SHEARLET COORBIT SPACES AND ASSOCIATED BANACH FRAMES

STEPHAN DAHLKE, GITTA KUTYNIOK, GABRIELE STEIDL, AND GERD TESCHKE

Abstract. In this paper, we study the relationships of the newly developed continuous shearlet transform with the coorbit space theory. It turns out that all the conditions that are needed to apply the coorbit space theory can indeed be satisfied for the shearlet group. Consequently, we establish new families of smoothness spaces, the shearlet coorbit spaces. Moreover, our approach yields Banach frames for these spaces in a quite natural way. We also study the approximation power of best n-term approximation schemes and present some first numerical experiments.

1. Introduction

One of the central issues in applied analysis is the problem of analyzing and approximating a given signal. The first step is always to decompose the signal with respect to a suitable set of building blocks. These building blocks may, e.g., consist of the elements of a basis, a frame, or even of the elements of huge dictionaries. Classical examples with many important practical applications are wavelet bases/frames and Gabor frames, respectively. The goal is always to find those building blocks that are most appropriate for the given signal. This means, e.g., that the building blocks give rise to sparse representations and/or that interesting features of the signal can be easily extracted.

In signal/image analysis/processing, the wavelet transform has already been very successfully applied and is therefore very often the method of choice. Nevertheless, there is still a serious problem: in image analysis, it is desirable to obtain directional information which is very complicated in the wavelet setting. To overcome this deficiency, several approaches have been suggested in the last few years such as ridgelets [2], curvelets [3], contourlets [9], shearlets [18] and many others. Among all these approaches, the shearlet transform stands out for the following reason. In recent studies [5], it has been shown that the shearlet transform is related with group theory, i.e., this transform can be derived from a strongly continuous, irreducible square-integrable representation of a certain group, the shearlet group. This property provides us with a link to another central problem in applied analysis, namely how to measure the smoothness of a given function. Classical approaches are, e.g., based on (strong or weak) derivatives (Hölder and Sobolev spaces), or moduli of smoothness (Besov spaces). However, by means of the concept of square-integrable group representations it is possible to derive a unified approach to many different smoothness measures: they can all be restated in terms of the decay of the voice transform associated with the representation. Moreover, by discretizing the representation in a judicious way, one obtains frames for these smoothness spaces which can therefore be interpreted as the natural building blocks for the

---

Date: May 9, 2007.

2000 *Mathematics Subject Classification.* 22D10, 42C15, 46E35, 47B25.

Key words and phrases. Shearlets, coorbit space theory, square-integrable group representations, Banach frames.

S.D acknowledges support from Deutsche Forschungsgemeinschaft (DFG), Grant Da 360/4–3.

G.K. acknowledges support from Preis der Justus-Liebig-Universität Gießen 2006 and from Deutsche Forschungsgemeinschaft Heisenberg-Fellowship Ku 1446/8-1.

G.S. acknowledges support from DFG Grant We 2062/2-3.

G.T. gratefully acknowledges partial support by Deutsche Forschungsgemeinschaft Grants Te 354/1-2, Te 354/3-1, Te 354/4-1, and Te 354/5-1.

underlying transformation. All these relationships have been clarified in the so-called *coorbit space theory* which has been derived by Feichtinger and Gröchenig in a series of papers [10, 11, 12, 13].

Now, once we know that the shearlet transform also stems from a square-integrable group representation, it is quite natural to study the relationships with the coorbit theory and to ask the following questions:

- Is it possible to apply the general coorbit theory to the shearlet group, i.e., can all the necessary assumptions indeed be satisfied?
- If so, what can be said about the structure of the associated new families of smoothness spaces, the shearlet coorbit spaces?
- What is the convergence order that can be achieved by approximation schemes based on the shearlet frames?
- Are the resulting shearlet frames really useful in practice, i.e., is it possible to decompose and to reconstruct images in an efficient way?

In this paper, we give at least partial answers to these questions. It turns out that indeed all the assumptions needed for the coorbit theory can be fulfilled. The resulting conditions look quite canonical and the associated shearlet frames can be derived in a very natural way. Note that curvelet–type decomposition spaces were considered in [1].

The second question is clearly substantial enough to fill a whole series of papers, however, in this article we prove a first result which says that at least the very important Schwartz space is contained in the shearlet coorbit spaces.

In the context of the third question, especially nonlinear approximation methods such as best n-term approximations are of interest. In this paper, we show that similar to Gabor frames [17] the approximation order of best n-term approximation schemes based on shearlet frames depends on the smoothness of the signal as measured in a second shearlet coorbit space.

The answer to the fourth question is again a long-term project. Nevertheless, we present some first numerical experiments that indicate the usability of the shearlet approach.

This paper is organized as follows. In Section 2, we introduce the shearlet group and establish the square-integrability of its representation in $L_2(\mathbb{R}^2)$. For a slightly different version of the shearlet group, the reduced shearlet group, similar questions have already been studied in [5]. However, the representation of the reduced shearlet group fails to be irreducible so that this group is not suitable in our setting. In Section 3, we briefly recall the basic concepts of the coorbit space theory as far as it is needed for our purposes. In Section 4, we study the relationships of the shearlet transform with the coorbit space theory. First of all, in Subsection 4.1, we show that all the conditions to construct the coorbit spaces can be satisfied, i.e., the new family of shearlet coorbit spaces is established. Then, in Subsection 4.2, we derive the associated (Banach) frames for these spaces. To this end, suitable discrete subsets, the so-called *U-dense sets*, have to be constructed and an additional integrability condition has to be satisfied. In Subsection 4.3, we prove that the Schwartz space is contained in our shearlet coorbit spaces. Then, in Subsection 4.4, we study the power of best n-term approximation schemes based on the shearlet frames. Finally, in Section 5 we investigate the performance of our new frame algorithms by applying them to some test images.

2. Continuous Shearlet Transform

In this section, we introduce the definition and basic properties of shearlet systems and of the Continuous Shearlet Transform from a group–theoretical point of view.

Recall that a *unitary representation* of a locally compact group G with the left–invariant Haar measure μ_G on a Hilbert space $\mathcal{H}$ is a homomorphism π from G into the group of unitary operators

$\mathcal{U}(\mathcal{H})$ on $\mathcal{H}$ which is continuous with respect to the strong operator topology. Given a unitary representation π of G on $\mathcal{H}$, a function $\psi \in \mathcal{H}$ is called *admissible*, if

$$\int_G |\langle \psi, \pi(g)\psi\rangle|^2 d\mu_G(g) < \infty.$$

The admissibility condition is important, since it yields to a resolution of identity that allows the reconstruction of a function $f \in \mathcal{H}$ from its *voice transform* $V_\psi : \mathcal{H} \to L_2(G)$ given by

$$V_\psi(f)(g) := \langle f, \pi(g)\psi\rangle\,. \tag{1}$$

Suppose that W is a closed subspace of $\mathcal{H}$. Then W is called an *invariant subspace* for π, if $\pi(g)W \subseteq W$ for all $g \in G$. If there exists a nontrivial invariant subspace for π, then π is called *reducible*, otherwise π is *irreducible*. If π is irreducible and there exists at least one nontrivial admissible function $\psi \in \mathcal{H}$, then π is called *square-integrable*. In this case there exists a positive, densely defined self-adjoint operator A on $\mathcal{H}$ such that the following orthogonality relation holds true:

$$\int_G \overline{\langle \pi(g)\tilde{f}_1, f_1\rangle}\langle \pi(g)\tilde{f}_2, f_2\rangle d\mu_G(g) = \langle A\tilde{f}_2, A\tilde{f}_1\rangle\langle f_1, f_2\rangle. \tag{2}$$

For more details on the theory of locally compact groups and group representations we refer to [14].

In this paper, we are interested in the shearlet group. For $a \in \mathbb{R}^* := \mathbb{R}\setminus\{0\}$ and $s \in \mathbb{R}$, let

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \operatorname{sgn}(a)\sqrt{|a|}\end{pmatrix} \quad \text{and} \quad S_s = \begin{pmatrix} 1 & s \\ 0 & 1\end{pmatrix}$$

denote the *parabolic scaling matrix* and the *shear matrix*, respectively, where $\operatorname{sgn}(a)$ denotes the sign of a. Further, for $t \in \mathbb{R}^2$ and $M \in GL(2,\mathbb{R})$, let

$$T_t\, f(x) := f(x-t) \quad \text{and} \quad D_M\, f(x) := |\det M|^{-\frac{1}{2}} f(M^{-1}x)$$

denote the *translation* and *dilation operator* on $L_2(\mathbb{R}^2)$, respectively. It is easy to check that $T_t\, D_{S_s A_a}$ is a unitary operator on $L_2(\mathbb{R}^2)$. The *(full) shearlet group* $\mathbb{S}$ is then defined to be the set $\mathbb{R}^* \times \mathbb{R} \times \mathbb{R}^2$ endowed with the group operation

$$(a,s,t)\,(a',s',t') = (aa', s + s'\sqrt{|a|}, t + S_s A_a t').$$

A left–invariant Haar measure of $\mathbb{S}$ is given by

$$\mu_{\mathbb{S}} = \frac{da}{|a|^3}\, ds\, dt.$$

Let $\pi : \mathbb{S} \to \mathcal{U}(L_2(\mathbb{R}^2))$ be defined by

$$\pi(a,s,t)\,\psi(x) := T_t\, D_{S_s A_a}\,\psi = |a|^{-\frac{3}{4}}\,\psi(A_a^{-1}S_s^{-1}(x-t)). \tag{3}$$

In the following, we use the abbreviation $\psi_{a,s,t} := \pi(a,s,t)\,\psi$. In [5] it was shown that the mapping π is a unitary representation of the *reduced shearlet group* $\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$, i.e., the group with only positive scalings $a > 0$. For the convenience of the reader we provide a short proof of this fact for the full shearlet group.

Lemma 2.1. *The mapping π defined by* (3) *is a unitary representation of* $\mathbb{S}$.

Proof. Let $\psi \in L^2(\mathbb{R}^2)$, $x \in \mathbb{R}^2$, and $(a,s,t),(a',s',t') \in \mathbb{S}$. Using that

$$A_{a'}^{-1}S_{s'}^{-1}A_a^{-1}S_s^{-1} = \begin{pmatrix} \frac{1}{a'} & -\frac{s'}{a'} \\ 0 & \frac{\operatorname{sgn}(a')}{\sqrt{|a'|}}\end{pmatrix}\begin{pmatrix} \frac{1}{a} & -\frac{s}{a} \\ 0 & \frac{\operatorname{sgn}(a)}{\sqrt{|a|}}\end{pmatrix} = \begin{pmatrix} \frac{1}{aa'} & -\frac{s}{aa'} - \frac{s'\operatorname{sgn}(a)}{a'\sqrt{|a|}} \\ 0 & \frac{\operatorname{sgn}(aa')}{\sqrt{|aa'|}}\end{pmatrix} = A_{aa'}^{-1}S_{s+s'\sqrt{|a|}}^{-1},$$

we obtain

$$\begin{aligned}
\pi(a,s,t)(\pi(a',s',t')\psi)(x) &= |a|^{-\frac{3}{4}}\pi(a',s',t')\psi(A_a^{-1}S_s^{-1}(x-t)) \\
&= |aa'|^{-\frac{3}{4}}\psi(A_{a'}^{-1}S_{s'}^{-1}(A_a^{-1}S_s^{-1}(x-t)-t')) \\
&= |aa'|^{-\frac{3}{4}}\psi(A_{a'}^{-1}S_{s'}^{-1}A_a^{-1}S_s^{-1}(x-(t+S_sA_at'))) \\
&= |aa'|^{-\frac{3}{4}}\psi(A_{aa'}^{-1}S_{s+s'\sqrt{|a|}}^{-1}(x-(t+S_sA_at'))) \\
&= \pi((a,s,t)(a',s',t'))\psi(x).
\end{aligned}$$

□

Let the Fourier transform be defined by

$$\mathcal{F}f(\omega) = \hat{f}(\omega) = \int_{\mathbb{R}^2} f(x)e^{-2\pi i\langle\omega,x\rangle}\,dx.$$

Then straightforward computation yields

$$\hat{\psi}_{a,s,t}(\omega) = |a|^{\frac{3}{4}}e^{-2\pi it\omega}\hat{\psi}\left(A_a^{\mathrm{T}}S_s^{\mathrm{T}}\omega\right) = |a|^{\frac{3}{4}}e^{-2\pi it\omega}\hat{\psi}\left(a\omega_1, \operatorname{sgn}(a)\sqrt{|a|}(s\omega_1+\omega_2)\right). \tag{4}$$

The following result shows that the unitary representation π defined in (3) is a square-integrable representation of $\mathbb{S}$. This is not the case for the reduced shearlet group, in particular the representation is not irreducible.

Theorem 2.2. *A function $\psi \in L_2(\mathbb{R}^2)$ is admissible if and only if it fulfills the* admissibility condition

$$C_\psi := \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\hat{\psi}(\omega_1,\omega_2)|^2}{\omega_1^2}\,d\omega_2\,d\omega_1 < \infty. \tag{5}$$

Then, for any $f \in L^2(\mathbb{R}^2)$, the following equality holds true:

$$\int_{\mathbb{S}} |\langle f,\psi_{a,s,t}\rangle|^2\,d\mu_{\mathbb{S}}(a,s,t) = C_\psi\ \|f\|_2^2\,. \tag{6}$$

In particular, the unitary representation π is irreducible and hence square-integrable.

Proof. Employing the Plancherel theorem and (4), we obtain

$$\begin{aligned}
\int_{\mathbb{S}} |\langle f,\psi_{a,s,t}\rangle|^2\,d\mu_{\mathbb{S}}(a,s,t) &= \int_{\mathbb{S}} |f*\psi^*_{a,s,0}(t)|^2\,dt\,ds\,\frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}^2} |\hat{f}(\omega)|^2|\hat{\psi^*}_{a,s,0}(\omega)|^2\,d\omega\,ds\,\frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}^2} |\hat{f}(\omega)|^2|a|^{\frac{3}{2}}|\hat{\psi}(A_a^TS_s^T\omega)|^2\,d\omega\,ds\,\frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^2}\int_{\mathbb{R}} |\hat{f}(\omega)|^2|a|^{-\frac{3}{2}}|\hat{\psi}(a\omega_1,\operatorname{sgn}(a)\sqrt{|a|}(\omega_2+s\omega_1))|^2\,ds\,d\omega\,da,
\end{aligned}$$

where $\psi^*_{a,s,0}(x) = \overline{\psi_{a,s,0}(-x)}$. Now we use that

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(ax)\,dx\,da = \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{|a|}\,f(y)\,dy\,da.$$

Substituting $\xi_2 := \operatorname{sgn}(a)\sqrt{|a|}(\omega_2 + s\omega_1)$, i.e., $\operatorname{sgn}(a)\sqrt{|a|}\omega_1\, ds = d\xi_2$, we obtain

$$\int_{\mathbb{S}} |\langle f, \psi_{a,s,t}\rangle|^2 \, d\mu_{\mathbb{S}}(a,s,t) = \int_{\mathbb{R}} \int_{\mathbb{R}^2} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \, \frac{a^{-2}}{|\omega_1|} \, |\hat{\psi}(a\omega_1, \xi_2)|^2 \, d\xi_2 \, d\omega \, da.$$

Next, we substitute $\xi_1 := a\omega_1$, i.e., $\omega_1\, da = d\xi_1$ which results in

$$\int_{\mathbb{S}} |\langle f, \psi_{a,s,t}\rangle|^2 \, d\mu_{\mathbb{S}}(a,s,t) = \int_{\mathbb{R}} \int_{\mathbb{R}^2} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \, \frac{\omega_1^2}{\xi_1^2 \omega_1^2} \, |\hat{\psi}(\xi_1, \xi_2)|^2 \, d\xi_2 \, d\omega \, d\xi_1 = C_\psi \, \|f\|_2^2.$$

Setting $f := \psi$, we see that ψ is admissible if and only if C_ψ is finite.

Now we show how (6) implies the irreducibility of π. Towards a contradiction, assume that there exists a closed, proper, nontrivial subspace W of $L_2(\mathbb{R}^2)$ such that $\pi(g)W \subseteq W$ for all $g \in \mathbb{S}$. Hence there exist nontrivial functions $\psi \in W$ and $f \in W^\perp$ such that

$$\langle f, \psi_{a,s,t}\rangle = 0 \quad \text{for all } (a,s,t) \in \mathbb{S}.$$

Employing (6) we obtain

$$0 = \int_{\mathbb{S}} |\langle f, \psi_{a,s,t}\rangle|^2 \, \frac{da}{|a|^3} \, ds \, dt = \|f\|_2^2 \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega_1, \omega_2)|^2}{\omega_1^2} \, d\omega_2 \, d\omega_1,$$

which is only possible if

$$0 = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega_1, \omega_2)|^2}{\omega_1^2} \, d\omega_2 \, d\omega_1.$$

This is a contradiction, since $\psi \neq 0$. $\square$

A function $\psi \in L_2(\mathbb{R}^2)$ fulfilling the admissibility condition (5) is called a *continuous shearlet*. Further, the voice transform (1) which we denote for our special case $G = \mathbb{S}$ by $\mathcal{SH}_\psi$ instead of V_ψ is given by $\mathcal{SH}_\psi : L_2(\mathbb{R}^2) \to L_2(\mathbb{S})$,

$$\mathcal{SH}_\psi f(a,s,t) = \langle f, \psi_{a,s,t}\rangle = (f * \psi^*_{a,s,0})(t), \tag{7}$$

and is called *Continuous Shearlet Transform*. An example of a continuous shearlet can be constructed as follows: Let ψ_1 be a continuous wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and $\operatorname{supp} \hat{\psi}_1 \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$, and let ψ_2 be such that $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and $\operatorname{supp} \hat{\psi}_2 \subseteq [-1, 1]$. Then the function $\psi \in L^2(\mathbb{R}^2)$ defined by

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

is a continuous shearlet. The corresponding family of continuous shearlets was exploited in [20] to show that the Continuous Shearlet Transform precisely resolves the wavefront set of a distribution. Generally speaking, for a given (small) *scale* a the Continuous Shearlet Transform provides information about the *location* t and *orientation* s of singularities of f.

3. Coorbit Theory

In this section, we want to briefly recall the basic facts concerning the coorbit theory as developed by Feichtinger and Gröchenig in a series of papers [10, 11, 12, 13]. This theory is based on square-integrable group representations and has the following important advantages:

- The theory is universal in the following sense: Given a Hilbert space $\mathcal{H}$ and a square-integrable representation of a group G, the whole abstract machinery can be applied.

- The approach provides us with natural families of smoothness spaces, the coorbit spaces. They are defined as the collection of all elements in the Hilbert space $\mathcal{H}$ for which the voice transform associated with the group representation has a certain decay. In many cases, e.g., for the affine group and the Weyl-Heisenberg group, these coorbit spaces coincide with classical smoothness spaces such as Besov and modulation spaces, respectively.
- The Feichtinger-Gröchenig theory does not only give rise to Hilbert frames in $\mathcal{H}$, but also to frames in scales of the associated coorbit spaces. Moreover, not only Hilbert spaces, but also Banach spaces can be handled.
- The discretization process that produces the frame does not take place in $\mathcal{H}$ (which might look ugly and complicated), but on the topological group at hand (which is usually a more handy object), and is transported to $\mathcal{H}$ by the group representation.

First of all, in Subsection 3.1, we explain how the coorbit spaces can be established. Then, in Subsection 3.2, we discuss the discretization problem, i.e., we outline the basic steps to construct Banach frames for these spaces. The facts are mainly taken from [15].

3.1. **Coorbit Spaces.** Fix an irreducible, unitary, continuous representation π of a σ-compact group G in a Hilbert space $\mathcal{H}$. Moreover, choose a weight function w, i.e., $w(gh) \leq w(g)w(h)$ and $w(g) \geq 1$ for all $g, h \in G$. Let us assume that the representation π is *w–integrable*, in other words

$$\mathcal{A}_w := \{\psi \in \mathcal{H} : \int_G |\langle \psi, \pi(g)\psi\rangle| w(g) d\mu_G(g) < \infty\} \quad \text{is non-trivial.} \tag{8}$$

Then, the first step is to construct a suitable large object that may serve as the reservoir for the coorbit spaces. For $1 \leq p < \infty$, let

$$L_{p,w}(G) := \{f \text{ measurable on } G : \|f\|_{L_{p,w}(G)} := \left(\int_G |f(g)|^p\, w(g)^p d\mu_G(g)\right)^{1/p} < \infty\},$$

and let $L_{\infty,w}$ be defined with the usual modifications. We consider the space

$$\mathcal{H}_{1,w} := \{f \in \mathcal{H} : V_\psi(f) = \langle f, \pi(\cdot)\psi\rangle \in L_{1,w}(G)\}, \tag{9}$$

with norm

$$\|f\|_{\mathcal{H}_{1,w}} := \|V_\psi f\|_{L_{1,w}}. \tag{10}$$

Then, the anti-dual $\mathcal{H}_{1,w}^\sim$, the space of all continuous conjugate-linear functionals on $\mathcal{H}_{1,w}$, will serve as the reservoir for selection. $\mathcal{H}_{1,w}$ and $\mathcal{H}_{1,w}^\sim$ are π-invariant Banach spaces with continuous embeddings

$$\mathcal{H}_{1,w} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{H}_{1,w}^\sim, \tag{11}$$

and their definition is independent of the analyzing vector $\psi \in \mathcal{A}_w$. Moreover, it follows that $\mathcal{H}_{1,w} = \mathcal{A}_w$ as sets. The inner product on $\mathcal{H} \times \mathcal{H}$ extends to a sesquilinear form on $\mathcal{H}_{1,w}^\sim \times \mathcal{H}_{1,w}$, therefore for $\psi \in \mathcal{H}_{1,w}$ and $f \in \mathcal{H}_{1,w}^\sim$ the *extended representation coefficients*

$$V_\psi(f)(x) := \langle f, \pi(x)\psi\rangle_{\mathcal{H}_{1,w}^\sim \times \mathcal{H}_{1,w}}$$

are well-defined. Now, for $1 \leq p \leq \infty$, we define the *coorbit spaces*

$$\mathcal{H}_{p,w} := \{f \in \mathcal{H}_{1,w}^\sim : V_\psi(f) \in L_{p,w}(G)\} \tag{12}$$

with norms

$$\|f\|_{\mathcal{H}_{p,w}} := \|V_\psi f\|_{L_{p,w}(G)}.$$

Indeed $\mathcal{H}_{1,w}$ is the same space as those defined in (9). Moreover, we have that $\mathcal{H} = \mathcal{H}_{2,1}$. Again, the definition is independent of the analyzing vector ψ and of the weight w in the sense that $\tilde{w}$ with $\tilde{w}(g) \leq Cw(g)$ for all $g \in G$ and with $\mathcal{A}_{\tilde{w}} \neq \{0\}$ gives rise to the same spaces.

3.2. **Banach Frames.** The Feichtinger-Gröchenig theory also provides us with a machinery to construct atomic decompositions and Banach frames for the coorbit spaces introduced above. Some further preparations are necessary. Given a compact set U with non-void interior and $e \in U$, the *set of basic atoms* is defined by

$$\mathcal{B}_w := \{\psi \in \mathcal{H} : \langle \psi, \pi(\cdot)\psi \rangle \in \mathcal{M}(L_{1,w}(G))\}, \tag{13}$$

where

$$\mathcal{M}(L_{1,w}) := \{F \text{ such that } MF(g) := \sup_{u \in gU} |F(u)| \in L_{1,w}(G)\}.$$

It follows that $\mathcal{B}_w \subset \mathcal{H}_{1,w}$. Moreover, a (countable) family $X = (g_\lambda)_{\lambda \in \Lambda}$ in G is said to be *U-dense* if $\cup_{\lambda \in \Lambda} g_\lambda U = G$, and *separated* if for some compact neighborhood Q of e we have $g_i Q \cap g_j Q = \emptyset, i \neq j$, and *relatively separated* if X is a finite union of separated sets. Finally, the *$U-$oscillation* is defined as

$$\operatorname{osc}_U(g) := \sup_{u \in U} |V_\psi(\psi)(ug) - V_\psi(\psi)(g)|. \tag{14}$$

Then, one can show the following decomposition theorem which says that discretizing the representation by means of an U-dense set produces an atomic decomposition for $\mathcal{H}_{p,w}$.

Theorem 3.1. *Assume that the irreducible, unitary representation π is w-integrable and choose $\psi \in \mathcal{B}_w$ normalized by $\|A\psi\| = 1$, where A is defined by* (2)*. Choose a neighborhood U so small that*

$$\|\operatorname{osc}_U\|_{L_{1,w}(G)} < 1. \tag{15}$$

Then for any U-dense and relatively separated set $X = (g_\lambda)_{\lambda \in \Lambda}$ the space $\mathcal{H}_{p,w}$ has the following atomic decomposition: If $f \in \mathcal{H}_{p,w}$, then it has the atomic decomposition

$$f = \sum_{\lambda \in \Lambda} c_\lambda(f)\pi(g_\lambda)\psi \tag{16}$$

where the sequence of coefficients depends linearly on f and satisfies

$$\|(c_\lambda(f))_{\lambda \in \Lambda}\|_{\ell_{p,w}} \leq C\|f\|_{\mathcal{H}_{p,w}} \tag{17}$$

with a constant C depending only on ψ and with $\ell_{p,w}$ being defined by

$$\ell_{p,w} := \{c = (c_\lambda)_{\lambda \in \Lambda} : \|c\|_{\ell_{p,w}} := \|cw\|_{\ell_p} < \infty\},$$

where $w = (w(g_\lambda))_{\lambda \in \Lambda}$. Conversely, if $(c_\lambda(f))_{\lambda \in \Lambda} \in \ell_{p,w}$, then $f = \sum_{\lambda \in \Lambda} c_\lambda \pi(g_\lambda)\psi$ is in $\mathcal{H}_{p,w}$ and

$$\|f\|_{\mathcal{H}_{p,w}} \leq C'\|(c_\lambda(f))_{\lambda \in \Lambda}\|_{\ell_{p,w}}. \tag{18}$$

Given such an atomic decomposition, the problem arises under which conditions a function f is completely determined by its *moments* $\langle f, \pi(g_\lambda)\psi \rangle$ and how f can be reconstructed from these moments. This is answered by the following theorem which establishes the existence of Banach frames.

Theorem 3.2. *Impose the same assumptions as in Theorem* 3.1*. Choose a neighborhood U of e such that*

$$\|\operatorname{osc}_U\|_{L_{1,w}(G)} < 1/\|V_\psi(\psi)\|_{L_{1,w}(G)}. \tag{19}$$

Then, for every U-dense and relatively separated family $X = (g_\lambda)_{\lambda \in \Lambda}$ in G the set $\{\pi(g_\lambda)\psi : \lambda \in \Lambda\}$ is a Banach frame for $\mathcal{H}_{p,w}$. This means that

i) *$f \in \mathcal{H}_{p,w}$ if and only if $(\langle f, \pi(g_\lambda)\psi \rangle_{\mathcal{H}_{1,w}^{\sim} \times \mathcal{H}_{1,w}})_{\lambda \in \Lambda} \in \ell_{p,w}$;*

ii) *there exist two constants* $0 < D \le D' < \infty$ *such that*

$$D\,\|f\|_{\mathcal{H}_{p,w}} \le \|(\langle f,\pi(g_\lambda)\psi\rangle_{\mathcal{H}_{1,w}^{\sim}\times\mathcal{H}_{1,w}})_{\lambda\in\Lambda}\|_{\ell_{p,w}} \le D'\,\|f\|_{\mathcal{H}_{p,w}}; \tag{20}$$

iii) *there exists a bounded, linear reconstruction operator* $\mathcal{S}$ *from* $\ell_{p,w}$ *to* $\mathcal{H}_{p,w}$ *such that* $\mathcal{S}\left((\langle f,\psi(g_\lambda)\psi\rangle_{\mathcal{H}_{1,w}^{\sim}\times\mathcal{H}_{1,w}})_{\lambda\in\Lambda}\right) = f.$

It remains to check how the conditions (15) and (19) can be ensured. One answer is given by the following lemma proved in [15].

Lemma 3.3. *Let* $V_\psi(\psi) \in L_{1,w}$ *and* $\mathrm{osc}_U \in L_{1,w}$ *for one compact neighborhood* U *of* e. *Then we have that* $\psi \in \mathcal{B}_w$. *If, in addition,* $V_\psi(\psi)$ *is continuous, then*

$$\lim_{U\to\{e\}} \|\,\mathrm{osc}_U\,\|_{L_{1,w}(G)} = 0. \tag{21}$$

For further information on coorbit space theory, the reader is referred to [10, 11, 12, 13, 15].

4. Shearlet Coorbit Theory

In this section we want to establish a coorbit theory based on the square-integrable representation (3) of the shearlet group.

4.1. **Shearlet Coorbit Space.** We consider weight functions $w(a,s,t) = w(a,s)$ that are locally integrable with respect to a and s, i.e., $w \in L_1^{loc}(\mathbb{R}^2)$. To obtain well–defined *shearlet coorbit spaces* (12) which we denote for our special setting $G = \mathbb{S}$ by $\mathcal{SC}_{p,w}$ instead of $\mathcal{H}_{p,w}$, more precisely, to get,

$$\mathcal{SC}_{p,w} := \{f \in \mathcal{H}_{1,w}^{\sim} : \mathcal{SH}_\psi(f) \in L_{p,w}(\mathbb{S})\},$$

we have only to ensure that there exists a function $\psi \in L_2(\mathbb{R}^2)$ such that $\mathcal{SH}_\psi(\psi) = \langle\psi,\pi(g)\psi\rangle \in L_{1,w}(\mathbb{S})$. To this end, we need a preliminary lemma on the support of ψ.

Lemma 4.1. *Let* $a_1 > a_0 \ge \alpha > 0$, $b > 0$ *and* $\operatorname{supp}\hat\psi \subseteq ([-a_1,-a_0]\cup[a_0,a_1])\times[-b,b]$. *Then* $\hat\psi\hat\psi_{a,s,0} \not\equiv 0$ *implies* $a \in \big[-\frac{a_1}{a_0},-\frac{a_0}{a_1}\big]\cup\big[\frac{a_0}{a_1},\frac{a_1}{a_0}\big]$ *and* $s\in[-c,c]$, *where* $c := \frac{b}{a_0}\big(1+\big(\frac{a_1}{a_0}\big)^{1/2}\big)$.

Proof. By (4) and since $a > 0$, we see that the following conditions are necessary for $\hat\psi(\omega)\hat\psi_{a,s,0}(\omega) \not\equiv 0$:

$$\text{i)}\quad \begin{array}{lcll} a_0 \le \omega_1 \le a_1 & \text{and} & \frac{a_0}{a} \le \omega_1 \le \frac{a_1}{a} & \text{or} \\ -a_1 \le \omega_1 \le -a_0 & \text{and} & -\frac{a_1}{a} \le \omega_1 \le -\frac{a_0}{a} & \end{array}$$

$$\text{ii)}\quad -b \le \omega_2 \le b \quad \text{and} \quad -a^{-1/2}b - s\omega_1 \le \omega_2 \le a^{-1/2}b - s\omega_1$$

Condition i) implies that

$$a \in \Big[\frac{a_0}{a_1},\frac{a_1}{a_0}\Big]. \tag{22}$$

For $s \ge 0$ and $a_0 \le \omega_1 \le a_1$ the second condition in ii) becomes

$$-a^{-1/2}b - sa_1 \le \omega_2 \le a^{-1/2}b - sa_0$$

and with (22) further

$$-\left(\frac{a_0}{a_1}\right)^{-1/2} b - sa_1 \le \omega_2 \le \left(\frac{a_0}{a_1}\right)^{-1/2} b - sa_0.$$

Together with the first condition in ii) this results in $s \le \frac{b}{a_0}\big(1+\big(\frac{a_1}{a_0}\big)^{1/2}\big)$. The same condition can be deduced for $s \ge 0$ and $-a_1 \le \omega_1 \le -a_0$.

For $s < 0$ and $a_0 \leq \omega_1 \leq a_1$ or $-a_1 \leq \omega_1 \leq -a_0$, we obtain that $s \geq -\frac{b}{a_0}\big(1 + \big(\frac{a_1}{a_0}\big)^{1/2}\big)$ is necessary for $\hat{\psi}(\omega)\hat{\psi}_{a,s,0}(\omega) \not\equiv 0$. Finally, the case $a < 0$ can be treated similarly which results in $a \in \big[-\frac{a_1}{a_0}, -\frac{a_0}{a_1}\big]$. This completes the proof. $\square$

Now we can prove the required property of $\mathcal{SH}_\psi(\psi)$ for the shearlet transform $\mathcal{SH}_\psi$.

Theorem 4.2. *Let ψ be a Schwartz function such that* $\operatorname{supp}\hat{\psi} \subseteq ([-a_1, -a_0] \cup [a_0, a_1]) \times [-b, b]$. *Then we have that $\mathcal{SH}_\psi(\psi) \in L_{1,w}(\mathbb{S})$, i.e.,*

$$\int_{\mathbb{S}} |\mathcal{SH}_\psi(\psi)(g)|\, w(g)\, d\mu(g) < \infty.$$

Proof. Straightforward computation gives

$$\begin{aligned}
\int_{\mathbb{S}} |\mathcal{SH}_\psi(\psi)(g)|\, w(g)\, d\mu(g) &= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}^2} |\langle \psi, \psi_{a,s,t}\rangle|\, w(a,s)\, dt ds \frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}^2} |\psi * \psi^*_{a,s,0}(t)|\, w(a,s)\, dt ds \frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}^2} |\mathcal{F}^{-1}\mathcal{F}\big(\psi * \psi^*_{a,s,0}\big)(t)|\, dt\, w(a,s)\, ds \frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \|\mathcal{F}\big(\psi * \psi^*_{a,s,0}\big)\|_{\mathcal{F}^{-1}L_1}\, w(a,s)\, ds \frac{da}{|a|^3} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \|\hat{\psi}\, \bar{\hat{\psi}}_{a,s,0}\|_{\mathcal{F}^{-1}L_1}\, w(a,s)\, ds \frac{da}{|a|^3},
\end{aligned}$$

where $\psi^* = \bar{\psi}(-\,\cdot)$ and $\|f\|_{\mathcal{F}^{-1}L_1} := \int_{\mathbb{R}^2} |\mathcal{F}^{-1}f(x)|\, dx$ for $f \in L_1$. By Lemma 4.1 this can be rewritten as

$$\int_{\mathbb{S}} |\mathcal{SH}_\psi(\psi)(g)|\, w(g)\; d\mu(g) = \left(\int_{-a_1/a_0}^{-a_0/a_1} + \int_{a_0/a_1}^{a_1/a_0}\right)\int_{-c}^{c} \|\hat{\psi}\, \hat{\psi}^*_{a,s,0}\|_{\mathcal{F}^{-1}L_1}\, w(a,s)\, ds \frac{da}{|a|^3},$$

which is obviously finite. $\square$

4.2. **Shearlet Banach Frames.** In order to find atomic decompositions and Banach frames for our shearlet coorbit spaces $\mathcal{SC}_{p,w}$, we have to determine the corresponding U–dense sets first.

Proposition 4.3. *Let U be a neighborhood of the identity in $\mathbb{S}$, and let $\alpha > 1$ and $\beta, \gamma > 0$ be defined such that*

$$[\alpha^{-\frac{1}{2}}, \alpha^{\frac{1}{2}}) \times [-\tfrac{\beta}{2}, \tfrac{\beta}{2}) \times [-\tfrac{\gamma}{2}, \tfrac{\gamma}{2})^2 \subseteq U. \tag{23}$$

Then the sequence

$$\{(\epsilon\alpha^j, \beta k \alpha^{\frac{j}{2}}, S_{\beta k \alpha^{\frac{j}{2}}} A_{\alpha^j}\gamma m) : j,k \in \mathbb{Z},\, m \in \mathbb{Z}^2,\, \epsilon \in \{-1, 1\}\}$$

is U-dense and relatively separated.

Proof. Let U be a neighborhood of the identity in $\mathbb{S}$. Then the existence of $\alpha > 1$ and $\beta, \gamma > 0$ satisfying (23), follows immediately from the fact that

$$\{[a^{-\frac{1}{2}}, a^{\frac{1}{2}}) \times [-\tfrac{b}{2}, \tfrac{b}{2}) \times [-\tfrac{c}{2}, \tfrac{c}{2})^2 : a > 1,\, b, c > 0\}$$

forms a fundamental system of neighborhoods of the identity in $\mathbb{S}$.

Now set

$$U_0 = [\alpha^{-\frac{1}{2}}, \alpha^{\frac{1}{2}}) \times [-\tfrac{\beta}{2}, \tfrac{\beta}{2}) \times [-\tfrac{\gamma}{2}, \tfrac{\gamma}{2})^2.$$

Observe that, by (23), it is sufficient to prove that the sequence $\{(\epsilon\alpha^j, \beta k\alpha^{\frac{j}{2}}, S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}\gamma m) : j, k \in \mathbb{Z},\ m \in \mathbb{Z}^2,\ \epsilon \in \{-1, 1\}\}$ is U_0-dense.

For this, fix any $(x, y, z) \in \mathbb{S}$. In the following we assume that $x \in \mathbb{R}^+$ in which case we have to set $\epsilon = 1$. If $x < 0$, the same arguments apply while choosing $\epsilon = -1$.

For all $\alpha > 1$ and $\beta, \gamma > 0$,

$$\begin{aligned}
&(\alpha^j, \beta k\alpha^{\frac{j}{2}}, S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}\gamma m) \cdot U_0 \\
&\quad = \{(\alpha^j u, \beta k\alpha^{\frac{j}{2}} + v\alpha^{\frac{j}{2}}, S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}\gamma m + S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j} w) : (u, v, w) \in U_0\} \\
&\quad = \{(\alpha^j u, \alpha^{\frac{j}{2}}(\beta k + v), S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}(\gamma m + w)) : (u, v, w) \in U_0\}.
\end{aligned}$$

Then $[\log_\alpha x - \frac{1}{2}, \log_\alpha x + \frac{1}{2})$ contains a unique integer j, and there exists a unique $u \in [\alpha^{-\frac{1}{2}}, \alpha^{\frac{1}{2}})$ such that $\log_\alpha x = \log_\alpha u + j$. Further, there exist a unique integer k and a unique $v \in [-\frac{\beta}{2}, \frac{\beta}{2})$ so that $\beta k + v = \alpha^{-\frac{j}{2}} y$. Finally, there exist unique $m_2 \in \mathbb{Z}$ and $w_2 \in [-\frac{\gamma}{2}, \frac{\gamma}{2})$ with $\gamma m_2 + w_2 = \alpha^{-\frac{j}{2}} z_2$ and unique $m_1 \in \mathbb{Z}$ and $w_1 \in [-\frac{\gamma}{2}, \frac{\gamma}{2})$ such that $\gamma m_1 + w_1 = \alpha^{-j} z_1 - \beta k(\gamma m_2 + w_2)$. Since

$$S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}(\gamma m + w) = \begin{pmatrix} \alpha^j & \beta k\alpha^j \\ 0 & \alpha^{\frac{j}{2}} \end{pmatrix} (\gamma m + w) = \begin{pmatrix} \alpha^j(\gamma m_1 + w_1) + \beta k\alpha^j(\gamma m_2 + w_2) \\ \alpha^{\frac{j}{2}}(\gamma m_2 + w_2) \end{pmatrix},$$

we have shown that

$$(x, y, z) = (\alpha^j u, \alpha^{\frac{j}{2}}(\beta k + v), S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}(\gamma m + w)) \in (\alpha^j, \beta k\alpha^{\frac{j}{2}}, S_{\beta k\alpha^{\frac{j}{2}}} A_{\alpha^j}\gamma m) \cdot U_0.$$

Finally, the uniqueness of the decomposition proves immediately that the chosen sequence is relatively separated. □

Next, we see that we can apply the whole machinery of Theorems 3.1 and 3.2 to our shearlet group setting if we can prove that $\|\operatorname{osc}_U\|_{L_{1,w}(\mathbb{S})}$ becomes arbitrarily small for a sufficiently small neighborhood U of $e = (1, 0, 0) \in \mathbb{S}$.

Theorem 4.4. *Let ψ be a Schwartz function with $\operatorname{supp}\hat{\psi} \subseteq ([-a_1, -a_0] \cup [a_0, a_1]) \times [-b, b]$. Then, for every $\varepsilon > 0$, there exists a sufficiently small neighborhood U of e so that*

$$\|\operatorname{osc}_U\|_{L_{1,w}(\mathbb{S})} \le \varepsilon. \tag{24}$$

Proof. By Theorem 4.2 we have that $\mathcal{SH}_\psi(\psi) \in L_{1,w}(\mathbb{S})$. Moreover, it is easy to check that $\mathcal{SH}_\psi(\psi)$ is continuous on $\mathbb{S}$. Thus, by Lemma 3.3, it remains to show that $\operatorname{osc}_U \in L_{1,w}$ for some compact neighborhood of e. By definition of osc_U and Parseval's identity we have that

$$\begin{aligned}
\operatorname{osc}_U(a, s, t) &= \sup_{(\alpha,\beta,\gamma)\in U} \left|\langle\hat{\psi}, \hat{\psi}_{a,s,t}\rangle - \langle\hat{\psi}, \hat{\psi}_{(\alpha,\beta,\gamma)(a,s,t)}\rangle\right| \\
&= \sup_{(\alpha,\beta,\gamma)\in U} \left| |a|^{\frac{3}{4}} \mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}} \cdot)\bar{\hat{\psi}}\right)(t) - |a\alpha|^{\frac{3}{4}} \mathcal{F}\left(\hat{\psi}(A_{a\alpha} S_{\beta+s\sqrt{\alpha}}^{\mathrm{T}} \cdot)\bar{\hat{\psi}}\right)(\gamma + S_\beta A_\alpha t)\right|,
\end{aligned}$$

where we can assume that $\alpha > 0$. By Lemma 4.1, we see that for (α, β) in a sufficiently small neighborhood of $(1, 0)$, the function $\hat{\psi}(A_{a\alpha} S_{\beta+s\sqrt{\alpha}}^{\mathrm{T}} \cdot)\bar{\hat{\psi}}$ becomes zero except for values a contained in two finite intervals away from zero and values s in a finite interval. Thus, it remains to show that $\int_{\mathbb{R}^2} \operatorname{osc}_U(a, s, t)\, dt \le C(a, s)$ with a finite constant $C(a, s)$. We split the integral into three parts

$$\int_{\mathbb{R}^2} \operatorname{osc}_U(a, s, t)\, dt = |a|^{\frac{3}{4}}(I_1 + I_2 + I_3),$$

where

$$\begin{aligned}
I_1 &:= \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \big||1-\alpha^{\frac{3}{4}}|\mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(t)\,dt \\
I_2 &:= \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \alpha^{\frac{3}{4}}\big|\mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(t) - \mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(\gamma + S_\beta A_\alpha t)\big|\,dt \\
I_3 &:= \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \alpha^{\frac{3}{4}}\big|\mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(\gamma + S_\beta A_\alpha t) - \mathcal{F}\left(\hat{\psi}(A_{a\alpha} S_{\beta+s\sqrt{\alpha}}^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(\gamma + S_\beta A_\alpha t)\big|\,dt.
\end{aligned}$$

1. Concerning I_1, we see that for every $\varepsilon > 0$ there exists α near 1 such that $|1-\alpha^{\frac{3}{4}}| \le \varepsilon$. Since $\hat{\psi} \in \mathcal{S}$, we have that $\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}} \in \mathcal{S}$ so that $\mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right) \in S$ for any (a,s). Consequently, $I_1 \le \varepsilon C(a,s)$ with a finite constant $C(a,s)$.

2. Concerning I_2, we consider

$$G_{a,s}(t) = G(t) = \mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(t) \in \mathcal{S}.$$

By Taylor expansion we obtain

$$\begin{aligned}
|G(\gamma + S_\beta A_\alpha t) - G(t)| &= |\nabla G\,(t + \theta(\gamma + S_\beta A_\alpha t - t))\,(\gamma + S_\beta A_\alpha t - t)| \\
&\le \|\nabla G\,(t + \theta(S_\beta A_\alpha t - t + \gamma))\|\;\|S_\beta A_\alpha t - t + \gamma\| \\
&\le \|\nabla G\,(t + \theta(S_\beta A_\alpha t - t + \gamma))\|\;(\|S_\beta A_\alpha - I\|\,\|t\| + \|\gamma\|)\,,
\end{aligned}$$

where $\theta \in [0,1)$. For any $\varepsilon > 0$, there exists a sufficiently small neighborhood U of e such that $\|S_\beta A_\alpha - I\| < \varepsilon$ and $\|\gamma\| \le \varepsilon$ for all $(\alpha,\beta,\gamma) \in U$. On the other hand, we have for $t = (t_1,t_2)$ that

$$\begin{aligned}
G_1(t) := \frac{\partial}{\partial t_1} G(t) &= \frac{\partial}{\partial t_1}\mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(t) \\
&= \mathcal{F}\left(-2\pi i\omega_1 \hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\bar{\hat{\psi}}\right)(t) \;=\; \mathcal{F}\left(\hat{\psi}(A_a S_s^{\mathrm{T}}\cdot)\overline{\partial_1\hat{\psi}}\right)(t).
\end{aligned}$$

Since $\psi \in \mathcal{S}$, we obtain that $\frac{\partial}{\partial t_1} G \in \mathcal{S}$ and similarly that $\frac{\partial}{\partial t_2} G \in \mathcal{S}$. Thus, since $\|(G_1,G_2)\| \le |G_1| + |G_2|$, we conclude that

$$I_2 \le \varepsilon \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \alpha^{\frac{3}{4}}\left(|G_1\,(t + \theta(S_\beta A_\alpha t - t + \gamma))\,| + |G_2\,(t + \theta(S_\beta A_\alpha t - t + \gamma))\,|\right)(\|t\| + 1)\;dt$$

Now $G_i \in \mathcal{S}$, $i = 1,2$ implies for all $m > 0$ and sufficiently small γ that

$$\begin{aligned}
|G_i\,(t + \theta(S_\beta A_\alpha t - t + \gamma))\,| &\le C_i(a,s)\;(1 + \|t + \theta(S_\beta A_\alpha t - t + \gamma)\|^2)^{-m} \\
&\le \tilde{C}_i(a,s)\;(1 + \|t + \theta(S_\beta A_\alpha t - t)\|^2)^{-m}.
\end{aligned}$$

To show that $I_2 \le \varepsilon C(a,s)$, we have to prove that

$$\sup_{(\alpha,\beta,0)\in U} (1 + \|t + \theta(S_\beta A_\alpha t - t)\|^2)^{-1} \le \tilde{C}(a,s)\;(1 + \|t\|^2)^{-1}.$$

Straightforward computation gives

$$\|t + \theta(S_\beta A_\alpha t - t)\|^2 = (pt_1 + q\beta t_2)^2 + r^2 t_2^2,$$

where $p := 1 - \theta(1-\alpha)$, $q := \theta\sqrt{\alpha}$ and $r := 1 - \theta(1 - \sqrt{\alpha})$, hence it remains to show that

$$(1 + (pt_1 + q\beta t_2)^2 + r^2 t_2^2)^{-1} \le C(1 + t_1^2 + t_2^2)^{-1}$$

for all $(\alpha,\beta,0) \in U$. The function $g(\beta) := (pt_1 + q\beta t_2)^2 + r^2 t_2^2$ has its minimum at $\beta = -\frac{pt_1}{qt_2}$. Let U be chosen such that $|\beta| \le \beta_0$ with some fixed sufficiently small β_0. If $\left|\frac{pt_1}{qt_2}\right| \le \beta_0$, i.e., $p^2 t_1^2/(q^2\beta_0^2) \le t_2^2$, then

$$\frac{1}{1+(pt_1+q\beta t_2)^2 + r^2 t_2^2} \le \frac{1}{1+r^2t_2^2} \le \frac{1}{\min\{1, r^2p^2/(2q^2\beta_0^2), r^2/2\}}\,\frac{1}{1+t_1^2+t_2^2}.$$

If $\left|\frac{pt_1}{qt_2}\right| > \beta_0$ and t_1, t_2 have the same sign, we see that

$$\frac{1}{1+(pt_1+q\beta t_2)^2 + r^2 t_2^2} \le \frac{1}{1+(pt_1-q\beta_0 t_2)^2 + r^2 t_2^2}.$$

Set $y := pt_1 - q\beta_0 t_2$ and note that y has the same sign as t_1 and t_2. Then $t_1 = (y+q\beta_0 t_2)/p$ and we have to show that

$$\frac{1}{1+y^2+r^2t_2^2} \le \frac{C}{1+(y+q\beta_0t_2)^2/p^2+t_2^2}.$$

Now $2yq\beta_0 t_2 \le y^2 + q^2\beta_0^2 t_2^2$ so that

$$\begin{aligned}
1+(y+q\beta_0t_2)^2/p^2+t_2^2 &\le 1+2y^2/p^2 + (2q^2\beta_0^2/p^2+1)t_2^2 \\
&\le \max\{1, 2/p^2, (1+2q^2\beta_0^2/p^2)/r^2\}\,(1+y^2+r^2t_2^2).
\end{aligned}$$

The case $\left|\frac{pt_1}{qt_2}\right| > \beta_0$ and t_1, t_2 having different signs can be treated in a similar way and we are done.

3. Concerning I_3, we have that

$$I_3 = \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \alpha^{\frac{3}{4}} \Big| \int_{\mathbb{R}^2} \left(\hat\psi(A_a S_s^{\mathrm{T}}\,\omega) - \hat\psi(A_{a\alpha} S^{\mathrm{T}}_{\beta+s\sqrt{\alpha}}\,\omega)\right)\, \bar{\hat\psi}(\omega)\, e^{-2\pi i\omega(\gamma+S_\beta A_\alpha t)}\, d\omega\Big|\, dt.$$

Using the Short Time Fourier Transform defined by

$$G_\psi f(x,\omega) = \int_{\mathbb{R}^2} f(t)\bar\psi(t-x) e^{-2\pi i\omega t}\, dt,$$

this can be rewritten as

$$I_3 = \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \alpha^{\frac{3}{4}} \big| G_{\hat\psi(A_a S_s^{\mathrm{T}}\,\cdot) - \hat\psi(A_{a\alpha} S^{\mathrm{T}}_{\beta+s\sqrt{\alpha}}\,\cdot)} \hat\psi(0, \gamma + S_\beta A_\alpha t)\big|\, dt$$

and since $G_\psi f(x,\omega) = e^{-2\pi i\omega x} G_{\hat\psi} \hat f(\omega, -x)$ further as

$$I_3 = \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} \alpha^{\frac{3}{4}} \big| G_{\psi(A_a^{-1} S_s^{-1}\,\cdot) - \psi(A_{a\alpha}^{-1} S^{-1}_{\beta+s\sqrt{\alpha}}\,\cdot)} \psi(\gamma + S_\beta A_\alpha t, 0)\big|\, dt.$$

By [16, p. 232], we have for $\psi \in \mathcal{S}$ and $|f(x)| \le C(1+\|x\|)^{-s}$, $s > 2$ that

$$|G_\psi f(x,\omega)| \le C \|f\|_{L_{\infty,m_s}} \|\psi\|_{L_{\infty,m_s}} (1+\|x\|)^{-s},$$

where $m_s(x) = (1+\|x\|)^s$ and $\|\psi\|_{L_{\infty,m_s}} = \operatorname*{ess\,sup}_{x\in\mathbb{R}^2} |\psi(x)|(1+\|x\|)^s$. Thus, since $\psi \in \mathcal{S}$,

$$\begin{aligned}
I_3 &\le C\,\|\psi\|_{L_{\infty,m_s}} \sup_{(\alpha,\beta,0)\in U} \|\psi(A_a^{-1}S_s^{-1}\,\cdot) - \psi(A_{a\alpha}^{-1} S^{-1}_{\beta+s\sqrt{\alpha}}\,\cdot)\|_{L_{\infty,m_s}} \\
&\quad \times \int_{\mathbb{R}^2} \sup_{(\alpha,\beta,\gamma)\in U} (1+\|\gamma + S_\beta A_\alpha t\|)^{-s}\, dt \;\le\; C(a,s).
\end{aligned}$$

This completes the proof. $\square$

4.3. **Embedding of Schwartz Spaces.** Let $\mathcal{S}_0$ denote the space of Schwartz-functions with the property that

$$|\hat{f}(\omega)| \le \frac{\omega_1^{2\alpha}}{(1+\|\omega\|^2)^{2\alpha}} \quad \text{for all } \alpha > 0.$$

The functions in $\mathcal{S}_0$ vanish of infinite order in $\omega_1 = 0$. In this subsection, we want to show that the space $\mathcal{S}_0$ is contained in our coorbit shearlet spaces $\mathcal{SC}_{p,w}$. To this end, we need the following two preliminary lemmas. The proof of the first lemma uses ideas of [19].

Lemma 4.5. *For $\alpha > 1$, let $|f(x)| \le \frac{C}{(1+\|x\|^2)^\alpha}$ and $|\psi(x)| \le \frac{C}{(1+\|x\|^2)^\alpha}$. Then the shearlet transform fulfils*

$$|\mathcal{SH}_\psi f(a,s,t)| \le C|a|^{-3/4} \frac{\max\{1,d^2\}}{\left(1+\|\frac{t}{\max\{1,d\}}\|^2\right)^{\alpha-1/2}},$$

where $d^2 := (s^2+2)\,\max\{a^2,|a|\}$.

Proof. We restrict our attention to the case $a > 0$. By assumption on f and ψ we have that

$$|\mathcal{SH}_\psi f(a,s,t)| \le Ca^{-3/4} \int_{\mathbb{R}^2} \frac{1}{(1+\|x\|^2)^\alpha} \; \frac{1}{1+\|A_a^{-1}S_s^{-1}(x-t)\|^2)^\alpha}\, dx$$

and since $\|A_a^{-1}S_s^{-1}(x-t)\| \ge \frac{\|x-t\|}{\|A_a\|\|S_s\|}$ further that

$$|\mathcal{SH}_\psi f(a,s,t)| \le Ca^{-3/4} \int_{\mathbb{R}^2} \frac{1}{(1+\|x\|^2)^\alpha} \frac{1}{\left(1+\frac{\|x-t\|^2}{\|A_a\|^2\|S_s\|^2}\right)^\alpha}\, dx.$$

Using that $\|A_a\|^2 = \max\{a,a^2\}$ and $\|S_s\|^2 = \rho(S_s^T S_s) = \rho\left(\begin{pmatrix} 1 & s \\ s & s^2+1 \end{pmatrix}\right) \le s^2+2$ this can be estimated as

$$|\mathcal{SH}_\psi f(a,s,t)| \le Ca^{-3/4} \int_{\mathbb{R}^2} \frac{1}{(1+\|x\|^2)^\alpha} \frac{1}{\left(1+\frac{\|x-t\|^2}{d^2}\right)^\alpha}\, dx.$$

We consider

$$F(t,d) := \int_{\mathbb{R}^2} \frac{1}{(1+\|x\|^2)^\alpha} \; \frac{1}{\left(1+\frac{\|x-t\|^2}{d^2}\right)^\alpha}\, dx.$$

We show that for $d \le 1$ the estimate

$$F(t,d) \le \frac{C}{(1+\|t\|^2)^{\alpha-1/2}} \tag{25}$$

holds true. Since

$$F\left(\frac{t}{d},\frac{1}{d}\right) = \frac{1}{d^2} F(t,d),$$

this implies for $d > 1$ that

$$|F(t,d)| = d^2 \left| F\left(-\frac{t}{d},\frac{1}{d}\right)\right| \le d^2 \frac{C}{(1+\|\frac{t}{d}\|^2)^{\alpha-1/2}}$$

and we are done.

It remains to prove (25) for $d \le 1$. For $\|t\| < 1$ we only have to show that $F(t,d)$ is bounded independently of d which is trivial. In the following let $\|t\| \ge 1$. We split the integral as follows:

$$F(t,d) = \Big(\int_{\Omega_1} + \int_{\Omega_2} + \int_{\Omega_{31}} + \int_{\Omega_{32}}\Big) \frac{1}{(1+\|x\|^2)^\alpha} \; \frac{1}{1+\|\frac{x-t}{d}\|^2)^\alpha}\, dx = I_1 + I_2 + I_{31} + I_{32}$$

where $\Omega_1 := \{x \in \mathbb{R}^2 : t(x - \frac{t}{2}) > 0 \text{ componentwise}\}$, $\Omega_2 := \{x \in \mathbb{R}^2 : t(x - \frac{t}{2}) \leq 0 \text{ componentwise}\}$ and $\Omega_{3i} := \{x \in \mathbb{R}^2 : t_i(x_i - \frac{t_i}{2}) > 0 \text{ and } t_j(x_j - \frac{t_j}{2}) \leq 0, j \neq i\}$, $i = 1, 2$.

Estimation of I_1: In Ω_1 we have that $|x_i| \geq \frac{|t_i|}{2}$ and thus $\|x\|^2 \geq \frac{1}{4}\|t\|^2$. Consequently, we get

$$\begin{aligned} I_1 &\leq \int_{\Omega_1} \frac{1}{(1 + \frac{1}{4}\|t\|^2)^\alpha} \frac{1}{(1 + \|\frac{x-t}{d}\|^2)^\alpha}\, dx \;\leq\; C\, \frac{1}{(1 + \|t\|^2)^\alpha} \int_{\Omega_1} \frac{1}{(1 + \|\frac{x-t}{d}\|^2)^\alpha}\, dx \\ &\leq C\, d^2 \frac{1}{(1 + \|t\|^2)^\alpha} \int \frac{1}{(1 + \|y\|^2)^\alpha}\, dy \;\leq\; C\, \frac{1}{(1 + \|t\|^2)^\alpha}. \end{aligned}$$

Estimation of I_2: In Ω_2, it holds $|x_i - t_i| \geq \frac{|t_i|}{2}$, i.e., $\|x - t\|^2 \geq \|t\|^2/4$ and thus

$$I_2 \leq \int_{\Omega_2} \frac{1}{(1 + \|x\|^2)^\alpha} \frac{1}{(1 + \|\frac{t}{2d}\|^2)^\alpha}\, dx.$$

Using that

$$\frac{1}{(1 + \|\frac{t}{2d}\|)^\alpha} = \left(\frac{1 + \|t\|^2}{1 + \frac{\|t\|^2}{4d^2}}\right)^\alpha \frac{1}{(1 + \|t\|^2)^\alpha} \leq C\, \frac{1}{(1 + \|t\|^2)^\alpha}$$

we obtain the estimate for I_2.

Estimation of I_{31}. By assumption we have as in the previous cases that

$$\begin{aligned} I_{32} &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\left(1 + \frac{t_1^2}{4} + x_2^2\right)^\alpha} \frac{1}{\left(1 + \left(\frac{x_1 - t_1}{d}\right)^2 + \frac{t_2^2}{4d^2}\right)^\alpha}\, dx_1 dx_2 \\ &= \frac{1}{\left(1 + \frac{t_1^2}{4}\right)^\alpha \left(1 + \frac{t_2^2}{4d^2}\right)^\alpha} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\left(1 + \frac{x_2^2}{1 + \frac{t_1^2}{4}}\right)^\alpha} \frac{1}{\left(1 + \frac{(x_1 - t_1)^2}{d^2\left(1 + \frac{t_2^2}{4d^2}\right)}\right)^\alpha}\, dx_1 dx_2 \\ &\leq \frac{C}{(1 + t_1^2)^\alpha (1 + t_2^2)^\alpha} \cdot d\sqrt{1 + \frac{t_2^2}{4d^2}} \sqrt{1 + \frac{t_1^2}{4}} \int_{\mathbb{R}} \frac{1}{(1 + y_2^2)^\alpha}\, dy_2 \int_{\mathbb{R}} \frac{1}{(1 + y_1^2)^\alpha}\, dy_1 \\ &\leq \frac{C}{(1 + t_1^2)^{\alpha - 1/2} (1 + t_2^2)^{\alpha - 1/2}} \;\leq\; \frac{C}{(1 + \|t\|^2)^{\alpha - 1/2}}\,. \end{aligned}$$

Analogously we can obtain the estimate for I_{32}. This finishes the proof. □

Another estimate of the shearlet transform is given by the following lemma.

Lemma 4.6. *Let* $\operatorname{supp} \hat{\psi} \in ([-a_1, -a_0] \cup [a_0, a_1]) \times [-b, b]$, *where* $0 < a_0 < a_1$ *and* $b > 0$ *with bounded* $\hat{\psi}$ *and let* $\hat{f}$ *fulfill*

$$|\hat{f}(\omega)| \leq \frac{\omega_1^{2\alpha}}{(1 + \|\omega\|^2)^{2\alpha}}\,, \quad \alpha > 0.$$

Then the following estimate holds true:

$$|\mathcal{SH}_\psi f(a, s, t)| \leq C |a|^{-\frac{3}{4}} \frac{|a|^\alpha}{(1 + a^2)^\alpha}\; \frac{1}{(1 + |s|)^\alpha}\,.$$

Proof. Again, let $a > 0$. By definition of the shearlet transform (7), Parseval's equality and (4) we obtain

$$
\begin{aligned}
|\mathcal{SH}_\psi f(a,s,t)| &\le \int_{\mathbb{R}^2} |\hat f(\omega)||\hat\psi(A_a S_s^T\omega)| a^{3/4} d\omega \\
&\le a^{-3/4}\Big(\int_{a_0}^{a_1} + \int_{-a_1}^{-a_0}\Big)\int_{-b}^{b} |\hat f(S_s^{-T}A_a^{-1}v)||\hat\psi(v)| dv \\
&\le a^{-3/4}\Big(\int_{a_0}^{a_1} + \int_{-a_1}^{-a_0}\Big)\int_{-b}^{b} |\hat f(S_s^{-T}A_a^{-1}v)| dv\,.
\end{aligned}
$$

We restrict our attention to the integral

$$
\begin{aligned}
I &= \int_{a_0}^{a_1}\int_{-b}^{b} \frac{(\frac{v_1}{a})^{2\alpha}}{(1+(\frac{v_1}{a})^2 + (\frac{v_2}{\sqrt a} - s\frac{v_1}{a})^2)^{2\alpha}}\, dv_2 dv_1 \\
&= \int_{a_0}^{a_1}\int_{-b}^{b} \frac{v_1^{2\alpha}}{(a + \frac{v_1^2}{a} + (v_2 - \frac{s}{\sqrt a}v_1)^2)^{2\alpha}}\, dv_2 dv_1\,.
\end{aligned}
$$

The estimation for the other integral follows analogously.

For $|s| \le 1$ we have

$$
I \le C\frac{a_1^{2\alpha}}{\left(a+\frac{a_0^2}{a}\right)^{2\alpha}} \le C\frac{a^{2\alpha}}{(1+a^2)^{2\alpha}} \le C\frac{a^{2\alpha}}{(1+a^2)^{2\alpha}(1+|s|)^{\alpha}} \le C\frac{a^{\alpha}}{(1+a^2)^{\alpha}}\frac{1}{(1+|s|)^{\alpha}}.
$$

In the following, let $|s| > 1$. Set $c := \frac{s}{\sqrt a}$.

Case 1. For $|c| \le \frac{2b}{a_0} = C_1$, i.e., $1 < |s| \le C_1\sqrt a$, we can estimate

$$
I \le C\frac{a_1^{2\alpha}}{(a+\frac{a_0^2}{a})^{2\alpha}} \le C\frac{1}{a^{2\alpha}(1+\frac{a_0}{a^2})^{2\alpha}} \le C\frac{1}{a^{2\alpha}} \le C\frac{1}{a^{\alpha}}\frac{1}{(1+|s|)^{\alpha}} \le C\frac{a^{\alpha}}{(1+a^2)^{\alpha}}\frac{1}{(1+|s|)^{\alpha}}\,.
$$

Case 2. For $|c| > C_1$, i.e., $|s| > C_1\sqrt a$ we obtain

$$
I \le C\frac{a_1^{2\alpha}}{(a+\frac{a_0^2}{a} + (b - |c|a_0)^2)^{2\alpha}} = C\left(\frac{a}{s^2}\right)^{2\alpha}\frac{1}{\left(\frac{a^2}{s^2} + \frac{a_0^2}{s^2} + (\frac{b}{|c|} - a_0)^2\right)^{2\alpha}}\,.
$$

If $a \le 1$, then

$$
I \le \frac{a^{2\alpha}}{(1+a^2)^{2\alpha}}\,\frac{1}{(1+s^2)^{2\alpha}} \le C\frac{a^{\alpha}}{(1+a^2)^{\alpha}}\,\frac{1}{(1+|s|)^{\alpha}}\,.
$$

Let $a > 1$. If $\frac{a^2}{s^2} \le 1$, i.e., $a^2 \le s^2$, then we obtain

$$
I \le C\frac{a^{2\alpha}}{s^{2\alpha}s^{2\alpha}} \le C\frac{1}{s^{2\alpha}} \le \frac{1}{|s|^{\alpha}}\,\frac{1}{a^{\alpha}} \le C\frac{1}{(1+|s|)^{\alpha}}\,\frac{a^{\alpha}}{(1+a^2)^{\alpha}}\,.
$$

If $\frac{a^2}{s^2} > 1$, then $a^2 > s^2 > 1$ and we can estimate

$$
I \le C\frac{1}{a^{2\alpha}}\,\frac{1}{\left(1+\frac{a_0^2}{a^2} + (\frac{b}{|c|} - a_0)^2\,\frac{s^2}{a^2}\right)^{2\alpha}} \le C\frac{1}{a^{\alpha}}\,\frac{1}{|s|^{\alpha}} \le C\frac{a^{\alpha}}{(1+a^2)^{\alpha}}\,\frac{1}{(1+|s|)^{\alpha}}\,.
$$

This completes the proof. □

By Lemma 4.5 and 4.6 we obtain the following theorem.

Theorem 4.7. *Let ψ be a Schwartz-function with $\operatorname{supp}\hat{\psi} \subseteq ([-a_1,-a_0] \cup [a_0,a_1]) \times [-b,b]$ where $0 < a_0 < a_1$ and $b > 0$. Then we have for any Schwartz-function $f \in \mathcal{S}_0$ that*

$$|\mathcal{SH}_\psi f(a,s,t)| \le C|a|^{-3/4} \frac{\max\{1,d\}}{\left(1+\|\frac{t}{\max\{1,d\}}\|^2\right)^{\alpha-1/4}} \frac{|a|^\alpha}{(1+a^2)^\alpha(1+|s|)^\alpha} \quad \text{for all } \alpha > 1,$$

where $d^2 := (s^2+2)\max\{a^2,|a|\}$ As a consequence we obtain that $\mathcal{S}_0 \subset \mathcal{SC}_{p,w}$ for $w(a,s,t) = w(a,s) = (1+|a|)^m(1+|s|)^n$, $n,m > 0$.

Proof. By Lemma 4.5 and 4.6 we verify with 2α instead of α that

$$\begin{aligned} |\mathcal{SH}_\psi f(a,s,t)|^2 &\le C|a|^{-3/4} \frac{\max\{1,d^2\}}{\left(1+\|\frac{t}{\max\{1,d\}}\|^2\right)^{2\alpha-1/2}} |a|^{-3/4} \frac{|a|^{2\alpha}}{(1+a^2)^{2\alpha}(1+|s|)^{2\alpha}}, \\ |\mathcal{SH}_\psi f(a,s,t)| &\le C|a|^{-3/4} \frac{\max\{1,d\}}{\left(1+\|\frac{t}{\max\{1,d\}}\|^2\right)^{\alpha-1/4}} \frac{|a|^\alpha}{(1+a^2)^\alpha(1+|s|)^\alpha}. \end{aligned}$$

Finally, we conclude for sufficiently large α that

$$\begin{aligned} &\int_{\mathbb{S}} |\mathcal{SH}_\psi f(a,s,t)|^p (1+|a|)^{mp}(1+|s|)^{np}\, dtds \frac{da}{|a|^3} \\ \le\ & C \int_{\mathbb{R}^*}\int_{\mathbb{R}}\int_{\mathbb{R}^2} \frac{\max\{1,d\}^p}{\left(1+\|\frac{t}{\max\{1,d\}}\|^2\right)^{p(\alpha-1/4)}}\, dt\, \frac{|a|^{-\frac{3p}{4}}|a|^{p\alpha}}{(1+a^2)^{p\alpha}(1+|s|)^{p\alpha}}\, (1+|a|)^{mp}(1+|s|)^{np}\, ds \frac{da}{|a|^3} \\ \le\ & C \int_{\mathbb{R}^*}\int_{\mathbb{R}} \frac{|a|^{p(\alpha-\frac{3}{4})-3}\max\{1,d^{p+2}\}}{(1+a^2)^{p\alpha}(1+|s|)^{p\alpha}}\, (1+|a|)^{mp}(1+|s|)^{np} dsda \int_{\mathbb{R}^2} \frac{1}{(1+\|y\|^2)^{p(\alpha-\frac{1}{4})}}\, dy \\ <\ & \infty \end{aligned}$$

and we are done. ∎

4.4. **Non-Linear Approximation.** In Section 4.2 we established atomic decompositions of functions from the shearlet coorbit spaces $\mathcal{SC}_{p,w}$ by means of special discretized shearlet systems $(\psi_\lambda)_{\lambda\in\Lambda}$, $\Lambda \subset \mathbb{S}$. From the computational point of view, this naturally leads us to the question of the quality of approximating schemes in $\mathcal{SC}_{p,w}$ using only a finite number of elements from $(\psi_\lambda)_{\lambda\in\Lambda}$.

In this section we will focus on the non-linear approximation scheme of *best N–term approximation*, i.e., of approximating functions f of $\mathcal{SC}_{p,w}$ in an "optimal" way by a linear combination of precisely N elements from $(\psi_\lambda)_{\lambda\in\Lambda}$. In order to study the quality of best N–term approximation we will prove estimates for the asymptotic behavior of the approximation error.

Let us now delve more into the specific setting we are considering here. Let U be a neighborhood of e in $\mathbb{S}$ satisfying condition (24) for $\varepsilon < 1$. Further, let $\Lambda \subset \mathbb{S}$ be a relatively separated, U-dense sequence, which exists by Proposition 4.3. Then the associated shearlet system

$$\{\psi_\lambda = \psi_{a,s,t} : \lambda = (a,s,t) \in \Lambda\} \tag{26}$$

can be employed for atomic decompositions of elements from $\mathcal{SC}_{p,w}$, where $1 \le p < \infty$. Indeed, by Theorems 3.1 and 3.2, for any $f \in \mathcal{SC}_{p,w}$, we have

$$f = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda \tag{27}$$

with $(c_\lambda)_{\lambda\in\Lambda}$ depending linearly on f, and

$$C_{1,p}\,\|f\|_{\mathcal{SC}_{p,w}} \le \|(c_\lambda)_{\lambda\in\Lambda}\|_{\ell_{p,w}} \le C_{2,p}\,\|f\|_{\mathcal{SC}_{p,w}} \tag{28}$$

with constants $C_{1,p}, C_{2,p}$ being independent of f. We intend to approximate functions f from the shearlet coorbit spaces $\mathcal{SC}_{p,w}$ by elements from the nonlinear manifolds Σ_n, $n \in \mathbb{N}$, which consist of all functions $S \in \mathcal{SC}_{p,w}$ whose expansions with respect to the shearlet system $(\psi_\lambda)_{\lambda\in\Lambda}$ from (26) have at most n nonzero coefficients, i.e.,

$$\Sigma_n := \left\{ S \in \mathcal{SC}_{p,w} : S = \sum_{\lambda\in\Gamma} d_\lambda \psi_\lambda,\ \Gamma \subseteq \Lambda,\ \#\Gamma \le n \right\}.$$

Then we are interested in the asymptotic behavior of the error

$$E_n(f)_{\mathcal{SC}_{p,w}} := \inf_{S\in\Sigma_n} \|f - S\|_{\mathcal{SC}_{p,w}}.$$

Usually, the order of approximation which can be achieved depends on the regularity of the approximated function as measured in some associated smoothness space. For instance, for nonlinear wavelet approximation, the order of convergence is determined by the regularity as measured in a specific scale of Besov spaces. For nonlinear approximation based on Gabor frames, it has been shown in [17] that the 'right' smoothness spaces are given by a specific scale of modulation spaces. An extension of these relations to systems arising from the Weyl-Heisenberg group and α-modulation spaces has been studied in [4].

In our case it turns out that a result from [17], i.e., an estimate in one direction, carries over. The basic ingredient in the proof of the theorem is the following lemma which has been shown in [17], see also [8].

Lemma 4.8. *Let $0 < p < q \le \infty$. Then there exists a constant $D_p > 0$ independent of q such that, for all decreasing sequences of positive numbers $a = (a_i)_{i=1}^\infty$, we have*

$$2^{-1/p}\|a\|_{\ell_p} \le \left(\sum_{n=1}^\infty \frac{1}{n} (n^{1/p-1/q} E_{n,q}(a))^p \right)^{1/p} \le D_p\,\|a\|_{\ell_p},$$

where $E_{n,q}(a) := (\sum_{i=n}^\infty a_i^q)^{1/q}$.

Now one can prove the following theorem, which provides an upper estimate for the asymptotic behavior of $E_n(f)_{\mathcal{SC}_{p,w}}$.

Theorem 4.9. *Let $(\psi_\lambda)_{\lambda\in\Lambda}$ be a discrete shearlet system as in (26), and let $1 \le p < q < \infty$. Then there exists a constant $C = C(p,q) < \infty$ such that, for all $f \in \mathcal{SC}_{p,w}$, we have*

$$\left(\sum_{n=1}^\infty \frac{1}{n} \left(n^{1/p-1/q} E_n(f)_{\mathcal{SC}_{q,w}} \right)^p \right)^{1/p} \le C\|f\|_{\mathcal{SC}_{p,w}}.$$

Proof. Let $\tau : \mathbb{N} \to \Lambda$ permutate the sequence $(c_\lambda w_\lambda)_{\lambda\in\Lambda}$ in (27) in a decreasing order, i.e., $|c_{\tau(n)} w_{\tau(n)}| \ge |c_{\tau(n+1)} w_{\tau(n+1)}|$ for all $n \in \mathbb{N}$. Then we obtain

$$E_n(f)_{\mathcal{SC}_{q,w}} \le \| \sum_{i=n+1}^\infty c_{\tau(i)} \psi_{\tau(i)} \|_{\mathcal{SC}_{q,w}}.$$

Applying (28) yields

$$E_n(f)_{\mathcal{SC}_{q,w}} \le C_{1,q}^{-1} \left(\sum_{i=n+1}^{\infty} |c_{\tau(i)} w_{\tau(i)}|^q \right)^{1/q} = C_{1,q}^{-1}\, E_{n+1,q}(|c_{\tau(i)} w_{\tau(i)}|) \le C_{1,q}^{-1}\, E_{n,q}(|c_{\tau(i)} w_{\tau(i)}|).$$

Finally, by Lemma 4.8 and (28),

$$\begin{aligned}\left(\sum_{n=1}^{\infty} \frac{1}{n} \left(n^{1/p-1/q} E_n(f)_{\mathcal{SC}_{q,w}} \right)^p \right)^{1/p} &\le \left(\sum_{n=1}^{\infty} \frac{1}{n} \left(n^{1/p-1/q}\, C_{1,q}^{-1}\, E_{n,q}(|c_{\tau(i)} w_{\tau(i)}|) \right)^p \right)^{1/p} \\ &\le C_{1,q}^{-1} D_p \, \|(|c_{\tau(i)}|)_{i=1}^{\infty}\|_{\ell_{p,w}} \\ &\le C_{1,q}^{-1} C_{2,p} D_p \, \|f\|_{\mathcal{SC}_{p,w}},\end{aligned}$$

which finishes the proof. □

5. Numerical Experiments

In Section 4, in which we have established the coorbit theory for the shearlet group, we have shown that a suitable discretization process produces atomic decompositions and frames in scales of shearlet coorbit spaces. This theoretical result shall be now verified in practice, i.e., we aim to derive an atomic decomposition by means of a shearlet system (for simplicity we abstain here from doing the same by means of a shearlet Banach frame). The basic ingredients are Proposition 4.3 providing a construction principle for the U-dense and relatively separated grid Λ and Theorem 4.4 in which we have shown that the essential condition, $\|\operatorname{osc}_U\|_{L_{1,w}(\mathbb{S})} < 1$, can indeed be fulfilled. Remember, this condition ensures that each function in the associated shearlet coorbit space has an atomic decomposition, i.e. for each $f \in \mathcal{SC}_{p,w}$ there exists a sequence $(c_\lambda)_{\lambda\in\Lambda}$ such that $f = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda$ with $\|f\|_{\mathcal{SC}_{p,w}} \sim \|c\|_{\ell_{p,w}(\Lambda)}$.

For sake of simplicity we consider the case $p = 2$ and $w \equiv 1$. In order to compute the shearlet decomposition of some given function/image f we introduce an operator F via $c \mapsto Fc = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda$. Then finding the right sequence $(c_\lambda)_{\lambda\in\Lambda}$ can be achieved by solving the inverse problem

$$f = Fc \; . \tag{29}$$

Since a direct inversion in (29) is impossible, we suggest to minimize the following cost functional

$$\Phi(c) = \|f - Fc\|_{\mathcal{SC}_2}^2 = \|V_\psi(f) - V_\psi(Fc)\|_{L_2(\mathbb{S})}^2 \; , \tag{30}$$

where the necessary condition for a minimum of (30) is given by the normal equation

$$(V_\psi F)^* V_\psi F c = (V_\psi F)^* V_\psi f \; , \tag{31}$$

which can be solved iteratively. Since the system $\{\psi_\lambda\}_{\lambda\in\Lambda}$ is usually an overcomplete shearlet system, one can find many different sequences c satisfying (31). A few of them have special properties for which they are preferred, e.g., a sequence with minimal ℓ_2 norm (often referred to as the generalized solution). A suitable iteration to approach a solution of (31) is the so-called Landweber iteration,

$$c^{m+1} = (I - \beta (V_\psi F)^* V_\psi F) c^m + \beta (V_\psi F)^* V_\psi f \; , \tag{32}$$

with $0 < \beta < 2/\|V_\psi F\|^2$. If $c^\dagger$ denotes the generalized solution of (29), then iteration (32) converges for arbitrarily chosen c^0 and exact right hand side $(V_\psi F)^* V_\psi f$ towards $P_{\ker(F)} c^0 + c^\dagger$. If, in particular, $c^0 \in \ker(F)^\perp$ (e.g. $c^0 = 0$ or $c^0 = (V_\psi F)^* V_\psi f$), then the iteration converges to $c^\dagger$.

As we have experienced, the grid Λ has for usual configurations (images of size 256x256 pixel and adequate choices of α, β and γ) a remarkable complexity. Therefore it is reasonable to search for sequences c that are sparser than the minimum ℓ_2 norm solution. This may dramatically reduce the computational complexity for computing approximations to c. To this end, we consider a slightly modified variational functional that ensures sparsity on c,

$$\Phi_q(c) = \|f - Fc\|^2_{\mathcal{SC}_2} + 2\rho\|c\|_{\ell_q} \ , \tag{33}$$

with $1 \leq q \leq 2$. As it can be retrieved in [7], an iteration approaching the minimizer of (33) is again given by a Landweber iteration, but now a shrinkage operation is applied in each step,

$$c^{m+1} = S_{q,\rho\beta}\left((I - \beta(V_\psi F)^* V_\psi F)c^m + \beta(V_\psi F)^* V_\psi f\right) \ , \tag{34}$$

where $0 < \beta < \|V_\psi F\|^{-2}$ and $S_{q,\rho\beta}$ denotes the shrinkage operator with respect to the ℓ_q norm and threshold $\rho\beta$ (for details we refer to [7]).

As a test image for which we aim to derive a sparse atomic shearlet decomposition we choose the eye–image in Figure 1. The basic shearlet atom is defined in the Fourier domain by $\hat{\psi}(\omega_1, \omega_2) = \phi_1(\omega_1)\phi_2(\omega_2)$, where ϕ_1 is some smooth variant of Meyer's wavelet (for its construction see [6]) and ϕ_2 is the Gaussian. As the initial iterate c^0 for iteration (34) we choose $F^* f$, where the computation of the coefficients $\mathcal{SH}_\psi(f)(a,s,t) = \langle f, \psi_{a,s,t}\rangle$ is realized via the convolution theorem (for a detailed discussion we refer to [21]),

$$\mathcal{SH}_\psi(f)(a,s,\cdot)^\wedge(\omega) = \hat{f}(\omega)\hat{\psi}^*_{a,s,0}(\omega) \ . \tag{35}$$

The values of (35) for several dilations a and shear values s are visualized in Figures 2 and 3 (note that c^0 is just a subset). The directional selectivity of the shearlet transform can clearly be recognized. As the final result, we present in Figure 4 several iterates Fc^m.

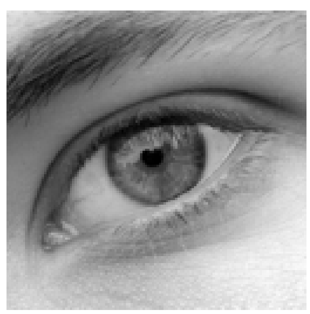

FIGURE 1. Original function/image f.

REFERENCES

[1] L. Borup and M. Nielsen, Frame decomposition of decomposition spaces, J. Fourier Anal. Appl., to appear.

[2] E. J. Candès and D. L. Donoho, *Ridgelets: a key to higher-dimensional intermittency?,* Phil. Trans. R. Soc. Lond. A. **357** (1999), 2495–2509.

[3] E. J. Candès and D. L. Donoho, *Curvelets - A surprisingly effective nonadaptive representation for objects with edges,* in Curves and Surfaces, L. L. Schumaker et al., eds., Vanderbilt University Press, Nashville, TN (1999).

[4] S. Dahlke, M. Fornasier, H. Rauhut, G. Steidl, and G. Teschke, *Generalized coorbit theory, Banach frames, and the relations to alpha–modulation spaces,* Preprint (2005).

[5] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke, *The uncertainty principle associated with the continuous shearlet transform*, IJWMIP, to appear.

[6] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia (1992).

[7] I. Daubechies, M. Defrise and C. DeMol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math. **57** (2004) 1413-1541.

[8] R. DeVore and V.N. Temlyakov, *Some remarks on greedy algorithms,* Adv. in Comput.Math. **5** (1996) 173–187.

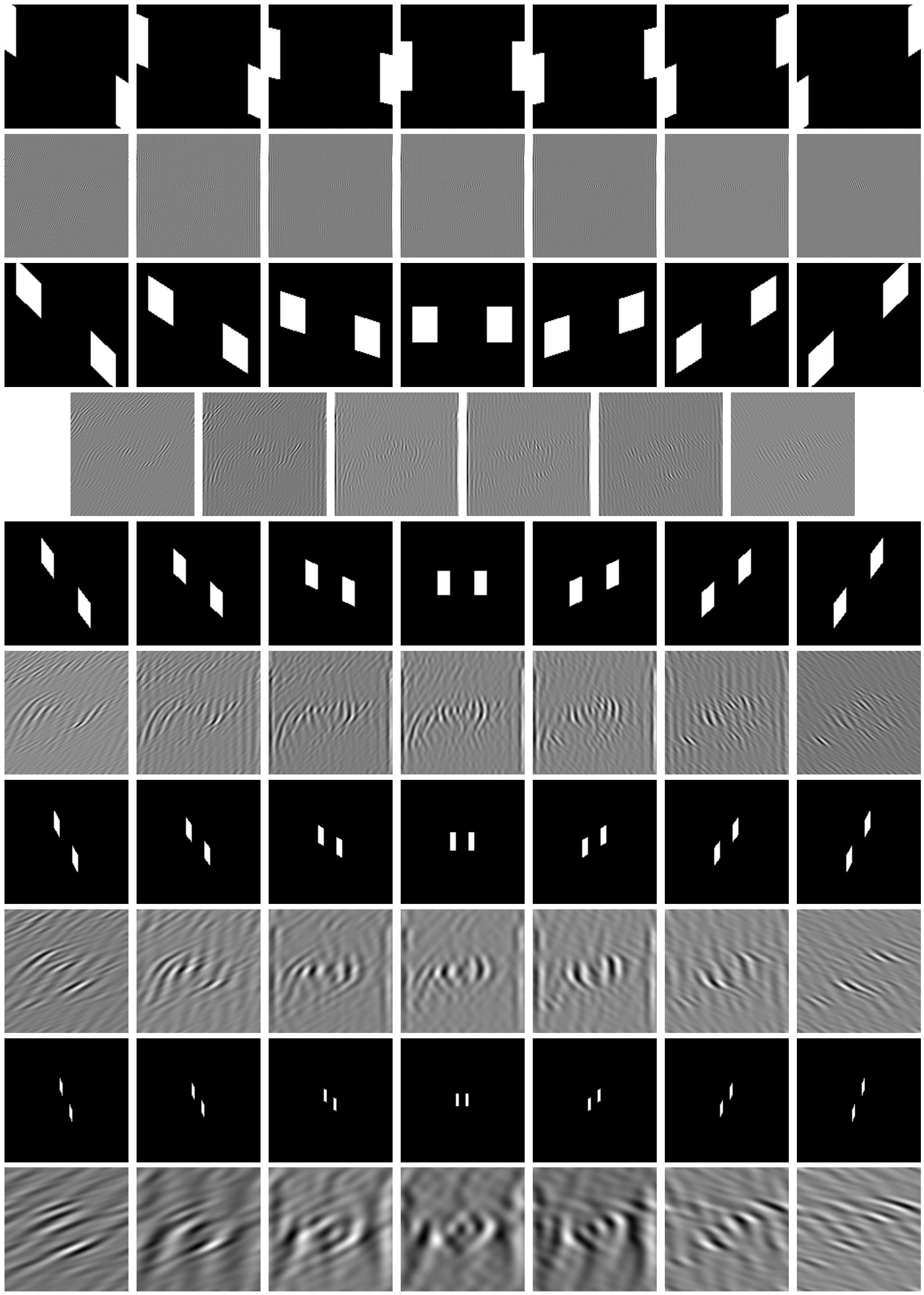

FIGURE 2. Shearlet coefficients for different values of a and s.

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

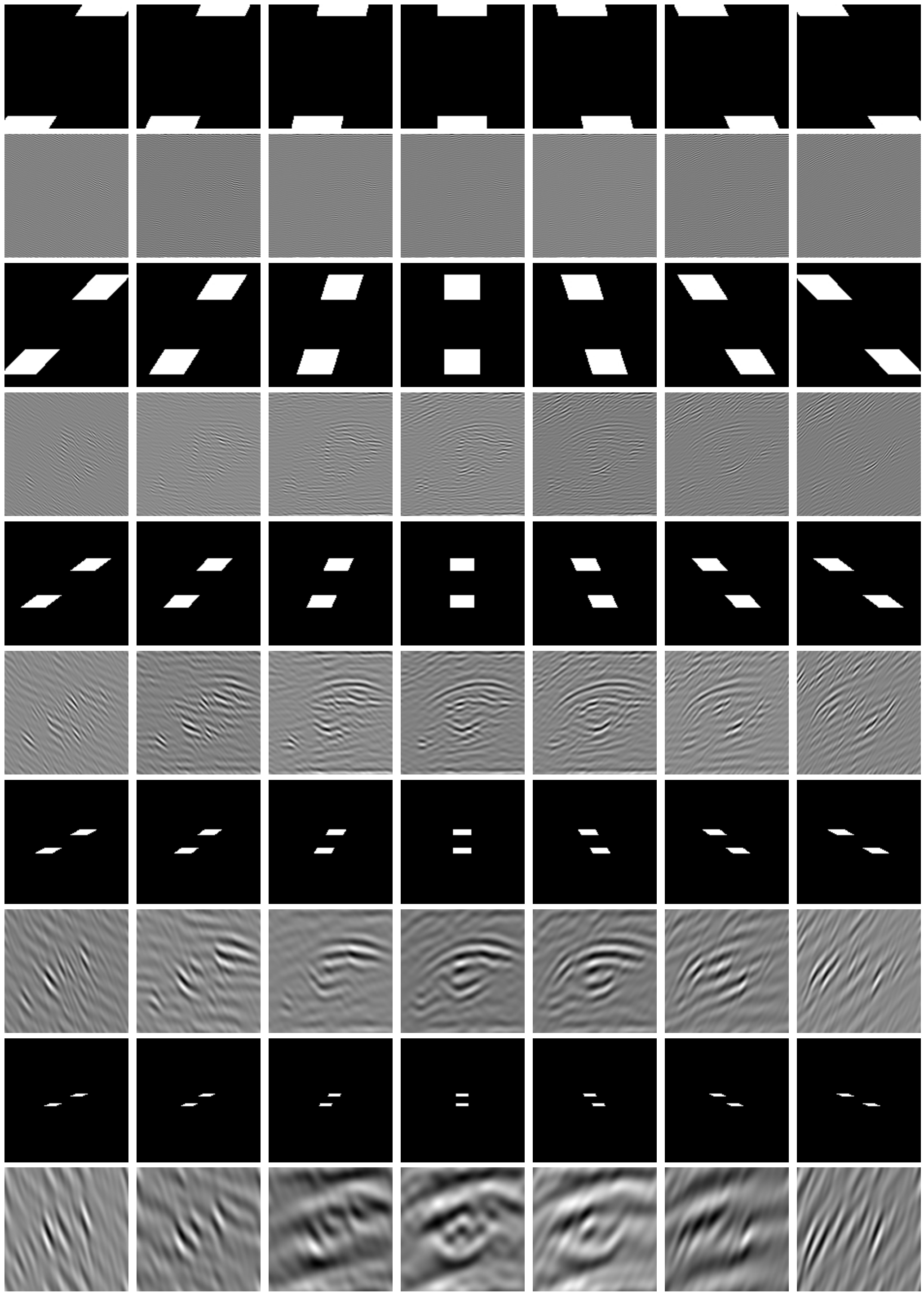

FIGURE 3. Shearlet coefficients for different values of a and s.

[10] H. G. Feichtinger and K. Gröchenig, *A unified approach to atomic decompositions via integrable group representations*, Proc. Conf. "Function Spaces and Applications", Lund 1986, Lecture Notes in Math. **1302** (1988), 52–73.

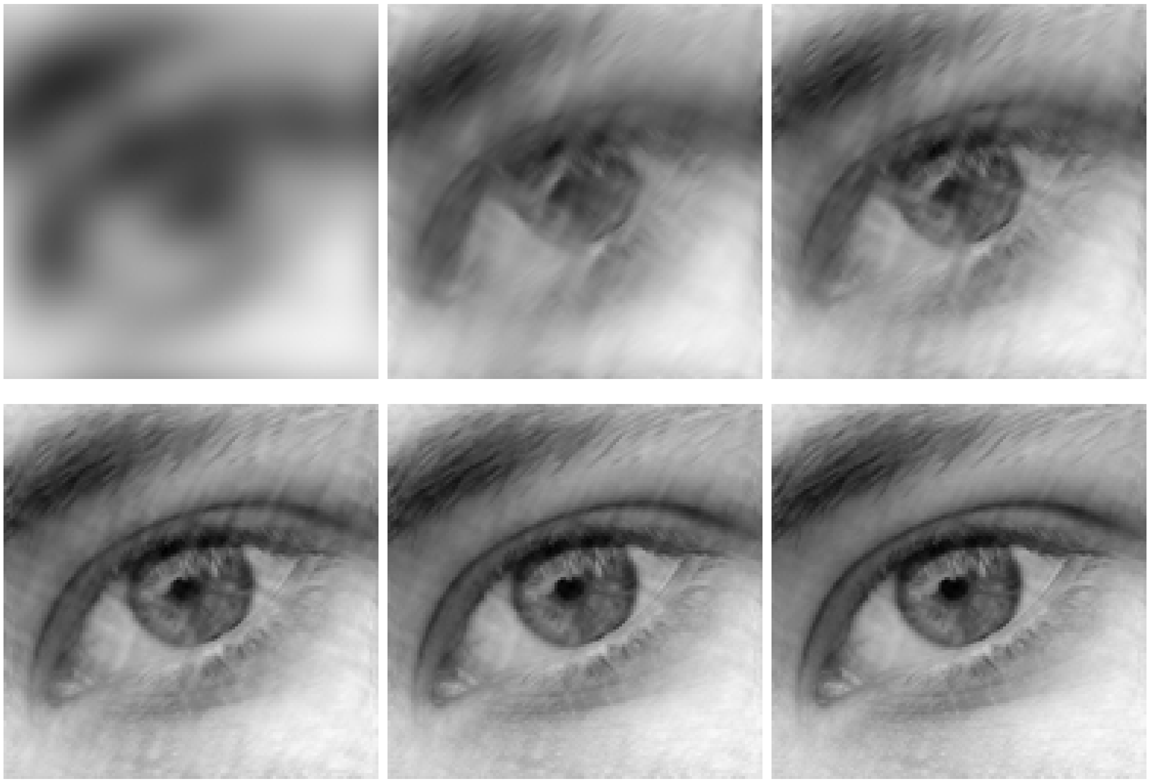

FIGURE 4. Several iterates of iteration (34) for $\rho = 0.01$ and $q = 1$.

[11] H. G. Feichtinger and K. Gröchenig, *Banach spaces related to integrable group representations and their atomic decomposition I,* J. Funct. Anal. **86** (1989), 307–340.

[12] H. G. Feichtinger and K. Gröchenig, *Banach spaces related to integrable group representations and their atomic decomposition II,* Monatsh. Math. **108** (1989), 129–148.

[13] H. G. Feichtinger and K. Gröchenig, *Non–orthogonal wavelet and Gabor expansions and group representations,* in: Wavelets and Their Applications, M.B. Ruskai et.al. (eds.), Jones and Bartlett, Boston, 1992, 353–376.

[14] G. B. Folland, *A Course in Abstract Harmonic Analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, 1995.

[15] K. Gröchenig, *Describing functions: Atomic decompositions versus frames,* Monatsh. Math. **112** (1991), 1–42.

[16] K. Gröchenig, *Foundations of Time–Frequency Analysis*, Birkhäuser, Boston, Basel, Berlin, 2001.

[17] K. Gröchenig and S. Samarah, *Nonlinear approximation with local Fourier bases,* Constr. Approx. **16** (2000) 317-331.

[18] K. Guo, G. Kutyniok, and D. Labate, *Sparse multidimensional representations using anisotropic dilation und shear operators,* in Wavelets und Splines (Athens, GA, 2005), G. Chen und M. J. Lai, eds., Nashboro Press, Nashville, TN (2006), 189–201.

[19] M. Holschneider, *Wavelet analysis of partial differential operators*, in Advances in Partial Differential Equations vol 11: Schrödinger Operators, Markov Semigroups, Operator Algebras, M. Demuth, E. Schrohe, B.W. Schulze and J. Sjöstrand, eds., Akademie-Verlag, 1996, 218-347.

[20] G. Kutyniok and D. Labate, *Resolution of the wavefront set using continuous shearlets,* Preprint (2006).

[21] G. Kutyniok and T. Sauer, *From Wavelets to Shearlets and back again*, Preprint (2007).

Philipps-Universität Marburg, FB12 Mathematik und Informatik, Hans-Meerwein Strasse, Lahnberge, 35032 Marburg, Germany
E-mail address: dahlke@mathematik.uni-marburg.de

Program in Applied and Computational Mathematics, Princeton University, Washington Road, Princeton, NJ 08544-1000, USA
E-mail address: kutyniok@math.princeton.edu

Universität Mannheim, Fakultät für Mathematik und Informatik, Institut für Mathematik, 68131 Mannheim, Germany
E-mail address: steidl@math.uni-mannheim.de

Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Research Group Inverse Problems in Science and Technology, Takustr. 7, 14195 Berlin-Dahlem, Germany
E-mail address: teschke@zib.de