 1997.
18. A. Seeger, C. Sogge, and E. Stein, *Regularity properties of Fourier integral operators* Annals of Math. **134** (1991), 231-251.
19. H.F. Smith, *A Hardy space for Fourier integral operators*, J. Geom. Anal. **8** (1998), 629–653.
20. H.F. Smith, *A parametrix construction for wave equations with $C_{1,1}$ coefficients*, Ann. Inst. Fourier, Grenoble **48** (1998), 797–835.
21. E. M. Stein, *Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, NJ, 1993.
22. M. E. Taylor, *pseudodifferential Operators*, Princeton University Press, Princeton, NJ, 1981.

Department of Mathematics, Missouri State University, Springfield, Missouri 65804
E-mail address: KanghuiGuo@MissouriState.edu

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