Appl. Comput. Harmon. Anal. 32 (2012) 28–44

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

The nonexistence of shearlet scaling functions

Robert Houska *

Department of Mathematics, William Jewell College, Liberty, MO 64068, USA

ARTICLE INFO

Article history:
Received 14 March 2010
Revised 1 March 2011
Accepted 3 March 2011
Available online 5 March 2011
Communicated by Christopher Heil

Keywords:
Directional representation system
MRA
Multiresolution analysis
Scaling function
Shearlet

ABSTRACT

Over the past five years, the directional representation system of shearlets has received much attention and has been shown to exhibit many advantageous properties. Over this time period, there have been a number of attempts to associate shearlet systems with a multiresolution analysis (MRA). However, one can argue that, in each of these attempts, the following statement regarding the resulting shearlet MRA notion is inaccurate: “There exist scaling functions satisfying various desirable properties, such as significant amounts of decay or regularity, nonnegativity, or advantageous refinement or representation conditions. Each such scaling function naturally induces an associated shearlet (either traditional or cone-adapted) that satisfies similar desirable properties. Each such scaling function/associated shearlet pair rationally induces a fast decomposition algorithm for discrete data.” In this article, we attempt to provide explanation for this situation by arguing the great difficulty of associating shearlet systems with such an MRA. We do so by considering two very natural and general notions of shearlet MRA—one which leads to traditional shearlets and one which leads to cone-adapted shearlets—each of which seems to be an excellent candidate to satisfy the above quoted statement. For each of these notions, we prove the nonexistence of associated scaling functions satisfying the above mentioned desirable properties.

1. Introduction

Over the past approximately twenty years, wavelets have been supplied with a rich mathematical theory and have established themselves as very popular tools for a wide variety of applications. Perhaps two of the most well-known applications of wavelets are their use in the new FBI fingerprint database [1] and in JPEG 2000, the new standard for image compression [2]. There are three important general properties of wavelets which can be seen as primary contributors to this success. First, wavelets are ideally suited for the study of a certain relevant class of functions: otherwise smooth functions that exhibit point singularities, which we call type I functions. For example, discrete wavelets provide “optimally sparse” representations of type I functions—i.e., very few terms from a discrete wavelet system are necessary to approximate such functions accurately (fewer than with “any” other representation system) [3]—and continuous wavelets are able to detect the locations of the singularities of type I functions [4]. Second, a wavelet can be associated with multiresolution analysis (MRA) and a scaling function [5]. Most importantly, there exist scaling functions that are compactly supported and arbitrarily smooth [6] which are associated with wavelets satisfying similar desirable properties. Third, each scaling function/associated wavelet pair (particularly those mentioned in the previous sentence) is associated with a fast decomposition algorithm which is very useful in the compression, analysis, etc. of discrete data [7].

* Fax: +1 816 415 7612.
E-mail address: rthouska@gmail.com.

1063-5203/$ – see front matter
doi:10.1016/j.acha.2011.03.001

The great success of wavelets has resulted in their use in applications for which they are very likely far from the ideal tool. A prominent example of such an application is image analysis/compression, for which the subideality of wavelet systems manifests itself as follows: Type I functions provide a very poor model for image data, which typically possesses edges. A more realistic model of such data and one that has been generally accepted is: otherwise C^2 bivariate functions which exhibit singularities along C^2 curves [8], which we call type II functions. Wavelet systems are ill-suited for the study of type II functions. For example, the representation of type II functions by discrete wavelet systems is far from optimally sparse—i.e., a relatively large number of terms from a discrete wavelet system are necessary to approximate such functions accurately (much larger than the theoretically optimal number) [3,9]—and while continuous wavelet systems are able to detect the locations of the singularities of type II functions, they are unable to detect the wavefront set—i.e., the locations and orientations of the singularities—of such functions [10].

1.1. Shearlet systems

This realization of wavelets as subideal has resulted in the introduction—over the past five to ten years—of a number of so-called directional representation systems (e.g., shearlets [11] and curvelets [8,10]). Among all these directional representation systems, shearlets are uniquely identified by the satisfaction of a long list of advantageous properties. In particular, they are generated by applying dilations and translations to a single generating function (or finite collection of such functions); they provide optimally sparse representations of and are able to identify the wavefront set of type II functions [12, 13]; and shearlets with compact support have been constructed [14].

Shearlet systems are affine-like systems generated by applying scaling dilations, "shear" transformations, and translations either in a continuous or discrete fashion to a collection $\{\psi_n\colon\ n=1,\dots,N\}\subset L^2(\mathbb{R}^2)$, where $N\in\mathbb{Z}^+$. We are interested in discrete shearlet systems, which come in one of two varieties. Let $a_1>0$ and $0<\alpha<1$. Write $a=\operatorname{diag}(a_1,a_1^\alpha)$, $\tilde a=\operatorname{diag}(a_1^\alpha,a_1)$,

$$B=\left\{b(l)=\begin{pmatrix}1&l\\0&1\end{pmatrix}\colon l\in\mathbb{Z}\right\},\quad\text{and}\quad \widetilde{B}=\left\{\tilde b(l)=\begin{pmatrix}1&0\\l&1\end{pmatrix}\colon l\in\mathbb{Z}\right\}.$$

If $\{\psi_n\colon\ n=1,\dots,N\}\subset L^2(\mathbb{R}^2)$ and if the collection

$$\left\{D_a^{-j}D_bT_k\psi_n\colon\ j\in\mathbb{Z},\ b\in B,\ k\in\mathbb{Z}^2,\ n=1,\dots,N\right\} \tag{1.1}$$

forms a frame for $L^2(\mathbb{R}^2)$, then $\{\psi_n\colon\ n=1,\dots,N\}$ is said to be a traditional shearlet and (1.1) is said to be a traditional shearlet system (see Section 2.1 for the various notation and [15] for the definitions of Bessel system, frame, dual frame, Riesz basis, etc.). While traditional shearlet systems have a rich mathematical structure, they are strongly biased toward one axis. To provide an essentially equal treatment of all directions, one often considers the second variant of discrete shearlet systems—cone-adapted shearlet systems. Write

$$\mathcal{C}=\left\{(\xi_1,\xi_2)\in\hat{\mathbb{R}}^2\colon\ |\xi_1|\geqslant 1,\ |\xi_1|\geqslant|\xi_2|\right\},$$
$$\widetilde{\mathcal{C}}=\left\{(\xi_1,\xi_2)\in\hat{\mathbb{R}}^2\colon\ |\xi_2|\geqslant 1,\ |\xi_2|\geqslant|\xi_1|\right\},$$

and

$$\mathcal{R}=\left\{(\xi_1,\xi_2)\in\hat{\mathbb{R}}^2\colon\ |\xi_1|,\ |\xi_2|\leqslant 1\right\}.$$

If $\{\psi_n\colon\ n=1,\dots,N\}\subset L^2(\mathbb{R}^2)$ and if the collection

$$\mathcal{S}=\left\{D_a^{-j}D_{b(l)}T_k\psi_n\colon\ j\in\mathbb{N},\ l=-\lceil a_1^{(1-\alpha)j}\rceil,\dots,\lceil a_1^{(1-\alpha)j}\rceil,\ k\in\mathbb{Z}^2,\ n=1,\dots,N\right\}$$

forms a frame for $L^2(\mathcal{C})^\vee$ (in the sense of the satisfaction of a frame inequality—$\mathcal{S}$ need not belong to $L^2(\mathcal{C})^\vee$), then $\{\psi_n\colon\ n=1,\dots,N\}$ is said to be a shearlet on $\mathcal{C}$. In this case, one can easily obtain $\{\tilde\psi_n\colon\ n=1,\dots,N\}\subset L^2(\mathbb{R}^2)$ such that

$$\widetilde{\mathcal{S}}=\left\{D_{\tilde a}^{-j}D_{\tilde b(l)}T_k\tilde\psi_n\colon\ j\in\mathbb{N},\ l=-\lceil a_1^{(1-\alpha)j}\rceil,\dots,\lceil a_1^{(1-\alpha)j}\rceil,\ k\in\mathbb{Z}^2,\ n=1,\dots,N\right\}$$

forms a frame for $L^2(\widetilde{\mathcal{C}})^\vee$. Moreover, one can also easily obtain a translation-generated frame $\mathcal{B}$ for $L^2(\mathcal{R})^\vee$. Finally, one can obtain from the collections $\mathcal{S}$, $\widetilde{\mathcal{S}}$, and $\mathcal{B}$ a frame for $L^2(\mathbb{R}^2)$, and we call this frame a cone-adapted shearlet system.

1.2. Shearlet MRA and our results

Let a_1, α, a, $b(l)$ ($l\in\mathbb{Z}$), and B be as in the previous subsection. Since shearlets are generated by applying dilations and translations to a finite set of generating functions (unlike, for instance, curvelets), one would expect to be able to successfully associate them with an MRA structure. By analogy with the wavelet case, the development of a notion of shearlet MRA which satisfies the following Desirable Properties is clearly of great interest:

(D1) There exist scaling functions satisfying various desirable properties, such as significant amounts of decay or regularity, nonnegativity, or advantageous refinement or representation conditions.
(D2) Each scaling function referred to in (D1) naturally induces an associated shearlet (either traditional or cone-adapted) that satisfies similar desirable properties.
(D3) Each scaling function/associated shearlet pair referred to in (D1) and (D2) rationally induces a fast decomposition algorithm for discrete data. Here, rational is as indicated in [16]. In particular, the action of the induced fast decomposition algorithm on discrete data should mirror the action of the scaling function/associated shearlet on continuous data.

1.2.1. Traditional shearlet MRA

In [17,18], traditional shearlet systems are considered within the context of composite wavelet systems (for which a notion of MRA and scaling function exist). The following is a generalized version of the definition of traditional shearlet MRA and scaling function from [17]:

Definition 1.1. A sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed linear subspaces of $L^2(\mathbb{R}^2)$ is said to be a traditional shearlet MRA if the following conditions hold:

- $V_j \subset V_{j+1}$, for all j;
- $V_j = D_a^{-j} V_0$, for all j;
- $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$;
- $\bigcup_{j\in\mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R}^2)$;
- there exists $\varphi \in V_0$ such that

$$\{D_b T_k \varphi\colon\ b \in B,\ k \in \mathbb{Z}^2\} \tag{1.2}$$

forms a frame for V_0.

In this case, we say φ is a traditional shearlet scaling function for the given MRA.

If $\{V_j\}_{j\in\mathbb{Z}}$ is a traditional shearlet MRA with scaling function φ, write $W_0 = V_0^{\perp} \cap V_1$. If there exists $\{\psi_n\colon\ n = 1,\dots,N\} \subset W_0$ ($N \in \mathbb{Z}^+$) such that

$$\{D_b T_k \psi_n\colon\ b \in B,\ k \in \mathbb{Z}^2,\ n = 1,\dots,N\}$$

forms a frame for W_0, then it follows from the MRA properties that

$$\{D_a^{-j} D_b T_k \psi_n\colon\ j \in \mathbb{Z},\ b \in B,\ k \in \mathbb{Z}^2,\ n = 1,\dots,N\}$$

forms a frame for $L^2(\mathbb{R}^2)$; i.e., that $\{\psi_n\colon\ n = 1,\dots,N\}$ is a traditional shearlet. In this situation, we say that φ and $\{\psi_n\colon\ n = 1,\dots,N\}$ are associated.

Essentially, the only traditional shearlet/associated scaling function pairs that have been constructed are of "Shannon-type"—i.e., their Fourier transforms are characteristic functions of measurable sets. The following shearlet/associated scaling function example is adapted from [17]: Define $\varphi, \psi_n \in L^2(\mathbb{R}^2)$ ($n = 1,\dots,15$) by $\hat{\varphi} = \chi_{E_0}$ and $\hat{\psi}_n = \chi_{E_n}$, where χ_E denotes the characteristic function of E, $E_n = E_n^+ \cup E_n^-$, $E_n^- = -E_n^+$, and E_n^+ is rectangle/triangle number n in Fig. 1. It follows that φ is an orthonormal traditional shearlet scaling function and $\{\psi_n\colon\ n = 1,\dots,15\}$ is an associated orthonormal traditional shearlet with dilation matrix $a = \mathrm{diag}(4,2)$ (where the adjective "orthonormal" has the obvious meaning).

1.2.2. Cone-adapted shearlet MRA

Refs. [19–21] each contains an attempt to associate cone-adapted shearlet systems with an MRA. However, one can argue that, in each case, the resulting shearlet MRA notion fails to satisfy at least one of (D2) or (D3). We propose the following notion of shearlet MRA and scaling function which is inspired by the Unitary Extension Principle (UEP) of wavelet theory [22] (in a different fashion than in [20]):

Assume, for this discussion, that $a_1^{\alpha}, a_1^{1-\alpha} \in \mathbb{Z}$. Let $\varphi \in L^2(\mathbb{R}^2)$, and, for $j \in \mathbb{N}$, write

$$B_j = \{b(l)\colon\ -a_1^{(1-\alpha)j} + 1 \leqslant l \leqslant a_1^{(1-\alpha)j}\}$$

and define $\omega_j \in L^1(\hat{\mathbb{R}}^2)$ by

$$\omega_j(\xi) = \sum_{b\in B_j} \left|\hat{\varphi}(\xi a^{-j} b)\right|^2.$$

Since $B \subset \widetilde{SL}_2(\mathbb{Z})$, we have

$$\{D_b T_k \varphi\colon\ k \in \mathbb{Z}^2\} = \{T_k D_b \varphi\colon\ k \in \mathbb{Z}^2\}, \tag{1.3}$$

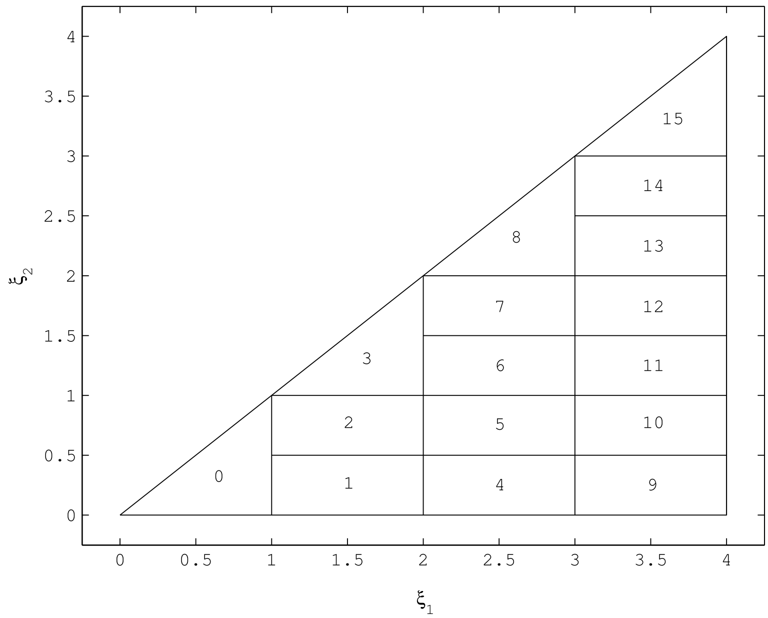

Fig. 1. The sets E_n^+ $(n = 0, \ldots, 15)$.

for all $b \in B$. If $f \in L_c^\infty(\hat{\mathbb{R}}^2)^\vee$, then Fatou's lemma, (1.3), and part (vi) of Lemma 2.1 imply that

$$\liminf_{j\to\infty} \sum_{b\in B_j,\, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-j} D_b T_k \varphi\right\rangle\right|^2 \geqslant \int_{\hat{\mathbb{R}}^2} \left|\hat{f}(\xi)\right|^2 \liminf_{j\to\infty} \omega_j(\xi)\, d\xi. \tag{1.4}$$

If $f \in L_c^\infty(\hat{\mathbb{R}}^2)^\vee$ and if $\langle \omega_j \rangle_{j\in\mathbb{N}}$ is essentially bounded on $\mathrm{supp}(\hat{f})$, then the Dominated Convergence Theorem, (1.3), and part (vi) of Lemma 2.1 imply that

$$\limsup_{j\to\infty} \sum_{b\in B_j,\, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-j} D_b T_k \varphi\right\rangle\right|^2 \leqslant \int_{\hat{\mathbb{R}}^2} \left|\hat{f}(\xi)\right|^2 \limsup_{j\to\infty} \omega_j(\xi)\, d\xi. \tag{1.5}$$

Let $N \in \mathbb{Z}^+$ and let Γ be a collection of the $a_1^{1+\alpha}$ distinct representatives of $\hat{\mathbb{Z}}^2/\hat{\mathbb{Z}}^2 a$ with $0 \in \Gamma$. Suppose $\{m_p^n\colon\ p = 0, \ldots, a_1^{1-\alpha} - 1,\ n = 0, \ldots, N\} \subset L^2(\hat{\mathbb{T}}^2)$ satisfy

$$\hat{\varphi}(\xi a) = \sum_{p=0}^{a_1^{1-\alpha}-1} m_p^0(\xi)\hat{\varphi}\big(\xi b(-p)\big),$$

for a.e. ξ, and

$$\sum_{n=0}^{N} m_p^n(\xi)\overline{m_{p'}^n\big(\xi + \gamma a^{-1}\big)} = \begin{cases} 1, & \text{if } p = p' \text{ and } \gamma = 0, \\ 0, & \text{otherwise}, \end{cases}$$

for a.e. ξ, all $p, p' \in \{0, \ldots, a_1^{1-\alpha} - 1\}$, and all $\gamma \in \Gamma$. Define $\{\psi_n\colon\ n = 1, \ldots, N\} \subset L^2(\mathbb{R}^2)$ by

$$\hat{\psi}_n(\xi a) = \sum_{p=0}^{a_1^{1-\alpha}-1} m_p^n(\xi)\hat{\varphi}\big(\xi b(-p)\big),$$

for a.e. ξ. It follows from standard UEP-type arguments that

$$\sum_{b\in B_{j+1},\, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-(j+1)} D_b T_k \varphi\right\rangle\right|^2 = \sum_{b\in B_j,\, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-j} D_b T_k \varphi\right\rangle\right|^2 + \sum_{n=1,\ldots,N,\, b\in B_j,\, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-j} D_b T_k \psi_n\right\rangle\right|^2, \tag{1.6}$$

for all $j \in \mathbb{N}$ and all $f \in L_c^\infty(\hat{\mathbb{R}}^2)^\vee$.

Suppose that $\langle\omega_j\rangle_{j\in\mathbb{N}}$ is essentially bounded on compact subsets of $\mathcal{C}$ and that

$$C \leqslant \liminf_{j\to\infty} \omega_j(\xi) \leqslant \limsup_{j\to\infty} \omega_j(\xi) \leqslant D,$$

for a.e. $\xi \in \mathcal{C}$. Using (1.4), (1.5), (1.6), and that $L_c^\infty(\mathcal{C})^\vee$ is dense in $L^2(\mathcal{C})^\vee$, it follows that

$$\{D_b T_k \varphi\colon b \in B_0,\ k \in \mathbb{Z}^2\} \cup \{D_a^{-j} D_b T_k \psi_n\colon j \in \mathbb{N},\ b \in B_j,\ k \in \mathbb{Z}^2,\ n = 1,\ldots,N\}$$

forms a frame (with constants $C \leqslant D$) for $L^2(\mathcal{C})^\vee$; i.e., that $\{\psi_n\colon n = 1,\ldots,N\}$ is a shearlet on $\mathcal{C}$ (essentially). In this case, we call φ a shearlet scaling function on $\mathcal{C}$ and say that φ and $\{\psi_n\colon n = 1,\ldots,N\}$ are associated. We note that, with φ and $\psi_1,\ldots,\psi_{15}$ as defined above, φ is a shearlet scaling function on $\mathcal{C}$ (with $C = D = 1$ and $a = \mathrm{diag}(4,2)$) and $\psi_1,\ldots,\psi_{15}$ is an associated shearlet on $\mathcal{C}$.

1.2.3. Our results

The notions of traditional shearlet scaling function and shearlet scaling function on $\mathcal{C}$ introduced above are both very natural and general, and each seems to be an excellent candidate to satisfy the Desirable Properties, particularly (D2) and (D3). However, our main results (Theorems 3.1, 3.2, 3.3, 4.1, and 4.2) show that both definitions fail completely in the satisfaction of (D1).

2. Preliminaries

In this section, we collect the various notation and results that are used throughout the article.

2.1. Notation

Throughout, a_1, α, a, $b(l)$ ($l \in \mathbb{Z}$), and B are as defined in Section 1.1 and φ denotes an element of $L^2(\mathbb{R}^2)$. Unless indicated otherwise, all norms are assumed to be L^2-norms.

We represent elements of the time domain, $\mathbb{R}^2$, by column vectors $x = \binom{x_1}{x_2}$ and elements of the frequency domain, $\hat{\mathbb{R}}^2$, by row vectors $\xi = (\xi_1, \xi_2)$. We use $\mathbb{N}$ and $\mathbb{Z}^+$ to denote $\{p \in \mathbb{Z}\colon p \geqslant 0\}$ and $\{p \in \mathbb{Z}\colon p > 0\}$, respectively. $\widetilde{SL}_2(\mathbb{Z})$ denotes the collection of all $c \in GL_2(\mathbb{Z})$ such that $|\det c| = 1$. If E is a measurable subset of $\hat{\mathbb{R}}^2$, $L_c^\infty(E)$ denotes the collection of all $f \in L^\infty(\hat{\mathbb{R}}^2)$ such that f is compactly supported and $\mathrm{supp}(f) \subset E$.

We use the Fourier transform $\mathcal{F} : L^2(\mathbb{R}^2) \to L^2(\hat{\mathbb{R}}^2)$ defined for $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ by

$$\mathcal{F}f(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^2} f(x) e^{-2\pi\imath\xi x}\,dx.$$

If $f \in L^2(\mathbb{R}^2)$, we denote $\mathcal{F}^{-1}f$ by $\check{f}$. For each $1 \leqslant p \leqslant \infty$ and each $y \in \mathbb{R}^2$, we define the translation operator $T_y : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$ by $T_y f(x) = f(x - y)$. If $\eta \in \hat{\mathbb{R}}^2$, $T_\eta : L^p(\hat{\mathbb{R}}^2) \to L^p(\hat{\mathbb{R}}^2)$ is defined similarly. For each $c \in GL_2(\mathbb{R})$, we define the dilation operator $D_c : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ by $D_c f(x) = |\det c|^{-1/2} f(c^{-1}x)$.

2.2. Shift invariant spaces

We require some basic terminology and results from shift invariant space theory. If $f, g \in L^2(\hat{\mathbb{R}}^2)$, the bracket product of f and g, denoted by $[f, g]$, is defined by

$$[f, g](\xi) = \sum_{k \in \hat{\mathbb{Z}}^2} f(\xi + k)\overline{g(\xi + k)},$$

for a.e. ξ, where $\hat{\mathbb{Z}}^2$ denotes the collection of all 1×2 row vectors with integer entries. It is straightforward to verify that the above sum converges absolutely for a.e. ξ and that $[f, g] \in L^1(\hat{\mathbb{T}}^2)$, where, for $1 \leqslant p \leqslant \infty$, $L^p(\hat{\mathbb{T}}^2)$ denotes the space of all measurable $\hat{\mathbb{Z}}^2$-periodic functions $f : \hat{\mathbb{R}}^2 \to \mathbb{C}$ satisfying $\|f\|_{L^p(\hat{\mathbb{T}}^2)} < \infty$, where

$$\|f\|_{L^p(\hat{\mathbb{T}}^2)} = \begin{cases} (\int_{[0,1]^2} |f(\xi)|^p\,d\xi)^{1/p}, & \text{if } 1 \leqslant p < \infty, \\ \operatorname{ess\,sup}_{\xi\in[0,1]^2} |f(\xi)|, & \text{if } p = \infty. \end{cases}$$

We have the following collection of basic results (see, for instance, [23–25] and [15]):

Lemma 2.1. *Let $f, g \in L^2(\mathbb{R}^2)$ and let $\{\varphi_i\colon i \in I\} \subset L^2(\mathbb{R}^2)$, where I is a countable indexing set. Denote the closed linear spans of the collections $\{T_k f\colon k \in \mathbb{Z}^2\}$, $\{T_k g\colon k \in \mathbb{Z}^2\}$, and $\Phi = \{T_k\varphi_i\colon k \in \mathbb{Z}^2, i \in I\}$ by X, Y, and Z, respectively.*

(i) *$X \perp Y$ if and only if $[\hat{f}, \hat{g}](\xi) = 0$, for a.e. ξ.*
(ii) *If X and Y are Bessel systems, then $[\hat{f}, \hat{g}] \in L^\infty(\hat{\mathbb{T}}^2)$.*
(iii) *If $m \in L^\infty(\hat{\mathbb{T}}^2)$, then $m\hat{f} \in \hat{X}$.*
(iv) *If Φ is a Bessel system with constant D, then*

$$\sum_{i\in I} \big\| [\hat{f}, \hat{\varphi}_i] \big\|^2_{L^2(\hat{\mathbb{T}}^2)} \leqslant D \|f\|^2.$$

(v) *If Φ forms a frame for Z with dual frame $\{T_k\theta_i:\ k \in \mathbb{Z}^2,\ i \in I\}$ and if $f \in Z$, then $\hat{f} = \sum_{i\in I} [\hat{f}, \hat{\theta}_i]\hat{\varphi}_i$, with unconditional convergence in $L^2(\hat{\mathbb{R}}^2)$.*
(vi) *If $f \in L^\infty_c(\hat{\mathbb{R}}^2)^\vee$, there exists $J = J(f) \in \mathbb{N}$ such that*

$$\sum_{k\in\mathbb{Z}^2} \big|\langle f, D_a^{-j} T_k h\rangle\big|^2 = \int_{\hat{\mathbb{R}}^2} \big|\hat{f}(\xi)\big|^2 \big|\hat{h}\big(\xi a^{-j}\big)\big|^2 \, d\xi,$$

for all $j \geqslant J$ and all $h \in L^2(\mathbb{R}^2)$.

2.3. Decay, regularity, and the Fourier transform

We require two somewhat nonstandard results regarding the connection, via the Fourier transform, between decay and regularity. We have the following definition:

Definition 2.1. Let $\gamma \in [0, 1)$ and $p \in \{1, 2\}$.

- The space $\mathcal{H}_p^\gamma(\hat{\mathbb{R}}^2)$ consists of all $f \in L^\infty(\hat{\mathbb{R}}^2)$ such that
$$\|f - T_{t\hat{e}_p} f\|_{L^\infty(\hat{\mathbb{R}}^2)} \leqslant M(t)|t|^\gamma,$$
for all $t \in \mathbb{R}$, where $M = M(f)$ is bounded and satisfies $M(t) \to 0$, as $t \to 0$ and $\hat{e}_p$ denotes the pth canonical basis vector of $\hat{\mathbb{R}}^2$.
- The space $\mathcal{H}^\gamma(\hat{\mathbb{R}}^2)$ consists of all $f \in L^\infty(\hat{\mathbb{R}}^2)$ such that
$$\|f - T_\eta f\|_{L^\infty(\hat{\mathbb{R}}^2)} \leqslant M(\eta)\|\eta\|^\gamma,$$
for all $\eta \in \hat{\mathbb{R}}^2$, where $M = M(f)$ is bounded and satisfies $M(\eta) \to 0$, as $\eta \to 0$.

If $0 < \gamma \leqslant 1$, we recall that $f \in L^\infty(\mathbb{R}^2)$ is said to be Hölder continuous with exponent γ if there exists $0 < K < \infty$ such that $\|f - T_y f\|_{L^\infty(\mathbb{R}^2)} \leqslant K\|y\|^\gamma$, for all $y \in \mathbb{R}^2$. We have the following result:

Lemma 2.2.

(i) *If $\gamma \in [0, 1)$ and $(1 + \|x\|^\gamma) f \in L^1(\mathbb{R}^2)$ (where x denotes the identity function on $\mathbb{R}^2$), then $\hat{f} \in \mathcal{H}^\gamma(\hat{\mathbb{R}}^2)$.*
(ii) *If $f \in L^2(\mathbb{R}^2)$ is compactly supported and Hölder continuous with exponent $\gamma > 1/2$, then $|\xi_2|^\gamma \hat{f} \in \mathcal{H}_1^{1/2}(\hat{\mathbb{R}}^2)$, where ξ_2 denotes the second coordinate function of $\hat{\mathbb{R}}^2$.*

Proof. To prove (i), let $\gamma \in [0, 1)$ and suppose that $(1 + \|x\|^\gamma) f \in L^1(\mathbb{R}^2)$. Since $f \in L^1(\mathbb{R}^2)$, the Riemann–Lebesgue lemma implies that $\hat{f} \in L^\infty(\hat{\mathbb{R}}^2)$. Using again that $f \in L^1(\mathbb{R}^2)$, we have

$$\big|\hat{f}(\xi) - T_\eta \hat{f}(\xi)\big| = \big|\big(\big(1 - e^{2\pi\imath\eta\cdot}\big)f\big)^\wedge(\xi)\big| \leqslant \int_{\mathbb{R}^2} \big|1 - e^{2\pi\imath\eta x}\big|\big|f(x)\big|\, dx,$$

for all ξ and η, implying that

$$\|\hat{f} - T_\eta \hat{f}\|_{L^\infty(\hat{\mathbb{R}}^2)} \leqslant M(\eta)\|\eta\|^\gamma, \tag{2.1}$$

for all $\eta \neq 0$, where

$$M(\eta) = \int_{\mathbb{R}^2} \|\eta\|^{-\gamma} \big|1 - e^{2\pi\imath\eta x}\big|\big|f(x)\big|\, dx.$$

Using that

$$\big|1 - e^{ix}\big| \leqslant |x|, \tag{2.2}$$

it follows easily that

$$\left|1-e^{ix}\right| \leqslant 2|x|^{\gamma}, \tag{2.3}$$

for all $x \in \mathbb{R}$. Using this observation and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\|\eta\|^{-\gamma}\left|1-e^{2\pi\imath\eta x}\right|\left|f(x)\right| &\leqslant \|\eta\|^{-\gamma}2|2\pi\eta x|^{\gamma}\left|f(x)\right| \\ &\leqslant \|\eta\|^{-\gamma}2^{\gamma+1}\pi^{\gamma}\|\eta\|^{\gamma}\|x\|^{\gamma}\left|f(x)\right| \\ &= 2^{\gamma+1}\pi^{\gamma}\|x\|^{\gamma}\left|f(x)\right|,\end{aligned} \tag{2.4}$$

for a.e. x and all $\eta \neq 0$. Using (2.2), the Cauchy–Schwarz inequality, and that $\gamma < 1$, we have

$$\begin{aligned}\|\eta\|^{-\gamma}\left|1-e^{2\pi\imath\eta x}\right|\left|f(x)\right| &= \|\eta\|^{1-\gamma}\|\eta\|^{-1}\left|1-e^{2\pi\imath\eta x}\right|\left|f(x)\right| \\ &\leqslant \|\eta\|^{1-\gamma}\|\eta\|^{-1}|2\pi\eta x|\left|f(x)\right| \\ &\leqslant \|\eta\|^{1-\gamma}\|\eta\|^{-1}2\pi\|\eta\|\|x\|\left|f(x)\right| \\ &= 2\pi\|\eta\|^{1-\gamma}\|x\|\left|f(x)\right| \\ &\to 0,\end{aligned} \tag{2.5}$$

as $\eta \to 0$, for a.e. x. Using (2.4), that $\|x\|^{\gamma}f \in L^1(\mathbb{R}^2)$, (2.5), and the Dominated Convergence Theorem, it follows that M is bounded and that $M(\eta) \to 0$, as $\eta \to 0$. In conjunction with (2.1), this proves (i).

To prove (ii), suppose that $f \in L^2(\mathbb{R}^2)$ is compactly supported and Hölder continuous with exponent $\gamma > 1/2$. For $s \in \mathbb{R}$, write $E_s = \operatorname{supp}(T_{se_2}f) = E_0 + se_2$. Define $g : \hat{\mathbb{R}}^2 \to \mathbb{C}$ by $g(\xi_1,\xi_2) = |\xi_2|^{\gamma}\hat{f}(\xi_1,\xi_2)$. If $h \in L^1(\mathbb{R}^2)$, we have

$$|s|^{-\gamma}\left|\left(1-e^{-2\pi\imath\xi_2 s}\right)\hat{h}(\xi)\right| = |s|^{-\gamma}\left|(h-T_{se_2}h)\widehat{\ }(\xi)\right| \leqslant M(h)(s),$$

where

$$M(h)(s) = |s|^{-\gamma}\int_{\mathbb{R}^2}\left|h(x)-T_{se_2}h(x)\right|dx,$$

for all $s \neq 0$ and all $\xi = (\xi_1,\xi_2)$. Substituting $1/(2\xi_2)$ for s in the above inequality, we obtain

$$2^{\gamma+1}|\xi_2|^{\gamma}\left|\hat{h}(\xi)\right| = \left|1/(2\xi_2)\right|^{-\gamma}\left|\left(1-e^{-2\pi\imath\xi_2/(2\xi_2)}\right)\hat{h}(\xi)\right| \leqslant M(h)\left(1/(2\xi_2)\right), \tag{2.6}$$

for all $\xi = (\xi_1,\xi_2)$ with $\xi_2 \neq 0$.

Since $f \in L^1(\mathbb{R}^2)$, (2.6) implies that $|g(\xi)| = |\xi_2|^{\gamma}|\hat{f}(\xi)| \leqslant M(f)(1/(2\xi_2))$, for all $\xi = (\xi_1,\xi_2)$ with $\xi_2 \neq 0$, where

$$\begin{aligned}M(f)(s) &= |s|^{-\gamma}\int_{\mathbb{R}^2}\left|f(x)-T_{se_2}f(x)\right|dx \\ &= |s|^{-\gamma}\int_{E_0\cup E_s}\left|f(x)-T_{se_2}f(x)\right|dx \\ &\leqslant |s|^{-\gamma}\int_{E_0\cup E_s}K_1|s|^{\gamma}\,dx \\ &\leqslant 2K_1|E_0|,\end{aligned}$$

for all $s \neq 0$, where $0 < K_1 < \infty$ is the Hölder constant of f. It follows that $g \in L^{\infty}(\hat{\mathbb{R}}^2)$. Using again that $f \in L^1(\mathbb{R}^2)$ and (2.6), we obtain

$$\begin{aligned}\left|g(\xi)-T_{t\hat{e}_1}g(\xi)\right| &= |\xi_2|^{\gamma}\left|\hat{f}(\xi)-T_{t\hat{e}_1}\hat{f}(\xi)\right| \\ &= |\xi_2|^{\gamma}\left|\left(\left(1-e^{2\pi\imath t\hat{e}_1\cdot}\right)f\right)\widehat{\ }(\xi)\right| \\ &\leqslant M\left(\left(1-e^{2\pi\imath t\hat{e}_1\cdot}\right)f\right)\left(1/(2\xi_2)\right),\end{aligned} \tag{2.7}$$

for all $\xi = (\xi_1,\xi_2)$ with $\xi_2 \neq 0$. Choose $0 < K_2 < \infty$ such that $|\hat{e}_1 x| \leqslant K_2$, for all $x \in \bigcup_{s\in\mathbb{R}} E_s$. Using (2.2) and (2.3), it follows that

$$
\begin{aligned}
M\big(\big(1-e^{2\pi\imath t\hat{e}_1\cdot}\big)f\big)(s) &= |s|^{-\gamma}\int_{\mathbb{R}^2}\big|1-e^{2\pi\imath t\hat{e}_1x}\big|\big|f(x)-T_{se_2}f(x)\big|\,dx \\
&= |s|^{-\gamma}\int_{E_0\cup E_s}\big|1-e^{2\pi\imath t\hat{e}_1x}\big|\big|f(x)-T_{se_2}f(x)\big|\,dx \\
&\leqslant \begin{cases} |s|^{-\gamma}\int_{E_0\cup E_s}|2\pi t\hat{e}_1x|K_1|s|^{\gamma}\,dx, & \text{if } |t|\leqslant 1, \\ |s|^{-\gamma}\int_{E_0\cup E_s}2|2\pi t\hat{e}_1x|^{1/2}K_1|s|^{\gamma}\,dx, & \text{if } |t|>1, \end{cases} \\
&\leqslant \begin{cases} 4\pi K_1K_2|E_0||t|, & \text{if } |t|\leqslant 1, \\ 4K_1(2\pi K_2)^{1/2}|E_0||t|^{1/2}, & \text{if } |t|>1, \end{cases}
\end{aligned}
$$

for all $s \neq 0$. Part (ii) of this lemma follows from (2.7) and the above inequality. □

3. Cone-adapted shearlet scaling functions

We use the following notation in this section: Let $L_0^-, L_0^+, L_1^-, L_1^+, \ldots \in \mathbb{Z}^+$ and $\beta, K \in (0,\infty)$ satisfy

$$\lim_{j\to\infty} a_1^{-\beta j}L_j^- = \lim_{j\to\infty} a_1^{-\beta j}L_j^+ = K.$$

For $j \in \mathbb{N}$, define $B_j = \{b(l)\colon\ l = -L_j^-, \ldots, L_j^+\}$ and $\omega_j \in L^1(\hat{\mathbb{R}}^2)$ by

$$\omega_j(\xi) = \sum_{b\in B_j}\big|\hat{\varphi}\big(\xi a^{-j}b\big)\big|^2.$$

In this section, we prove three results (Theorems 3.1–3.3) regarding the convergence of $\langle\omega_j\rangle_{j\in\mathbb{N}}$. The case $\beta = 1-\alpha < 1$ is relevant to shearlet scaling functions on $\mathcal{C}$ and, as we see in Section 4, the case $\beta > 1$ is relevant to traditional shearlet scaling functions. Moreover, one can imagine more general notions of cone-adapted shearlet and cone-adapted shearlet scaling function that correspond to α and β such that $1-\alpha \neq \beta \leqslant 1$; we therefore study these cases as well. We have the following results:

Theorem 3.1. *Assume $\beta < 1$ and write $\gamma = \max\{\beta/(2\alpha), \beta/(2-2\beta)\}$. Suppose $(1+\|x\|^{\gamma})\varphi \in L^1(\mathbb{R}^2)$.*

(i) *If $\partial_2^p\hat{\varphi}(0) \neq 0$, for some $0 \leqslant p < \gamma$, then $\lim_{j\to\infty}\omega_j(\xi) = \infty$, for a.e. ξ.*
(ii) *If $\gamma \in \mathbb{Z}$, if $\partial_2^p\hat{\varphi}(0) = 0$, for all $0 \leqslant p < \gamma$, and if $\partial_2^{\gamma}\hat{\varphi}(0) \neq 0$, then $\langle\omega_j\rangle_{j\in\mathbb{N}}$ converges uniformly to P on compact subsets of $\hat{\mathbb{R}}^2$, where $P : \hat{\mathbb{R}}^2 \to [0,\infty)$ is defined by*

$$P(\xi) = \begin{cases} \frac{2K|\partial_2^{\gamma}\hat{\varphi}(0)|^2}{\gamma!^2}\xi_2^{2\gamma}, & \text{if } \alpha+\beta<1, \\ \frac{2|\partial_2^{\gamma}\hat{\varphi}(0)|^2}{\gamma!^2}\sum_{p=0}^{\gamma}\frac{K^{2p+1}}{2p+1}\binom{2\gamma}{2p}\xi_1^{2p}\xi_2^{2\gamma-2p}, & \text{if } \alpha+\beta=1, \\ \frac{2K^{2\gamma+1}|\partial_2^{\gamma}\hat{\varphi}(0)|^2}{\gamma!^2(2\gamma+1)}\xi_1^{2\gamma}, & \text{if } \alpha+\beta>1, \end{cases}$$

for all $\xi = (\xi_1, \xi_2)$.
(iii) *If $\partial_2^p\hat{\varphi}(0) = 0$, for all $0 \leqslant p \leqslant \gamma$, then $\langle\omega_j\rangle_{j\in\mathbb{N}}$ converges uniformly to 0 on compact subsets of $\hat{\mathbb{R}}^2$.*

Theorem 3.2. *Suppose that $\beta = 1$ and that $\varphi \in L^1(\mathbb{R}^2)$. Write $\mathcal{S} = \{|\xi_2|\colon\ \hat{\varphi}(0,\xi_2) \neq 0\}$ and, if $\mathcal{S} \neq \emptyset$, write $S = \inf\mathcal{S}$.*

(i) *If $S = 0$, then $\lim_{j\to\infty}\omega_j(\xi) = \infty$, for a.e. ξ.*
(ii) *If $S > 0$, then $\lim_{j\to\infty}\omega_j(\xi) = \infty$, for all (ξ_1,ξ_2) with $|\xi_1| > S/K$. If, in addition, $\|x\|^{1/2}\varphi \in L^1(\mathbb{R}^2)$, then $\langle\omega_j\rangle_{j\in\mathbb{N}}$ converges uniformly to 0 on compact subsets of $\{(\xi_1,\xi_2)\colon\ |\xi_1| < S/K\}$.*
(iii) *If $\mathcal{S} = \emptyset$ and if $\|x\|^{1/2}\varphi \in L^1(\mathbb{R}^2)$, then $\langle\omega_j\rangle_{j\in\mathbb{N}}$ converges uniformly to 0 on compact subsets of $\hat{\mathbb{R}}^2$.*

Theorem 3.3. *Suppose $\varphi \in L^1(\mathbb{R}^2)$.*

(i) *If $\hat{\varphi}(0,\eta_2) \neq 0$, for some η_2, and if $\beta > 1$, then $\lim_{j\to\infty}\omega_j(\xi) = \infty$, for a.e. ξ.*
(ii) *If $\hat{\varphi}(0,\xi_2) = 0$, for all ξ_2, if $\beta < 2$, and if $\|x\|^{\beta/2}\varphi \in L^1(\mathbb{R}^2)$, then $\langle\omega_j\rangle_{j\in\mathbb{N}}$ converges uniformly to 0 on compact subsets of $\hat{\mathbb{R}}^2$.*
(iii) *If $\hat{\varphi}(0,\xi_2) = 0$, for all ξ_2, and if one of the following two conditions holds:*

- *φ is compactly supported and Hölder continuous with exponent $\gamma > 1/2$;*
- *$\varphi \in C_0(\mathbb{R}^2) \cap C^1(\mathbb{R}^2)$ and $(1 + \|x\|^{1/2})\partial^\nu \varphi \in L^1(\mathbb{R}^2)$, for all multi-indices ν with $|\nu| \leqslant 1$;*

then $\langle \sum_{b\in B} |\hat{\varphi}(\cdot a^{-j}b)|^2\rangle_{j\in\mathbb{N}}$ converges uniformly to 0 on compact subsets of $\hat{\mathbb{R}}^2$. In particular, the same is true of $\langle \omega_j\rangle_{j\in\mathbb{N}}$.

We make the following remarks regarding Theorems 3.1–3.3: First, in part (ii) of Theorem 3.2, more can be said regarding the convergence of $\langle \omega_j\rangle_{j\in\mathbb{N}}$ near $\{(\xi_1, \xi_2)\colon |\xi_1| = S/K\}$. However, since the case $\beta = 1$ is of least interest, we investigate this case no further. Second, if $\varphi \in L^1(\mathbb{R}^2)$ and $\varphi(x) \geqslant 0$, for a.e. x, then

$$\hat{\varphi}(0) = \int_{\mathbb{R}^2} \varphi(x)\, dx > 0,$$

provided $\varphi \neq 0$. This comment is relevant, in particular, to φ of Haar-type or of nonnegative spline-type. Finally, we note that, combined with the preceding remark and the various observations of Section 1.2.2, Theorems 3.1–3.3 provide very strong evidence against the existence of shearlet scaling functions on $\mathcal{C}$ (and any other notion of cone-adapted shearlet scaling function based on similar ideas) satisfying (D1). We now prove Theorem 3.1.

Proof. Since $(1 + \|x\|^\gamma)\varphi \in L^1(\mathbb{R}^2)$, standard results regarding the connection, via the Fourier transform, between decay and regularity (e.g. Theorem 8.22 of [26]) together with part (i) of Lemma 2.2 imply that $\hat{\varphi} \in C^{\lfloor\gamma\rfloor}(\hat{\mathbb{R}}^2)$ and that $\partial_2^{\lfloor\gamma\rfloor}\hat{\varphi} \in \mathcal{H}^{\gamma-\lfloor\gamma\rfloor}(\hat{\mathbb{R}}^2)$.

Proof of parts (i) and (ii). Suppose that, for some $0 \leqslant p \leqslant \gamma$, we have $\partial_2^q\hat{\varphi}(0) = 0$, for all $q = 0, \dots, p-1$, and $\partial_2^p\hat{\varphi}(0) \neq 0$. Using Taylor's theorem, it follows that

$$\left|\hat{\varphi}(\xi)\right|^2 = E(\xi)\xi_1 + F(\xi)\xi_2^{2p}, \tag{3.1}$$

for all $\xi = (\xi_1, \xi_2)$, where $E : \hat{\mathbb{R}}^2 \to \mathbb{R}$ and $F : \hat{\mathbb{R}}^2 \to [0, \infty)$ are bounded on compact subsets of $\hat{\mathbb{R}}^2$ and

$$\lim_{\xi\to 0} F(\xi) = \frac{|\partial_2^p\hat{\varphi}(0)|^2}{p!^2}. \tag{3.2}$$

If $j \in \mathbb{N}$, $-L_j^- \leqslant l \leqslant L_j^+$, and $\xi = (\xi_1, \xi_2)$, note that

$$\xi a^{-j}b(l) = \left(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j}l\xi_1\right) \tag{3.3}$$

and, since $\beta < 1$, that

$$\left|a_1^{-\alpha j}\xi_2 + a_1^{-j}l\xi_1\right| \leqslant a_1^{-\alpha j}|\xi_2| + a_1^{-j}\max\left\{L_j^-, L_j^+\right\}|\xi_1| \to 0, \tag{3.4}$$

as $j \to \infty$. Using (3.1) and (3.3), we obtain

$$\omega_j(\xi) = \widetilde{E}_j(\xi) + \widetilde{F}_j(\xi), \tag{3.5}$$

for all $j \in \mathbb{N}$ and $\xi = (\xi_1, \xi_2)$, where

$$\widetilde{E}_j(\xi) = \xi_1 a_1^{-j}\sum_{l=-L_j^-}^{L_j^+} E\left(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j}l\xi_1\right)$$

and

$$\widetilde{F}_j(\xi) = \sum_{l=-L_j^-}^{L_j^+} F\left(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j}l\xi_1\right)\left(a_1^{-\alpha j}\xi_2 + a_1^{-j}l\xi_1\right)^{2p}.$$

Using (3.4), that E is bounded on compact subsets of $\hat{\mathbb{R}}^2$, and that $\beta < 1$, it follows that

$$\langle \widetilde{E}_j\rangle_{j\in\mathbb{N}} \text{ converges uniformly to 0 on compact subsets of } \hat{\mathbb{R}}^2. \tag{3.6}$$

For $j \in \mathbb{N}$ and $\xi = (\xi_1, \xi_2)$, write $\widetilde{G}_j(\xi) = \sum_{l=-L_j^-}^{L_j^+} (a_1^{-\alpha j}\xi_2 + a_1^{-j} l\xi_1)^{2p}$. Using the binomial theorem, we obtain

$$\begin{aligned}
\widetilde{G}_j(\xi) &= \sum_{l=-L_j^-}^{L_j^+} \sum_{q=0}^{2p} \binom{2p}{q} \big(a_1^{-\alpha j}\xi_2\big)^{2p-q} \big(a_1^{-j} l\xi_1\big)^q \\
&= \sum_{q=0}^{2p} \sigma_{qj} \binom{2p}{q} \xi_1^q \xi_2^{2p-q} \\
&= \sigma_{0j} \sum_{q=0}^{2p} \frac{\sigma_{qj}}{\sigma_{0j}} \binom{2p}{q} \xi_1^q \xi_2^{2p-q} \\
&= \sigma_{2pj} \sum_{q=0}^{2p} \frac{\sigma_{qj}}{\sigma_{2pj}} \binom{2p}{q} \xi_1^q \xi_2^{2p-q},
\end{aligned} \tag{3.7}$$

for all $j \in \mathbb{N}$ and all (ξ_1, ξ_2), where $\sigma_{qj} = a_1^{-((1-\alpha)q+2p\alpha)j} \sum_{l=-L_j^-}^{L_j^+} l^q$. If $\lambda \in [0, \infty)$, integral estimation implies that

$$\lim_{L \to \infty} \frac{\sum_{l=0}^{L} l^\lambda}{L^{\lambda+1}} = \frac{1}{\lambda + 1}. \tag{3.8}$$

Using (3.8), it follows that

$$\begin{aligned}
\sigma_{qj} &= a_1^{-((1-\alpha)q+2p\alpha)j} \big(L_j^-\big)^{q+1} \frac{\sum_{l=1}^{L_j^-} (-l)^q}{(L_j^-)^{q+1}} + a_1^{-((1-\alpha)q+2p\alpha)j} \big(L_j^+\big)^{q+1} \frac{\sum_{l=0}^{L_j^+} l^q}{(L_j^+)^{q+1}} \\
&\to \begin{cases} \infty, & \text{if } q = 0,\ 2p\alpha < \beta, \\ 2K, & \text{if } q = 0,\ 2p\alpha = \beta, \\ \infty, & \text{if } q = 2p,\ 2p/(2p+1) < \beta, \\ 2K^{2p+1}/(2p+1), & \text{if } q = 2p,\ 2p/(2p+1) = \beta, \end{cases}
\end{aligned} \tag{3.9}$$

as $j \to \infty$. If $r, q \in \{0, \dots, 2p\}$ and r is even, (3.8) implies that

$$\begin{aligned}
\frac{\sigma_{qj}}{\sigma_{rj}} &= a_1^{-(q-r)(1-\alpha-\beta)j} \frac{a_1^{-(q+1)\beta j} (L_j^-)^{q+1} \frac{\sum_{l=1}^{L_j^-} (-l)^q}{(L_j^-)^{q+1}} + a_1^{-(q+1)\beta j} (L_j^+)^{q+1} \frac{\sum_{l=0}^{L_j^+} l^q}{(L_j^+)^{q+1}}}{a_1^{-(r+1)\beta j} (L_j^-)^{r+1} \frac{\sum_{l=1}^{L_j^-} (-l)^r}{(L_j^-)^{r+1}} + a_1^{-(r+1)\beta j} (L_j^+)^{r+1} \frac{\sum_{l=0}^{L_j^+} l^r}{(L_j^+)^{r+1}}} \\
&\to \begin{cases} 0, & \text{if } \alpha + \beta < 1,\ r < q, \\ 0, & \text{if } \alpha + \beta > 1,\ r > q, \\ 0, & \text{if } \alpha + \beta = 1,\ q \text{ odd}, \\ (r+1)K^{q-r}/(q+1), & \text{if } \alpha + \beta = 1,\ q \text{ even}, \end{cases}
\end{aligned} \tag{3.10}$$

as $j \to \infty$. Finally, note that $\beta/(2 - 2\beta) \sim \beta/(2\alpha)$ if and only if $\alpha + \beta \sim 1$ and that $2p/(2p+1) \sim \beta$ if and only if $p \sim \beta/(2 - 2\beta)$, where $\sim$ is any of the relations $<$, $=$, or $>$.

To prove (i), suppose $p < \gamma$. By (3.2) and (3.4)–(3.6), it suffices to show $\lim_{j\to\infty} \widetilde{G}_j(\xi) = \infty$, for a.e. ξ. This follows by considering each of the cases $\alpha + \beta < 1$, $\alpha + \beta > 1$, and $\alpha + \beta = 1$ separately and using (3.7), (3.9), (3.10), and the last sentence of the previous paragraph.

To prove (ii), suppose $p = \gamma$. By (3.2) and (3.4)–(3.6), it suffices to show $\langle \widetilde{G}_j \rangle_{j\in\mathbb{N}}$ converges uniformly to $(\gamma!^2/|\partial_2^\gamma \hat{\varphi}(0)|^2)P$ on compact subsets of $\hat{\mathbb{R}}^2$. This follows by considering each of the cases $\alpha + \beta < 1$, $\alpha + \beta > 1$, and $\alpha + \beta = 1$ separately and using (3.7), (3.9), (3.10), and the last sentence of the penultimate paragraph. □

Proof of part (iii). Suppose $\partial_2^p \hat{\varphi}(0) = 0$, for all $0 \leqslant p \leqslant \gamma$. Note that, for each $\lambda > 0$, there exists $0 < K' = K'(\lambda) < \infty$ such that

$$|\xi_1 + \xi_2|^\lambda \leqslant K'\big(|\xi_1|^\lambda + |\xi_2|^\lambda\big), \tag{3.11}$$

for all (ξ_1, ξ_2). Using this observation and Taylor's theorem, it follows that

$$\left|\hat{\varphi}(\xi)\right|^2 \leqslant E(\xi)|\xi_1| + F(\xi)|\xi_1|^{2\gamma} + F(\xi)|\xi_2|^{2\gamma}, \tag{3.12}$$

for all $\xi = (\xi_1, \xi_2)$, where $E, F : \hat{\mathbb{R}}^2 \to [0, \infty)$ are bounded on compact subsets of $\hat{\mathbb{R}}^2$ and

$$\lim_{\xi \to 0} F(\xi) = 0. \tag{3.13}$$

Using (3.12) and (3.3), we obtain

$$\omega_j(\xi) \leqslant \widetilde{E}_j(\xi) + \widetilde{F}_j(\xi) + \widetilde{G}_j(\xi), \tag{3.14}$$

for all $j \in \mathbb{N}$ and $\xi = (\xi_1, \xi_2)$, where

$$\widetilde{E}_j(\xi) = |\xi_1| a_1^{-j} \sum_{l=-L_j^-}^{L_j^+} E\big(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\big),$$

$$\widetilde{F}_j(\xi) = |\xi_1|^{2\gamma} a_1^{-2\gamma j} \sum_{l=-L_j^-}^{L_j^+} F\big(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\big),$$

and

$$\widetilde{G}_j(\xi) = \sum_{l=-L_j^-}^{L_j^+} F\big(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\big)\big|a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\big|^{2\gamma}.$$

By definition of γ, we have $\beta \leqslant 2\alpha\gamma < 2\gamma$. Using also that $\beta < 1$, (3.4), and that E and F are bounded on compact subsets of $\hat{\mathbb{R}}^2$, it follows that

$$\langle \widetilde{E}_j \rangle_{j \in \mathbb{N}}, \langle \widetilde{F}_j \rangle_{j \in \mathbb{N}} \text{ converge uniformly to 0 on compact subsets of } \hat{\mathbb{R}}^2. \tag{3.15}$$

By (3.11), there exists $0 < K' < \infty$ such that

$$\begin{aligned} \sum_{l=-L_j^-}^{L_j^+} \big|a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\big|^{2\gamma} &\leqslant \sum_{l=-L_j^-}^{L_j^+} K'\big(\big|a_1^{-\alpha j}\xi_2\big|^{2\gamma} + \big|a_1^{-j} l \xi_1\big|^{2\gamma}\big) \\ &= K'|\xi_2|^{2\gamma} a_1^{-2\alpha\gamma j}\big(L_j^- + L_j^+ + 1\big) \\ &\quad + K'|\xi_1|^{2\gamma} a_1^{-2\gamma j}\left(\sum_{l=0}^{L_j^-} l^{2\gamma} + \sum_{l=0}^{L_j^+} l^{2\gamma}\right), \end{aligned} \tag{3.16}$$

for all $j \in \mathbb{N}$ and all (ξ_1, ξ_2). Since $2\alpha\gamma \geqslant \beta$, we have

$$a_1^{-2\alpha\gamma j}\big(L_j^- + L_j^+ + 1\big) \leqslant a_1^{-\beta j}\big(L_j^- + L_j^+ + 1\big) \to 2K, \tag{3.17}$$

as $j \to \infty$. Since $\gamma \geqslant \beta/(2 - 2\beta)$, we have $2\gamma/(2\gamma + 1) \geqslant \beta$. Using also (3.8), we obtain

$$\begin{aligned} a_1^{-2\gamma j}\left(\sum_{l=0}^{L_j^-} l^{2\gamma} + \sum_{l=0}^{L_j^+} l^{2\gamma}\right) &= \big(a_1^{-2\gamma j/(2\gamma+1)} L_j^-\big)^{2\gamma+1} \frac{\sum_{l=0}^{L_j^-} l^{2\gamma}}{(L_j^-)^{2\gamma+1}} \\ &\quad + \big(a_1^{-2\gamma j/(2\gamma+1)} L_j^+\big)^{2\gamma+1} \frac{\sum_{l=0}^{L_j^+} l^{2\gamma}}{(L_j^+)^{2\gamma+1}} \\ &\leqslant \big(a_1^{-\beta j} L_j^-\big)^{2\gamma+1} \frac{\sum_{l=0}^{L_j^-} l^{2\gamma}}{(L_j^-)^{2\gamma+1}} + \big(a_1^{-\beta j} L_j^+\big)^{2\gamma+1} \frac{\sum_{l=0}^{L_j^+} l^{2\gamma}}{(L_j^+)^{2\gamma+1}} \\ &\to \frac{2K^{2\gamma+1}}{2\gamma + 1}, \end{aligned} \tag{3.18}$$

as $j \to \infty$. Part (iii) follows from (3.4) and (3.13)–(3.18). □

We require the following three lemmas in the proofs of Theorems 3.2 and 3.3:

Lemma 3.1. *Suppose that $\varphi \in L^1(\mathbb{R}^2)$ and that $\hat{\varphi}(0, \eta_2) \neq 0$, for some η_2.*

(i) *If $\beta = 1$, then $\lim_{j\to\infty} \omega_j(\xi) = \infty$, for all $\xi = (\xi_1, \xi_2)$ such that $|\xi_1| > |\eta_2|/K$.*
(ii) *If $\beta > 1$, then $\lim_{j\to\infty} \omega_j(\xi) = \infty$, for a.e. ξ.*

Proof. Since $\varphi \in L^1(\mathbb{R}^2)$, the Riemann–Lebesgue lemma implies that $\hat{\varphi} \in C(\hat{\mathbb{R}}^2)$. Since also $\hat{\varphi}(0, \eta_2) \neq 0$, there exist $\epsilon, \gamma > 0$ such that

$$\xi \in E = [-\epsilon, \epsilon] \times [\eta_2 - \epsilon, \eta_2 + \epsilon]$$

implies $|\hat{\varphi}(\xi)|^2 \geqslant \gamma$. For $\xi \in \hat{\mathbb{R}}^2$ and $j \in \mathbb{N}$, write $C_j(\xi) = \{b \in B_j\colon\ \xi a^{-j} b \in E\}$. Then, $\omega_j(\xi) \geqslant \gamma |C_j(\xi)|$, for all $j \in \mathbb{N}$ and all ξ, where $|C_j(\xi)|$ denotes the cardinality of $C_j(\xi)$. If $\beta = 1$ and $|\xi_1| \geqslant |\eta_2|/K$ or if $\beta > 1$ and $\xi_1 \neq 0$, it follows from inspection of (3.3) that $|C_j(\xi_1, \xi_2)| \to \infty$, as $j \to \infty$. The lemma follows. □

Lemma 3.2. *Let $S > 0$. Suppose that $\beta = 1$, that $(1 + \|x\|^{1/2})\varphi \in L^1(\mathbb{R}^2)$, and that $\hat{\varphi}(0, \xi_2) = 0$, for all ξ_2 with $|\xi_2| \leqslant S$. Then, $\langle \omega_j \rangle_{j\in\mathbb{N}}$ converges uniformly to 0 on compact subsets of $\{(\xi_1, \xi_2)\colon\ |\xi_1| < S/K\}$.*

Proof. The Riemann–Lebesgue lemma and part (i) of Lemma 2.2 imply that $\hat{\varphi} \in C(\hat{\mathbb{R}}^2) \cap \mathcal{H}_1^{1/2}(\hat{\mathbb{R}}^2)$. Using also the assumption regarding the vanishing of $\hat{\varphi}$, it follows that

$$\left|\hat{\varphi}(\xi)\right|^2 \leqslant M(\xi_1)|\xi_1|, \tag{3.19}$$

for all $\xi = (\xi_1, \xi_2)$ with $|\xi_2| \leqslant S$, where M is bounded and satisfies

$$M(\xi_1) \to 0, \quad \text{as } \xi_1 \to 0. \tag{3.20}$$

Let E be a compact subset of $\{(\xi_1, \xi_2)\colon\ |\xi_1| < S/K\}$. Using that $\beta = 1$ and the compactness of E, it follows that $|a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1| \leqslant S$, for all large enough $j \in \mathbb{N}$, all $l = -L_j^-, \dots, L_j^+$, and all $\xi = (\xi_1, \xi_2) \in E$. For these j and ξ, using (3.3) and (3.19), we obtain

$$\begin{aligned}\omega_j(\xi) &= \sum_{l=-L_j^-}^{L_j^+} \left|\hat{\varphi}\left(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\right)\right|^2 \\ &\leqslant \sum_{l=-L_j^-}^{L_j^+} M\left(a_1^{-j}\xi_1\right)\left|a_1^{-j}\xi_1\right| \\ &= |\xi_1| M\left(a_1^{-j}\xi_1\right) a_1^{-j}\left(L_j^- + L_j^+ + 1\right).\end{aligned}$$

The lemma follows from the above inequality, $\beta = 1$, and (3.20). □

Lemma 3.3. *If $\lambda > 1$, then*

$$\sum_{l=-\infty}^{\infty} \frac{1}{(1 + |x + ly|)^\lambda} \leqslant \frac{2}{(\lambda - 1)|y|(1 - |y|)^{\lambda-1}},$$

for all $x, y \in \mathbb{R}$ with $0 < |y| < 1$.

Lemma 3.3 follows easily from integral estimation; we therefore omit its proof. Theorem 3.2 follows immediately from part (i) of Lemmas 3.1 and 3.2. We now prove Theorem 3.3.

Proof. Part (i) follows immediately from part (ii) of Lemma 3.1.

To prove part (ii), suppose that $\hat{\varphi}(0, \xi_2) = 0$, for all ξ_2, that $\beta < 2$, and that $\|x\|^{\beta/2}\varphi \in L^1(\mathbb{R}^2)$. We proceed similarly to as in the proof of Lemma 3.2. The Riemann–Lebesgue lemma and part (i) of Lemma 2.2 imply that $\hat{\varphi} \in C(\hat{\mathbb{R}}^2) \cap \mathcal{H}_1^{\beta/2}(\hat{\mathbb{R}}^2)$. Using also the assumption regarding the vanishing of $\hat{\varphi}$, it follows that

$$\left|\hat{\varphi}(\xi)\right|^2 \leqslant M(\xi_1)|\xi_1|^\beta, \tag{3.21}$$

for all $\xi = (\xi_1, \xi_2)$, where M is bounded and satisfies

$$M(\xi_1) \to 0, \quad \text{as } \xi_1 \to 0. \tag{3.22}$$

Using (3.3) and (3.21), we obtain

$$\begin{aligned}\omega_j(\xi) &= \sum_{l=-L_j^-}^{L_j^+} \left|\hat{\varphi}\left(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\right)\right|^2 \\ &\leqslant \sum_{l=-L_j^-}^{L_j^+} M\left(a_1^{-j}\xi_1\right)\left|a_1^{-j}\xi_1\right|^\beta \\ &= |\xi_1|^\beta M\left(a_1^{-j}\xi_1\right) a_1^{-\beta j}\left(L_j^- + L_j^+ + 1\right),\end{aligned}$$

for all $j \in \mathbb{N}$ and all $\xi = (\xi_1, \xi_2)$. Part (ii) follows from the above inequality and (3.22).

To prove part (iii), suppose that $\hat{\varphi}(0, \xi_2) = 0$, for all ξ_2, and that one of the following two conditions holds:

- φ is compactly supported and Hölder continuous with exponent $\gamma > 1/2$;
- $\varphi \in C_0(\mathbb{R}^2) \cap C^1(\mathbb{R}^2)$ and $(1 + \|x\|^{1/2})\partial^\nu \varphi \in L^1(\mathbb{R}^2)$, for all multi-indices ν with $|\nu| \leqslant 1$.

We proceed similarly to as in the previous paragraph. Standard results regarding the connection, via the Fourier transform, between regularity and decay (e.g. Theorem 8.22 of [26]) and Lemma 2.2 imply that $\hat{\varphi}$, $|\xi_2|^\gamma \hat{\varphi} \in C(\hat{\mathbb{R}}^2) \cap \mathcal{H}_1^{1/2}(\hat{\mathbb{R}}^2)$, for some $\gamma > 1/2$. Using also the assumption regarding the vanishing of $\hat{\varphi}$ and (3.11), it follows that

$$\left|\hat{\varphi}(\xi)\right|^2 \leqslant \frac{M(\xi_1)|\xi_1|}{(1 + |\xi_2|)^{2\gamma}},$$

for all $\xi = (\xi_1, \xi_2)$, where M is bounded and satisfies $M(\xi_1) \to 0$, as $\xi_1 \to 0$. Using this and Lemma 3.3, we have

$$\begin{aligned}\sum_{b \in B}\left|\hat{\varphi}\left(\xi a^{-j} b\right)\right|^2 &= \sum_{l=-\infty}^{\infty} \left|\hat{\varphi}\left(a_1^{-j}\xi_1, a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1\right)\right|^2 \\ &\leqslant \sum_{l=-\infty}^{\infty} \frac{M(a_1^{-j}\xi_1)|a_1^{-j}\xi_1|}{(1 + |a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1|)^{2\gamma}} \\ &= M\left(a_1^{-j}\xi_1\right)\left|a_1^{-j}\xi_1\right| \sum_{l=-\infty}^{\infty} \frac{1}{(1 + |a_1^{-\alpha j}\xi_2 + a_1^{-j} l \xi_1|)^{2\gamma}} \\ &\leqslant M\left(a_1^{-j}\xi_1\right)\left|a_1^{-j}\xi_1\right| \frac{2}{(2\gamma - 1)|a_1^{-j}\xi_1|(1 - |a_1^{-j}\xi_1|)^{2\gamma - 1}} \\ &= \frac{2}{(2\gamma - 1)(1 - a_1^{-j}|\xi_1|)^{2\gamma - 1}} M\left(a_1^{-j}\xi_1\right),\end{aligned}$$

for all $j \in \mathbb{N}$ and all $\xi = (\xi_1, \xi_2)$ with $0 < |\xi_1| a_1^{-j} < 1$. Part (iii) follows. □

4. Traditional shearlet scaling functions

In this section, we prove two results (Theorems 4.1 and 4.2) which provide very strong evidence against the existence of traditional shearlet scaling functions satisfying (D1). Throughout this section, we assume $\varphi \neq 0$ and we let V denote the closed linear span of the collection $\Phi = \{D_b T_k \varphi\colon\ b \in B,\ k \in \mathbb{Z}^2\}$. The following is our first result.

Theorem 4.1. *The following three properties cannot all be satisfied*:

(i) $D_a V \subset V$.
(ii) Φ *forms a frame for* V.
(iii) *One of the following is satisfied*:
  (a) φ *is compactly supported and Hölder continuous with exponent* $\gamma > 1/2$.
  (b) $\varphi \in C_0(\mathbb{R}^2) \cap C^1(\mathbb{R}^2)$ *and* $(1 + \|x\|^{1/2})\partial^\nu \varphi \in L^1(\mathbb{R}^2)$, *for all multi-indices* ν *with* $|\nu| \leqslant 1$.
  (c) $\varphi \in L^1(\mathbb{R}^2)$ *and* $\varphi(x) \geqslant 0$, *for a.e.* x.

It turns out that we can omit the regularity assumptions from property (iii) of the above theorem, provided we, first, strengthen the refinement property (property (i)) to a "B-finite" refinement property and, second, strengthen the representation property (property (ii)) to a "semi-biorthogonal" frame property. Both of these (which are defined/indicated below) are very desirable properties for a traditional shearlet scaling function to possess. We have the following definition:

Definition 4.1. We say Φ forms a semi-biorthogonal frame for V if it forms a frame for V and if there exists a finite subset G of B and a dual frame to Φ of the form $\Theta = \{D_b T_k \theta\colon\ b \in B,\ k \in \mathbb{Z}^2\}$ such that $\varphi \perp D_b T_k \theta$, whenever $b \in B \setminus G$ and $k \in \mathbb{Z}^2$.

We make the following remarks regarding the above definition: First, if Φ forms a frame for V, then its canonical dual frame $\{(D_b T_k \varphi)^\sim\colon\ b \in B,\ k \in \mathbb{Z}^2\}$ automatically satisfies $(D_b T_k \varphi)^\sim = D_b T_k (\varphi)^\sim$, for all $b \in B$ and $k \in \mathbb{Z}^2$. Second, if Φ forms a Riesz basis for V, then it forms a semi-biorthogonal frame for V. We can now state our second result.

Theorem 4.2. *Assume $a_1^{\alpha}, a_1^{1-\alpha} \in \mathbb{Z}$. The following three properties cannot all be satisfied*:

(i) *There exists a finite subset F of B such that $D_a\varphi$ belongs to the closed linear span of $\{D_b T_k \varphi\colon\ b \in F,\ k \in \mathbb{Z}^2\}$.*
(ii) *Φ forms a semi-biorthogonal frame for V.*
(iii) $(1 + \|x\|^{(1-\alpha)/2})\varphi \in L^1(\mathbb{R}^2)$.

We now prove Theorem 4.1.

Proof. To obtain a contradiction, suppose that properties (i)–(iii) are satisfied. Let $C \leqslant D$ denote the frame constants of Φ, and, for $j \in \mathbb{N}$, write $V_j = D_a^{-j} V$. Since the operator D_a is unitary, the collection

$$\left\{D_a^{-j} D_b T_k \varphi\colon\ b \in B,\ k \in \mathbb{Z}^2\right\}$$

forms a frame for V_j with constants $C \leqslant D$, for each $j \in \mathbb{N}$.

We claim that $\hat{\varphi}(0, \xi_2) = 0$, for all ξ_2. Otherwise, choose $0 \neq f \in L_c^\infty(\hat{\mathbb{R}}^2)^\vee$. Using the Bessel property, (1.3), part (vi) of Lemma 2.1, Fatou's lemma, and part (i) of Theorem 3.3, it follows that

$$\begin{aligned} D\|f\|^2 &\geqslant \liminf_{j\to\infty} \sum_{b\in B, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-j} D_b T_k \varphi\right\rangle\right|^2 \\ &= \liminf_{j\to\infty} \int_{\hat{\mathbb{R}}^2} \left|\hat{f}(\xi)\right|^2 \sum_{b\in B} \left|\hat{\varphi}\left(\xi a^{-j} b\right)\right|^2 d\xi \\ &\geqslant \int_{\hat{\mathbb{R}}^2} \left|\hat{f}(\xi)\right|^2 \liminf_{j\to\infty} \sum_{b\in B} \left|\hat{\varphi}\left(\xi a^{-j} b\right)\right|^2 d\xi = \infty, \end{aligned}$$

a contradiction. By the second comment following Theorem 3.3, it follows that property (iii)(c) cannot be satisfied.

For $j \in \mathbb{N}$, let $P_j : L^2(\mathbb{R}^2) \to V_j$ denote the orthogonal projection of $L^2(\mathbb{R}^2)$ onto V_j. Choose $f \in L_c^\infty(\hat{\mathbb{R}}^2)^\vee$ such that $P_0 f \neq 0$. Using properties (i) and (ii), that either (iii)(a) or (iii)(b) is satisfied, (1.3), part (vi) of Lemma 2.1, the vanishing property of $\hat{\varphi}$, and part (iii) of Theorem 3.3, for large enough $j \in \mathbb{N}$, we have

$$\begin{aligned} 0 &< \|P_0 f\|^2 \leqslant \|P_j f\|^2 \\ &\leqslant C^{-1} \sum_{b\in B,\, k\in\mathbb{Z}^2} \left|\left\langle P_j f, D_a^{-j} D_b T_k \varphi\right\rangle\right|^2 \\ &= C^{-1} \sum_{b\in B,\, k\in\mathbb{Z}^2} \left|\left\langle f, D_a^{-j} D_b T_k \varphi\right\rangle\right|^2 \\ &= C^{-1} \int_{\hat{\mathbb{R}}^2} \left|\hat{f}(\xi)\right|^2 \sum_{b\in B} \left|\hat{\varphi}\left(\xi a^{-j} b\right)\right|^2 d\xi \to 0, \end{aligned}$$

as $j \to \infty$. This contradiction completes the proof. □

For the remainder of the section, we assume $a_1^{\alpha}, a_1^{1-\alpha} \in \mathbb{Z}$. We note that

$$b(l)b\left(l'\right) = b\left(l + l'\right) \quad \text{and} \quad ab(l)a^{-1} = b\left(a^{1-\alpha} l\right), \tag{4.1}$$

for all $l, l' \in \mathbb{Z}$. We use the following notation: For $j, L \in \mathbb{Z}^+$ and $E \subset B$, write $B(L) = \{b(l)\colon\ |l| \leqslant L\}$ and

$$E^j = \left\{ \prod_{p=0}^{j-1} a^p b_p a^{-p}\colon\ b_p \in E,\ \text{for each } p \right\}.$$

It follows from (4.1) that $E^j \subset B$, for all j. We require the following lemma in the proof of Theorem 4.2:

Lemma 4.1. *We have the following*:

(i) *If $j, L \in \mathbb{Z}^+$, then $B(L)^j \subset B(La_1^{(1-\alpha)j})$.*
(ii) *If E is a finite subset of B and if there exists $\{m_b\colon\ b \in E\} \subset L^\infty(\hat{\mathbb{T}}^2)$ such that*

$$\hat{\varphi}(\xi a) = \sum_{b \in E} m_b(\xi)\hat{\varphi}(\xi b), \tag{4.2}$$

for a.e. ξ, then, for each $j \in \mathbb{Z}^+$, $D_a^j\varphi$ belongs to the closed linear span of $\{D_b T_k \varphi\colon\ b \in E^j, k \in \mathbb{Z}^2\}$.

Proof. If $j, L \in \mathbb{Z}^+$ and $l_0, \dots, l_{j-1} \in \mathbb{Z}$ with $|l_p| \leqslant L$ (for all p), using (4.1), we have

$$\prod_{p=0}^{j-1} a^p b(l_p) a^{-p} = b\left(\sum_{p=0}^{j-1} l_p a_1^{(1-\alpha)p}\right)$$

and, since $a_1^{1-\alpha} \in \{2, 3, 4, \dots\}$,

$$\left|\sum_{p=0}^{j-1} l_p a_1^{(1-\alpha)p}\right| \leqslant L\sum_{p=0}^{j-1} a_1^{(1-\alpha)p} = L\frac{a_1^{(1-\alpha)j} - 1}{a_1^{1-\alpha} - 1} \leqslant La_1^{(1-\alpha)j}.$$

Part (i) follows.

To prove (ii), suppose that E is a finite subset of B and that there exists $\{m_b\colon\ b \in E\} \subset L^\infty(\hat{\mathbb{T}}^2)$ such that (4.2) holds, for a.e. ξ. We claim that, for each $j \in \mathbb{Z}^+$, there exists $\{m_b^j\colon\ b \in E^j\} \subset L^\infty(\hat{\mathbb{T}}^2)$ such that

$$\hat{\varphi}(\xi a^j) = \sum_{b \in E^j} m_b^j(\xi)\hat{\varphi}(\xi b), \tag{4.3}$$

for a.e. ξ. We proceed by induction on j. The case $j = 1$ follows by assumption. Fix $j \in \mathbb{Z}^+$ and suppose there exists $\{m_b^j\colon\ b \in E^j\} \subset L^\infty(\hat{\mathbb{T}}^2)$ such that (4.3) is satisfied, for a.e. ξ. Using (4.3) and (4.2), we obtain

$$\begin{aligned} \hat{\varphi}(\xi a^{j+1}) &= \hat{\varphi}(\xi a a^j) \\ &= \sum_{b \in E^j} m_b^j(\xi a)\hat{\varphi}(\xi a b) \\ &= \sum_{b \in E^j} m_b^j(\xi a)\hat{\varphi}(\xi a b a^{-1} a) \\ &= \sum_{b \in E^j} m_b^j(\xi a)\sum_{b' \in E} m_{b'}(\xi a b a^{-1})\hat{\varphi}(\xi a b a^{-1} b'), \end{aligned} \tag{4.4}$$

for a.e. ξ. If $b \in E^j$ and $b' \in E$, it follows that $aba^{-1}b' \in E^{j+1}$ and, since $a, aba^{-1} \in GL_2(\mathbb{Z})$, that $m_b^j(\cdot a), m_{b'}(\cdot aba^{-1}) \in L^\infty(\hat{\mathbb{T}}^2)$. These observations imply that the final quantity in (4.4) may be written as $\sum_{b \in E^{j+1}} m_b^{j+1}(\xi)\hat{\varphi}(\xi b)$, for a suitable collection $\{m_b^{j+1}\colon\ b \in E^{j+1}\} \subset L^\infty(\hat{\mathbb{T}}^2)$. This completes the induction and proves the claim.

For each $j \in \mathbb{Z}^+$, using (4.3) we obtain

$$\begin{aligned} (D_a^j\varphi)\widehat{\ }(\xi) &= |\det a|^{j/2}\hat{\varphi}(\xi a^j) \\ &= |\det a|^{j/2}\sum_{b \in E^j} m_b^j(\xi)\hat{\varphi}(\xi b) \\ &= \sum_{b \in E^j} |\det a|^{j/2} m_b^j(\xi)(D_b\varphi)\widehat{\ }(\xi), \end{aligned}$$

for a.e. ξ. Part (ii) follows from (1.3), the above equality, and part (iii) of Lemma 2.1. □

We now prove Theorem 4.2.

Proof. Suppose, in order to obtain a contradiction, that properties (i)–(iii) are satisfied. Let G, θ, and Θ be as in Definition 4.1. Choose $L \in \mathbb{Z}^+$ such that $F, G \subset B(L)$ and let $0 < D < \infty$ be the Bessel constant of Θ. For $p \in \mathbb{Z}^+$, let $V(p)$ and $W(p)$ denote the closed linear spans of $\{D_{b(l)} T_k \varphi\colon\ |l| \leqslant p,\ k \in \mathbb{Z}^2\}$ and $\{D_{b(l)} T_k \theta\colon\ |l| \geqslant p,\ k \in \mathbb{Z}^2\}$, respectively. Using property (ii), that $G \subset B(L)$, (4.1), and that $B \subset SL_2(\mathbb{Z})$, it follows that

$$V(p) \perp W(p+L+1), \tag{4.5}$$

for all $p \in \mathbb{Z}^+$. Using property (ii) and that $a \in GL_2(\mathbb{Z})$, it follows that $\{T_k D_a \varphi\colon\ k \in \mathbb{Z}^2\}$ forms a Bessel system. Using these two observations, properties (i) and (ii), (1.3), and parts (i), (ii), and (v) of Lemma 2.1, we obtain $\{m_b\colon\ b \in B(2L)\} \subset L^\infty(\hat{\mathbb{T}}^2)$ such that

$$\hat{\varphi}(\xi a) = \sum_{b \in B(2L)} m_b(\xi) \hat{\varphi}(\xi b),$$

for a.e. ξ. By Lemma 4.1, $D_a^j \varphi \in V(2La_1^{(1-\alpha)j})$, for each $j \in \mathbb{Z}^+$. Using (4.5), property (ii), (1.3), and parts (i), (iv), and (v) of Lemma 2.1, for each $j \in \mathbb{Z}^+$, we obtain $\{m_b^j\colon\ b \in B(L_j)\} \subset L^2(\hat{\mathbb{T}}^2)$ such that

$$|\det a|^{j/2} \hat{\varphi}\big(\xi a^j\big) = \big(D_a^j \varphi\big)^\frown(\xi) = \sum_{b \in B(L_j)} m_b^j(\xi) \hat{\varphi}(\xi b), \tag{4.6}$$

for a.e. ξ, and

$$\sum_{b \in B(L_j)} \big\|m_b^j\big\|^2 \leqslant D \big\|D_a^j \varphi\big\|^2 = D\|\varphi\|^2, \tag{4.7}$$

where $L_j = 2La_1^{(1-\alpha)j} + L$.

For $j \in \mathbb{Z}^+$ define $\omega_j \in L^1(\hat{\mathbb{R}}^2)$ by $\omega_j(\xi) = \sum_{b \in B(L_j)} |\hat{\varphi}(\xi a^{-j} b)|^2$. Using the argument of the second paragraph of the proof of Theorem 4.1, it follows that $\hat{\varphi}(0, \xi_2) = 0$, for all ξ_2. Using also property (iii) and part (ii) of Theorem 3.3, it follows that

$$\langle \omega_j \rangle_{j \in \mathbb{Z}^+} \text{ converges uniformly to 0 on compact subsets of } \hat{\mathbb{R}}^2. \tag{4.8}$$

Let E be a bounded and measurable subset of $\hat{\mathbb{R}}^2$ satisfying $E \cap (E+k) = \emptyset$, for all $k \in \hat{\mathbb{Z}}^2 \setminus \{0\}$. Using (4.6) and two applications of the Schwarz inequality, we obtain

$$\begin{aligned}
\int_E \big|\hat{\varphi}(\xi)\big|\, d\xi &= |\det a|^{-j/2} \int_E \Bigg|\sum_{b \in B(L_j)} m_b^j\big(\xi a^{-j}\big) \hat{\varphi}\big(\xi a^{-j} b\big)\Bigg|\, d\xi \\
&\leqslant |\det a|^{-j/2} \int_E \Bigg(\sum_{b \in B(L_j)} \big|m_b^j\big(\xi a^{-j}\big)\big|^2\Bigg)^{1/2} \Bigg(\sum_{b \in B(L_j)} \big|\hat{\varphi}\big(\xi a^{-j} b\big)\big|^2\Bigg)^{1/2} d\xi \\
&\leqslant |\det a|^{-j/2} \Bigg(\int_E \sum_{b \in B(L_j)} \big|m_b^j\big(\xi a^{-j}\big)\big|^2\, d\xi\Bigg)^{1/2} \Bigg(\int_E \sum_{b \in B(L_j)} \big|\hat{\varphi}\big(\xi a^{-j} b\big)\big|^2\, d\xi\Bigg)^{1/2} \\
&= \Bigg(|\det a|^{-j} \int_E \sum_{b \in B(L_j)} \big|m_b^j\big(\xi a^{-j}\big)\big|^2\, d\xi\Bigg)^{1/2} \Bigg(\int_E \omega_j(\xi)\, d\xi\Bigg)^{1/2},
\end{aligned} \tag{4.9}$$

for all $j \in \mathbb{Z}^+$. Using a change of variable, choice of E, that $a \in GL_2(\mathbb{Z})$, and (4.7), it follows that

$$\begin{aligned}
|\det a|^{-j} \int_E \sum_{b \in B(L_j)} \big|m_b^j\big(\xi a^{-j}\big)\big|^2\, d\xi &= \int_{Ea^{-j}} \sum_{b \in B(L_j)} \big|m_b^j(\xi)\big|^2\, d\xi \\
&\leqslant \int_{[0,1]^2} \sum_{b \in B(L_j)} \big|m_b^j(\xi)\big|^2\, d\xi \\
&= \sum_{b \in B(L_j)} \big\|m_b^j\big\|^2 \\
&\leqslant D\|\varphi\|^2,
\end{aligned}$$

for each $j \in \mathbb{Z}^+$. Letting $j \to \infty$ in (4.9) and using (4.8), choice of E, and the above inequality, we obtain $\int_E |\hat{\varphi}(\xi)|\, d\xi = 0$. Varying E, it follows that $\varphi = 0$. This contradiction completes the proof. □

5. Remark

Let $z, d_1, d_2 > 0$ and write $d = \mathrm{diag}(d_1, d_2)$ and

$$B(z) = \left\{ \begin{pmatrix} 1 & lz \\ 0 & 1 \end{pmatrix} \colon l \in \mathbb{Z} \right\}.$$

Let $I \in \mathbb{Z}^+$ and $\{\varphi_i \colon i = 1, \ldots, I\} \subset L^2(\mathbb{R}^2)$. The notions of traditional shearlet, traditional shearlet MRA and scaling function, shearlet on $\mathcal{C}$, shearlet scaling function on $\mathcal{C}$, etc. can be generalized, for instance, by replacing φ with $\{\varphi_i \colon i = 1, \ldots, I\}$, B with $B(z)$, and/or the translation lattice $\mathbb{Z}^2$ with $d\mathbb{Z}^2$. Versions of Theorems 3.1–4.2 hold for these generalized notions as well (in some cases, additional assumptions must be made on z and d).

Acknowledgments

The author wishes to thank Guido Weiss and Ed Wilson for their various helpful suggestions.

References

[1] C.M. Brislawn, Fingerprints go digital, Notices Amer. Math. Soc. 42 (1995) 1278–1283.
[2] C.M. Brislawn, M.D. Quirk, Image compression with the JPEG-2000 standard, in: R.G. Driggers (Ed.), Encyclopedia of Optical Engineering, Marcel Dekker, New York, 2003, pp. 780–785.
[3] S. Mallat, A Wavelet Tour of Signal Processing, third edition, Academic Press, 2009.
[4] Y. Meyer, Wavelets and Operators, Cambridge Studies in Advanced Mathematics, vol. 37, Cambridge University Press, Cambridge, 1995.
[5] S. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$, Trans. Amer. Math. Soc. 315 (1989) 69–87.
[6] I. Daubechies, Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math. 41 (1988) 909–996.
[7] S. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (1989) 674–693.
[8] E.J. Candès, D.L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities, Comm. Pure Appl. Math. 57 (2004) 219–266.
[9] D.L. Donoho, Sparse components of images and optimal atomic decomposition, Constr. Approx. 17 (2001) 353–382.
[10] E.J. Candès, D.L. Donoho, Continuous curvelet transform: I. Resolution of the wavefront set, Appl. Comput. Harmon. Anal. 19 (2005) 162–197.
[11] K. Guo, G. Kutyniok, D. Labate, Sparse multidimensional representations using anisotropic dilation and shear operators, in: Wavelets and Splines, Nashville, Athens, GA, 2005, Nashboro Press, 2006, pp. 189–201.
[12] K. Guo, D. Labate, Optimally sparse multidimensional representation using shearlets, SIAM J. Math. Anal. 39 (2007) 298–318.
[13] G. Kutyniok, D. Labate, Resolution of the wavefront set using continuous shearlets, Trans. Amer. Math. Soc. 361 (2009) 2719–2754.
[14] P. Kittipoom, G. Kutyniok, W.-Q. Lim, Construction of compactly supported shearlet frames, preprint.
[15] O. Christensen, An Introduction to Frames and Riesz Bases, Birkhäuser, Boston, 2003.
[16] D. L. Donoho, G. Kutyniok, M. Shahram, X. Zhuang, A rational design of a digital shearlet transform, preprint.
[17] D. Labate, W.-Q. Lim, G. Kutyniok, G. Weiss, Sparse multidimensional representation using shearlets, in: A.F. Laine, M. Papadakis, M.A. Unser (Eds.), Wavelets XI, San Diego, CA, 2005, SPIE, Bellingham, WA, 2009, p. 59140U.
[18] K. Guo, D. Labate, W.-Q. Lim, G. Weiss, E. Wilson, Wavelets with composite dilations and their MRA properties, Appl. Comput. Harmon. Anal. 20 (2006) 202–236.
[19] W.-Q. Lim, The discrete shearlet transform: A new directional transform and compactly supported shearlet frames, IEEE Trans. Image Process. 19 (2010) 1166–1180.
[20] B. Han, G. Kutyniok, Z. Shen, A unitary extension principle for shearlet systems, preprint.
[21] G. Kutyniok, T. Sauer, Adaptive directional subdivision schemes and shearlet multiresolution analysis, SIAM J. Math. Anal. 41 (2009) 1436–1471.
[22] A. Ron, Z. Shen, Affine systems in $L^2(\mathbb{R}^d)$: The analysis of the analysis operator, J. Funct. Anal. 148 (1997) 408–447.
[23] C. Deboor, R.A. Devore, A. Ron, The structure of finitely generated shift-invariant spaces in $L^2(\mathbb{R}^d)$, J. Funct. Anal. 119 (1994) 37–78.
[24] M. Bownik, The structure of shift-invariant subspaces of $L^2(\mathbb{R}^n)$, J. Funct. Anal. 177 (2000) 282–309.
[25] E. Hernández, H. Šikić, G. Weiss, E. Wilson, On the properties of the integer translates of a square integrable function in $L^2(\mathbb{R})$, preprint.
[26] G.B. Folland, Real Analysis: Modern Techniques and Their Applications, second edition, John Wiley and Sons, Inc., New York, 1999.