

**DIRECTIONAL MULTISCALE ANALYSIS USING  
SHEARLET  
THEORY AND APPLICATIONS**

---

A Dissertation  
Presented to  
the Faculty of the Department of Mathematics  
University of Houston

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

---

By  
Pooran Singh Negi  
August 2012

**DIRECTIONAL MULTISCALE ANALYSIS USING  
SHEARLET  
THEORY AND APPLICATIONS**

---

Pooran Singh Negi

APPROVED:

---

Dr. Demetrio Labate,  
Chairman

---

Dr. Emanuel Papadakis

---

Dr. Bernhard Bodmann

---

Dr. Robert Azencott

---

Dr. Saurabh Prasad

---

Dean, College of Natural Sciences and Mathematics

# Acknowledgements

I have been very fortunate to pursue my desire for doing research in mathematics. Coming back to academia has been an enjoyable journey. The work in this thesis is in collaboration with my advisor Dr. Demetrio Labate. His unrelenting support and mentoring has allowed me to enjoy doing research and grow in this field. I am deeply indebted to him for accepting me as his Ph.D student.

I am grateful to Dr. Robert Azencott, Dr. Bernhard Bodmann, Dr. Emanuel Papadakis, and Dr. Saurabh Prasad for being part of my thesis committee and providing valuable suggestions.

I am thankful to people at department of mathematics for all the support and guidance, specially to Dr. Shanyu Ji who allowed me to pursue mathematics. I would like to thank all my friends who have been source of laughter, joy and support. I would also like to thank Pallavi Arora for being there every single moment sharing life, research, and her strong belief in me. Finally back home in India, I am deeply indebted to my family specially to my parents Sri. Kripal Singh Negi and Smt. Bhagwati Devi for allowing me to follow my heart.

**DIRECTIONAL MULTISCALE ANALYSIS USING  
SHEARLET  
THEORY AND APPLICATIONS**

---

An Abstract of a Dissertation  
Presented to  
the Faculty of the Department of Mathematics  
University of Houston

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

---

By  
Pooran Singh Negi  
August 2012

# Abstract

Shearlets emerged in recent years in applied harmonic analysis as a general framework to provide sparse representations of multidimensional data. This construction was motivated by the need to provide more efficient algorithms for data analysis and processing, overcoming the limitations of traditional multiscale methods. Particularly, shearlets have proved to be very effective in handling directional features compared to ideas based on separable extension, used in multi-dimensional Fourier and wavelet analysis. In order to efficiently deal with the edges and the other directionally sensitive (anisotropic) information, the analyzing shearlet elements are defined not only at various locations and scales but also at various orientations.

Many important results about the theory and applications of shearlets have been derived during the past 5 years. Yet, there is a need to extend this approach and its applications to higher dimensions, especially 3D, where important problems such as video processing and analysis of biological data in native resolution require the use of 3D representations. The focus of this thesis is the study of shearlet representations in 3D, including their numerical implementation and application to problems of data denoising and enhancement. Compared to other competing methods like 3D curvelet and surfacelet, our numerical experiments show better Peak Signal to Noise Ratio (abbreviated as PSNR) and visual quality.

In addition, to further explore the ability of shearlets to provide an ideal framework for sparse data representations, we have introduced and analyzed a new class of smoothness spaces associated with the shearlet decomposition and their relationship with Besov and curvelet spaces. Smoothness spaces associated to a multi-scale representation system are important for analysis and design of better image processing algorithms.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Shearlet Smoothness Spaces</b>	<b>7</b>
2.1	Notation and definitions . . . . .	8
2.2	Decomposition spaces . . . . .	9
2.2.1	Coverings in Banach spaces . . . . .	9
2.2.2	Decomposition spaces and smoothness spaces . . . . .	11
2.3	The shearlet representation . . . . .	16
2.4	Shearlet-type decomposition . . . . .	27
2.4.1	Shearlet-type covering . . . . .	27
2.4.2	Minimal admissible covering . . . . .	30
2.4.3	Shearlet smoothness spaces . . . . .	32
2.4.4	Embedding results . . . . .	34
2.4.5	Equivalence with curvelet spaces . . . . .	35
<b>3</b>	<b>3D Shearlet Representations</b>	<b>39</b>
3.1	Shearlet in 3D . . . . .	40
<b>4</b>	<b>3D Discrete Shearlet Transform(3D DShT)</b>	<b>46</b>
4.1	3D DShT algorithm . . . . .	48

*CONTENTS*

---

4.2	Implementation issues . . . . .	52
4.3	Correlation with theory . . . . .	53
<b>5</b>	<b>Application and Numerical Experiments</b>	<b>57</b>
5.1	Video denoising . . . . .	57
5.2	Video enhancement . . . . .	62
5.3	Denoising with mixed dictionary . . . . .	66
5.3.1	Iterative shrinkage algorithm . . . . .	69
5.3.2	Experiments . . . . .	70
	<b>Bibliography</b>	<b>76</b>

# List of Figures

2.1	(a) The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by the shearlets. (b) Frequency support $\Sigma_{j,\ell}$ of a shearlet $\psi_{j,\ell,k}$ , for $\xi_1 > 0$ . The other half of the support, for $\xi_1 < 0$ , is symmetrical. . . . .	19
2.2	Equivalence of shearlet and curvelet coverings. . . . .	37
3.1	From left to right, the figure illustrates the pyramidal regions $\mathcal{P}_1$ , $\mathcal{P}_2$ , and $\mathcal{P}_3$ in the frequency space $\widehat{\mathbb{R}}^3$ . . . . .	40
3.2	Frequency support of a representative shearlet function $\psi_{j,\ell,k}$ , inside the pyramidal region $\mathcal{P}_1$ . The orientation of the support region is controlled by $\ell = (\ell_1, \ell_2)$ ; its shape is becoming more elongated as $j$ increases ( $j = 4$ in this plot). . . . .	43
4.1	3D DShT Decomposition of <i>Tempete</i> movie. The figure illustrates some representative 2D frames reconstructed from the 3D DShT decomposition of the movie. All detail frames are extracted from directional subbands contained in the pyramidal region $D_{\mathcal{C}_1}$ . Detail frames, which show highly directional features, are shown in inverted gray scale. . . . .	55
4.2	Analysis of the nonlinear approximation error using the 3D DShT algorithm. (a) Cross section of the piecewise constant radial function $f$ (on $\mathbb{R}^3$ ). (b) Approximation error $\ f - f_M\ _2$ . . . . .	56

5.1	Video Denoising of Mobile Video Sequence. The figure compares the denoising performance of the denoising algorithm based on the 3D DShT, denoted as 3DSHEAR, on a representative frame of the video sequence <i>Mobile</i> against various video denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB, corresponding to $\sigma = 30$ ), denoised frame using 3DSHEAR (PSNR= <b>28.68 dB</b> ), SURF (PSNR=28.39 dB), 2DSHEAR (PSNR=25.97 dB), and DWT (PSNR=24.93 dB). . . . .	60
5.2	Video Denoising of Coast Guard Video Sequence. The figure illustrates the denoising performance on a representative frame of the video sequence using various denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB), denoised frame using 3DSHEAR (PSNR= <b>27.36 dB</b> ), SURF (PSNR=26.82 dB), 2DSHEAR (PSNR=25.20 dB), DWT (PSNR=24.34 dB). . . . .	63
5.3	Video Enhancement. Representative frames from the Barbara video sequence (above) and from the Anterior ultrasound video sequence (below) illustrate the performance of the shearlet-based enhancement algorithm. This is compared against a similar wavelet-based enhancement algorithm. . . . .	65
5.4	Separable Surrogate Functional (SSF) iterative shrinkage algorithm to solve 5.3.3	69
5.5	Video Denoising of Tempete Video Sequence. The figure compares the denoising performance of the denoising algorithm based on the 3D DShT, denoted as Shear, on a representative frame of the video sequence <i>Tempete</i> against various video denoising routines. Starting from the top left: original frame, noisy frame (PSNR=22.14 dB, corresponding to $\sigma = 20$ ), denoised frame using DWT (PSNR= 22.16 dB), DWT/DCT (PSNR=24.09 dB), LP (PSNR=23.10 dB), LP/DCT (PSNR=24.45 dB), Shear (PSNR=25.87 dB), Shear/DCT (PSNR= <b>27.47 dB</b> ), Curvelet (PSNR= 22.60 dB) and Curvelet/DCT (PSNR=25.29 dB). . . . .	73
5.6	Video Denoising of Oil Painting Video Sequence. The figure compares the denoising performance of the denoising algorithm based on the 3D DShT, denoted as Shear, on a representative frame of the video sequence <i>Oil Painting</i> against various video denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB, corresponding to $\sigma = 30$ ), denoised frame using DWT (PSNR= 24.81 dB), DWT/DCT (PSNR=26.03 dB), LP (PSNR=25.52 dB), LP/DCT (PSNR=26.37 dB), Shear (PSNR=27.12 dB), Shear/DCT (PSNR= <b>29.07 dB</b> ), Curvelet (PSNR= 26.86 dB) and Curvelet/DCT (PSNR=25.94 dB). . . . .	75

# List of Tables

5.1	Table I: Video denoising performance using different video sequences. . . .	61
5.2	Table II: Comparison of running times for different 3D transforms. . . .	61
5.3	Table III: Mix Dictionary Denoising results (PSNR) using <i>Tempete video</i> . . . .	70
5.4	Table IV: Mix Dictionary Denoising results (PSNR) using <i>Oil painting video</i> . . .	71
5.5	Table V: Comparison of running times for different routines. . . . .	71

# Chapter 1

## Introduction

Over the past twenty years, wavelets and multiscale methods have been extremely successful in applications from harmonic analysis, approximation theory, numerical analysis, and image processing. However, it is now well established that, despite their remarkable success, wavelets are not very efficient when dealing with multidimensional functions and signals. This limitation is due to their poor directional sensitivity and limited capability in dealing with the anisotropic features which are frequently dominant in multidimensional applications. To overcome this limitation, a variety of methods have been recently introduced to better capture the geometry of multidimensional data, leading to reformulate wavelet theory and applied Fourier analysis within the setting of an emerging theory of *sparse representations*. It is indicative of this change of perspective that the latest edition of the classical wavelet textbook by S. Mallat was titled “A wavelet tour of signal processing. The *sparse* way.”

The shearlet representation, originally introduced in [1, 2], has emerged in recent years as one of the most effective frameworks for the analysis and processing of multidimensional

---

data. This representation is part of a new class of multiscale methods introduced during the last 10 years with the goal to overcome the limitations of wavelets and other traditional methods through a framework which combines the standard multiscale decomposition and the ability to efficiently capture anisotropic features. Other notable such methods include the *curvelets* [3] and the *contourlets* [4]. Similar to the *curvelets* of Donoho and Candès [3], the elements of the shearlet system form a pyramid of well localized waveforms ranging not only across various scales and locations, like wavelets, but also at various orientations and with highly anisotropic shapes. In particular, the directionality of the shearlet systems is controlled through the use of *shearing matrices* rather than rotations, which are employed by curvelets. This offers the advantage of preserving the discrete integer lattice and enables a natural transition from the continuous to the discrete setting. The contourlets, on the other hand, are a purely discrete framework, with the emphasis in the numerical implementation rather than the continuous construction.

Indeed, both curvelets and shearlets have been shown to form Parseval frames of  $L^2(\mathbb{R}^2)$  which are (nearly) optimally sparse in the class of *cartoon-like images*, a standard model for images with edges [3, 5]. Specifically, if  $f_M$  is the  $M$  term approximation obtained by selecting the  $M$  largest coefficients in the shearlet or curvelet expansion of a cartoon-like image  $f$ , then the approximation error satisfies the asymptotic estimate

$$\|f - f_M^S\|_2^2 \asymp M^{-2}(\log M)^3, \quad \text{as } M \rightarrow \infty.$$

Up to the log-like factor, this is the optimal approximation rate, in the sense that no other orthonormal systems or even frames can achieve a rate better than  $M^{-2}$ . By contrast, wavelet approximations can only achieve a rate  $M^{-1}$  for functions in this class [3]. Concerning the topic of sparse approximations, it is important to recall that the relevance of this notion goes far beyond the applications to compression. In fact, constructing sparse

---

representations for data in a certain class entails the intimate understanding of their true nature and structure, so that sparse representations also provide the most effective tool for tasks such as feature extraction and pattern recognition [6, 7].

The special properties of the shearlet approach have been successfully exploited in several imaging application. For example, the combination of multiscale and directional decomposition using shearing transformations is used to design powerful algorithms for image denoising in [7, 8]; the directional selectivity of the shearlet representation is exploited to derive very competitive algorithms for edge detection and analysis in [9]; the sparsity of the shearlet representation is used to derive a very effective algorithm for the regularized inversion of the Radon transform in [10]. We also recall that a recent construction of compactly supported shearlet appears to be especially promising in PDE's and other applications [11, 12].

While directional multiscale systems such as curvelet and shearlet have emerged several years ago, only very recently the analysis of sparse representations using these representations has been extended beyond dimension 2. This extension is of great interest since many applications from areas such as medical diagnosis, video surveillance and seismic imaging require to process 3D data sets, and sparse 3D representations are very useful for the design of improved algorithms for data analysis and processing.

Notice that the formal extension of the construction of multiscale directional systems from 2D to 3D is not the major challenge. In fact, 3D versions of curvelet have been introduced in [13], with the focus being on their discrete implementations. Another discrete method is based on the system of *surfacelets* that were introduced as 3D extensions of contourlets in [14]. However, the analysis of the sparsity properties of curvelet or shearlet (or any other similar systems) in the 3D setting does not follow directly from the 2D argument.

---

Only very recently [15, 16] it was shown that 3D shearlet representations exhibit essentially optimal approximation properties for piecewise smooth functions of 3 variables. Namely, for 3D functions  $f$  which are smooth away from discontinuities along  $C^2$  surfaces, it was shown that the  $M$  term approximation  $f_M^S$  obtained by selecting the  $N$  largest coefficients in the 3D Parseval frame shearlet expansion of  $f$  satisfies the asymptotic estimate

$$\|f - f_M^S\|_2^2 \asymp M^{-1}(\log M)^2, \quad \text{as } M \rightarrow \infty. \quad (1.0.1)$$

Up to the logarithmic factor, this is the *optimal decay rate* and significantly outperforms wavelet approximations, which only yield a  $M^{-1/2}$  rate for functions in this class.

It is useful to recall that optimal approximation properties for a large class of images can also be achieved using adaptive methods by using, for example, the *bandelelets* [17] or the *grouplets* [18]. The shearlet approach, on the other hand, is non-adaptive. Remarkably, shearlet are able to achieve approximation properties which are essentially as good as an adaptive approach when dealing with the class of cartoon-like images.

Many basic questions concerning the study of sparse and efficient representations are closely related to the study of the function spaces associated with these representations. For example, wavelets are ‘naturally’ associated with Besov spaces, and the notion of sparseness in the wavelet expansion is equivalent to an appropriate smoothness measure in Besov spaces [50]. Similarly, the Gabor systems, which are widely used in time-frequency analysis, are naturally associated to the class of modulation spaces [34, 45]. In the case of shearlets, a sequence of papers by Dahlke, Kutyniok, Steidl and Teschke [40, 53, 54] have recently introduced a class of shearlet spaces within the framework of the coorbit space theory of Feichtinger and Gröchenig [43, 44]. This approach exploits the fact that the shearlet transform stems from a square integrable group representation to derive an appropriate notion of shearlet coorbit spaces. In particular, it is shown that all the conditions needed

---

in the general coorbit space theory to obtain atomic decompositions and Banach frames can be satisfied in the new shearlet setting, and that the shearlet coorbit spaces of function on  $\mathbb{R}^2$  are embedded into Besov spaces.

We explore an alternative approach to the construction of smoothness spaces associated with the shearlet representation. Unlike the theory of shearlet coorbit spaces, our approach does not require any group structure and is closely associated with the geometrical properties of the spatial-frequency decomposition of the shearlet construction.

**Outline of the thesis:**

This thesis is organized as follows:

In Chapter 2 we review the basics of decomposition spaces recently introduced by L. Borup and M. Nielsen [35]. This theory is the backbone of the results about Shearlet Smoothness Spaces which are presented in this chapter. Section 2.2 lists all the preliminary result from [35]. Section 2.4 contains results obtained in collaboration with Mantovani about shearlet smoothness spaces. This work is currently part of a submitted paper [65].

In Chapter 3 the theory of 3D Shearlet Representation is presented. 3D shearlet are constructed as a system of waveform that are well localized, bandlimited, orientable and highly elongated at fine scale due to action of anisotropic dilation matrix.

In Chapter 4 the discrete implementation of the 3D system of shearlet is introduced. One main focus is the construction of directional shearing filters associated to this multiscale transform. This work was published in a proceeding paper [66] and a journal paper [67].

Chapter 5 contains the application of 3D Shearlet transform to 3D data denoising and enhancement. Two different methods are considered: data processing using a fixed

---

shearlet dictionary and also processing using combinations of sparse dictionaries including 3D Shearlet and 3D DCT (Discrete Cosine Transform). This work is part of a published journal paper [67] and part of a submitted paper [68].

# Chapter 2

## Shearlet Smoothness Spaces

The introduction of smoothness spaces is motivated by recent results in image processing showing the advantage of using smoothness spaces associated with directional multiscale representations for the design and performance analysis of improved image restoration algorithms. Method presented in current work is derived from the theory of decomposition spaces originally introduced by Feichtinger and Gröbner [41, 42] and recently revisited in the recent work by Borup and Nielsen [35], who have adapted the theory of decomposition spaces to design a very elegant framework for the construction of smoothness spaces closely associated with particular structured decompositions in the Fourier domain. As will be made clear below, this approach can be considered as a refinement of the classical construction of Besov spaces, which are associated with the dyadic decomposition of the Fourier space. Beside its mathematical interest, the construction of the shearlet smoothness spaces presented in this work is also motivated by some recent applications in image restoration where it is shown that the introduction of these smoothness spaces allows one to

take advantage of the optimally sparse approximation properties of directional representations such as shearlets and curvelets when dealing with images and other multidimensional data [37, 51]. In [37] for example, a denoising procedure based on Stein-block thresholding is applied within the class of piecewise  $C^2$  images away from piecewise  $C^2$  singularities, a function space which can be precisely described using curvelet or shearlet smoothness spaces.

The chapter is organized as follows. After recalling the basic definitions and results from the theory of decomposition spaces (Section 2.2) and from the theory of shearlets (Section 2.3), the new shearlet decomposition spaces for functions on  $\mathbb{R}^2$  are introduced in Section 2.4. In particular, we show that there is a Parseval frame forming an atomic decomposition for these spaces and that they are completely characterized by appropriate smoothness conditions on the frame coefficients. We also examine the embeddings of shearlet smoothness spaces into Besov spaces and their relationship with the so-called curvelet spaces.

## 2.1 Notation and definitions

Before proceeding, it is useful to establish some notation and definitions which are used in the following.

Let us adopt the convention that  $x \in \mathbb{R}^d$  is a column vector, i.e.,  $x = \begin{pmatrix} x_1, \dots, x_d \end{pmatrix}^t$ , and that  $\xi \in \widehat{\mathbb{R}}^d$  (in the frequency domain) is a row vector, i.e.,  $\xi = (\xi_1, \dots, \xi_d)$ . A vector  $x$  multiplying a matrix  $a \in GL_d(\mathbb{R})$  on the right is understood to be a column vector, while a vector  $\xi$  multiplying  $a$  on the left is a row vector. Thus,  $ax \in \mathbb{R}^d$  and  $\xi a \in \widehat{\mathbb{R}}^d$ . The

**Fourier transform** of  $f \in L^1(\mathbb{R}^d)$  is defined as

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \xi x} dx,$$

where  $\xi \in \widehat{\mathbb{R}^d}$ , and the inverse Fourier transform is

$$\check{f}(x) = \int_{\widehat{\mathbb{R}^d}} f(\xi) e^{2\pi i \xi x} d\xi.$$

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

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

This is equivalent to the reproducing formula  $f = \sum_i \langle f, \psi_i \rangle \psi_i$ , for all  $f \in \mathcal{H}$ , where the series converges in the norm of  $\mathcal{H}$ . This shows that a Parseval frame provides a basis-like representation even though a Parseval frame need not be a basis in general. The reader can follow [36, 38] for more details about frames.

## 2.2 Decomposition spaces

The following sections contain the main facts from the theory of Decomposition Spaces originally introduced by Feichtinger and Gröbner [41, 42], which will be used to introduce our new definition of Shearlet Smoothness Spaces in Sec. 2.4.

### 2.2.1 Coverings in Banach spaces

A collection  $\{Q_i : i \in I\}$  of measurable and limited sets in  $\mathbb{R}^d$  is an **admissible covering** if  $\cup_{i \in I} Q_i = \mathbb{R}^d$ , and if there is a  $n_0 \in \mathbb{N}$  such that  $\#\{j \in I : Q_i \cap Q_j \neq \emptyset\} \leq n_0$  for all

$i \in I$ . Given an admissible covering  $\{Q_i : i \in I\}$  of  $\mathbb{R}^d$ , a **bounded admissible partition of unity** (BAPU) is a family of functions  $\Gamma = \{\gamma_i : i \in I\}$  satisfying:

- $\text{supp } \gamma_i \subset Q_i \quad \forall i \in I,$
- $\sum_{i \in I} \gamma_i(\xi) = 1, \quad \xi \in \mathbb{R}^d,$
- $\sup_{i \in I} |Q_i|^{1/p-1} \|\mathcal{F}^{-1} \gamma_i\|_{L^p} < \infty, \quad \forall p \in (0, 1].$

Given  $\gamma_i \in \Gamma$ , let us define the multiplier  $\gamma_i(D)f = \mathcal{F}^{-1}(\gamma_i \mathcal{F}f)$ ,  $f \in L^2(\mathbb{R}^d)$ . The conditions in the above definition ensure that  $\gamma_i(D)$  defines a bounded operator for band limited functions in  $L^p(\mathbb{R}^d)$ ,  $0 < p \leq \infty$ , uniformly in  $i \in I$  (cf. Prop.1.5.1 in [55]).

The following definitions will also be needed. Let  $\mathcal{Q} = \{Q_i : i \in I\}$  be an admissible covering. A normed sequence space  $Y$  on  $I$  is called **solid** if  $b = b_i \in Y$  and  $|a_i| \leq |b_i|$  for all  $i \in I$  implies that  $a = a_i \in Y$ ; the same space is called  **$\mathcal{Q}$ -regular** if  $h \in Y$  implies that, for each  $i \in I$ ,  $\tilde{h}(i) = \sum_{j \in \tilde{i}} h(j) \in Y$ , with  $\tilde{i} := \{j \in I : Q_i \cap Q_j \neq \emptyset\}$ ; the space is called **symmetric** if it is invariant under permutations  $\rho : I \rightarrow I$ .

Let  $\mathcal{Q} = \{Q_i : i \in I\}$  be an admissible covering. A strictly positive function  $w$  on  $\mathbb{R}^d$  is called  **$\mathcal{Q}$ -moderate** if there exists  $C > 0$  such that  $w(x) \leq C w(y)$  for all  $x, y \in Q_i$  and all  $i \in I$ . A strictly positive  **$\mathcal{Q}$ -moderate weight** on  $I$  (derived from  $w$ ) is a sequence  $v_i = w(x_i), i \in I$ , with  $x_i \in Q_i$  and  $w$  a  $\mathcal{Q}$ -moderate function.

For a solid (quasi-)Banach sequence space  $Y$  on  $I$ , we define the weighted space  $Y_v$  as

$$Y_v = \{\{d_i\}_{i \in I} : \{d_i v_i\}_{i \in I} \in Y\}. \quad (2.2.1)$$

Given a subset  $J$  of the index set  $I$ , we use the notation  $\tilde{J} := \{i \in I : \exists j \in J \text{ s.t. } Q_i \cap Q_j \neq \emptyset\}$ . We also define inductively  $\tilde{J}^{(k+1)} := \widetilde{\tilde{J}^{(k)}}$ ,  $k \geq 0$ , where we set  $\tilde{J}^{(0)}$  be equal to

*J.* Observe that for a single element  $i \in I$  we have  $\tilde{i} := \{j \in I : Q_i \cap Q_j \neq \emptyset\}$ . Now define

$$\widetilde{Q}_i^{(k)} := \bigcup_{j \in \tilde{i}^{(k)}} Q_j, \quad \text{and} \quad \widetilde{\gamma}_i := \sum_{j \in \tilde{i}} \gamma_j,$$

where  $\{\gamma_i : i \in I\}$  is an associated BAPU.

Finally, a notion of equivalence for coverings is also needed. Let  $\mathcal{Q} = \{Q_i : i \in I\}$  and  $\mathcal{P} = \{P_h : h \in H\}$  be two admissible coverings.  $\mathcal{Q}$  is called **subordinate** to  $\mathcal{P}$  if for every index  $i \in I$  there exists  $j \in J$  such that  $Q_i \subset P_j$ .  $\mathcal{Q}$  is called **almost subordinate** to  $\mathcal{P}$ , and will be denoted by  $\mathcal{Q} \leq \mathcal{P}$ , if there exists  $k \in \mathbb{N}$  such that  $\mathcal{Q}$  is subordinate to  $\{\widetilde{P}_j^{(k)} : j \in J\}$ . If  $\mathcal{Q} \leq \mathcal{P}$  and  $\mathcal{P} \leq \mathcal{Q}$ , we say that  $\mathcal{Q}$  and  $\mathcal{P}$  are two **equivalent coverings**, and we will denote with  $\mathcal{Q} \sim \mathcal{P}$ . As shown in the next section, this notion is related to a notion of equivalence for functions spaces.

### 2.2.2 Decomposition spaces and smoothness spaces

There is a natural way of defining a function space associated with an admissible covering which was originally introduced in [42]. Specifically, let  $\mathcal{Q} = \{Q_i : i \in I\}$  be an admissible covering and  $\Gamma$  a corresponding BAPU. Let  $Y$  be a solid (quasi-) Banach sequence space on  $I$ , for which  $\ell^0(I)$  (the finite sequences on  $I$ ) is dense in  $Y$ . Then for  $p \in (0, \infty]$ , the **decomposition space**  $D(\mathcal{Q}, L^p, Y)$  is defined as the set of elements  $f \in \mathcal{S}'(\mathbb{R}^d)$  such that

$$\|f\|_{D(\mathcal{Q}, L^p, Y)} = \left\| \left\{ \|\gamma_i(D) f\|_{L^p} \right\}_{i \in I} \right\|_Y < \infty.$$

It follows from the definition that, for  $p \in (0, \infty)$ ,  $\mathcal{S}(\mathbb{R}^d)$  is dense in  $D(\mathcal{Q}, L^p, Y)$ . Also, one can show that the definition of decomposition space is independent of the particular BAPU, provided that  $Y$  is  $\mathcal{Q}$ -regular [42]. Following is an important result about the equivalence of decomposition spaces (cf. [35, Theorem 2.11]).

**Theorem 1.** *Let  $\mathcal{P} = \{P_i : i \in I\}$  and  $\mathcal{Q} = \{Q_j : j \in J\}$  be two equivalent admissible coverings, and  $\Gamma = \{\gamma_i : i \in I\}$  and  $\Phi = \{\phi_j : j \in J\}$  be corresponding BAPUs. If  $\{v_i ; i \in I\}$  and  $\{u_j ; j \in J\}$  are weights derived from the same moderate function  $w$ , then*

$$D(\mathcal{Q}, L^p, Y_v) = D(\mathcal{P}, L^p, Y_u)$$

*with equivalent norms.*

In this work, the main interest is in a special class of admissible coverings of the frequency space  $\widehat{\mathbb{R}}^d$  which are generated from the action of affine maps on an open set. This idea was originally developed in [35] where a detailed treatment can be found. This section briefly reviews the aspects of this theory which are useful to derive results in the following sections.

Let  $\mathcal{T} = \{A_k \cdot + c_k\}_{k \in \mathbb{N}}$  be a family of invertible affine transformations on  $\widehat{\mathbb{R}}^d$  and suppose that there are two bounded open set  $P, Q \in \widehat{\mathbb{R}}^d$ , with  $P$  compactly contained in  $Q$  such that the sets  $\{QT : T \in \mathcal{T}\}$  and  $\{PT : T \in \mathcal{T}\}$  are admissible coverings. If, in addition, there is a constant  $K$  such that

$$(QA_k + c_k) \cap (QA_{k'} + c_{k'}) \neq \emptyset \Rightarrow \|A_{k'}^{-1}A_k\|_{\ell^\infty} < K, \quad (2.2.2)$$

then  $\mathcal{Q} = \{TQ : T \in \mathcal{T}\}$  is a **structured admissible covering** and  $\mathcal{T}$  a **structured family of affine transformations**. Following result about structured admissible covering and structured family of affine transformations holds :

**Proposition 2** ([35]). *Let  $\mathcal{Q} = \{QT : T \in \mathcal{T}\}$  be a structured admissible covering and  $\mathcal{T}$  a structured family of affine transformations. Then there exist:*

- (a) a BAPU  $\{\gamma_T : T \in \mathcal{T}\} \subset \mathcal{S}(\widehat{\mathbb{R}}^d)$  corresponding to  $\mathcal{Q}$ ;
- (b) a system  $\{\phi_T : T \in \mathcal{T}\} \subset \mathcal{S}(\widehat{\mathbb{R}}^d)$  satisfying:

- $\text{supp } \phi_T \subset QT, \quad \forall T \in \mathcal{T},$
- $\sum_{T \in \mathcal{T}} \phi_T^2(\xi) = 1, \quad \xi \in \widehat{\mathbb{R}}^d,$
- $\sup_{T \in \mathcal{T}} |T|^{1/p-1} \|\mathcal{F}^{-1} \phi_T\|_{L^p} < \infty, \quad \forall p \in (0, 1].$

A family of function fulfilling the three conditions in point b) of Proposition 2 will be called a **squared BAPU** .

**Remark 2.2.1.** Notice that in the case of structured admissible coverings the characterization of **equivalent coverings** is simplified. In fact, let  $\mathcal{P} = \{PT : T \in \mathcal{T}\}$  and  $\mathcal{Q} = \{QT : T \in \mathcal{T}\}$  be two admissible structured covering with respect to the same family of transformation  $\mathcal{T}$ . Then  $\mathcal{P} \sim \mathcal{Q}$  if  $\#N_P < \infty$  and  $\#N_Q < \infty$ , where  $N_P := \{T \in \mathcal{T} : P \cap QT \neq \emptyset\}$  and  $N_Q := \{T \in \mathcal{T} : Q \cap PT \neq \emptyset\}$ . In fact, that means that  $P \subset \bigcup_{T \in N_P} QT$  and  $Q \subset \bigcup_{T \in N_Q} PT$ , hence  $PS \subset \bigcup_{T \in N_P} QTS$  and  $QS \subset \bigcup_{T \in N_Q} PTS, \forall S \in \mathcal{T}$ .

Let  $\mathcal{Q} = \{QT : T \in \mathcal{T}\}$  be a structured admissible covering and  $\mathcal{T}$  a structured family of affine transformations. Suppose that  $K_a$  is a cube in  $\widehat{\mathbb{R}}^d$  (aligned with the coordinate axes) with side-length  $2a$  satisfying  $Q \subset K_a$ . Corresponding to  $K_a$ , we define the system

$$\{\eta_{n,T} = (\phi_T e_{n,T})^\vee : n \in \mathbb{Z}^d, T \in \mathcal{T}\}, \quad (2.2.3)$$

where

$$e_{n,T}(\xi) = (2a)^{-d/2} |T|^{-1/2} \chi_{K_a}(\xi T^{-1}) e^{i \frac{\pi}{a} n \xi T^{-1}}, \quad n \in \mathbb{Z}^d, T \in \mathcal{T},$$

and  $\phi_T$  is a squared BAPU. The following fact is easy to verify.

**Proposition 3.** *The system  $\{\eta_{n,T} : n \in \mathbb{Z}^d, T \in \mathcal{T}\}$  is a Parseval frame of  $L^2(\mathbb{R}^d)$ .*

When the affine transformations  $\mathcal{T}$  are invertible linear transformations (i.e., all translations factors are  $c_k = 0$ ), then the Parseval frame  $\{\eta_{n,T}\}$  is in fact a collection of Meyer-type wavelets. Furthermore, one can go beyond the construction of Parseval frames of  $L^2(\mathbb{R}^d)$ , and use the frame coefficients  $\{\langle f, \eta_{n,T} \rangle\}$  to characterize the decomposition spaces  $D(\mathcal{Q}, L^p, Y_v)$ . For that, it is useful to introduce the notation:

$$\eta_{n,T}^{(p)} = |T|^{1/2-1/p} \eta_{n,T} \quad (2.2.4)$$

Then the following result from [35] holds.

**Proposition 4.** *Let  $\mathcal{Q} = \{TQ : T \in \mathcal{T}\}$  be a structured admissible covering,  $Y$  a solid (quasi-)Banach sequence space on  $\mathcal{T}$  and  $v$  a  $\mathcal{Q}$ -moderate weight. For  $0 < p \leq \infty$  we have the characterization*

$$\|f\|_{D(\mathcal{Q}, L^p, Y_v)} \approx \left\| \left\{ \left( \sum_{n \in \mathbb{Z}^d} |\langle f, \eta_{n,T}^{(p)} \rangle|^p \right)^{1/p} \right\}_{T \in \mathcal{T}} \right\|_{Y_v}.$$

*Usual modifications apply when  $p = \infty$ .*

Notice that the constants in the above characterization are uniform with respect  $p \in [p_0, \infty]$  for any  $p_0 > 0$ .

As Proposition 4 indicates, there is a coefficient space associated with the decomposition spaces  $D(\mathcal{Q}, L^p, Y_v)$ . Hence, we define the coefficient space  $d(\mathcal{Q}, \ell^p, Y_v)$  as the set of coefficients  $c = \{c_{n,T} : n \in \mathbb{Z}^d, T \in \mathcal{T}\} \subset \mathbb{C}$ , satisfying

$$\|c\|_{d(\mathcal{Q}, \ell^p, Y_v)} = \left\| \left\{ \left( \sum_{n \in \mathbb{Z}^d} |c_{n,T}|^p \right)^{1/p} \right\}_{T \in \mathcal{T}} \right\|_{Y_v}.$$

Using this notation, we can define the operators between these spaces. For  $f \in D(\mathcal{Q}, L^p, Y_v)$  the **coefficient operator** is the operator  $C : D(\mathcal{Q}, L^p, Y_v) \rightarrow d(\mathcal{Q}, \ell^p, Y_v)$  defined by

$$C f = \{\langle f, \eta_{n,T}^{(p)} \rangle\}_{n,T}.$$

For  $\{c_{n,T}\}_{n,T} \in d(\mathcal{Q}, \ell^p, Y_v)$  the **reconstruction operator** is the mapping  $R : d(\mathcal{Q}, \ell^p, Y_v) \rightarrow D(\mathcal{Q}, L^p, Y_v)$  defined by

$$R \{c_{n,T}\}_{n,T} = \sum_{n \in \mathbb{Z}^d} c_{n,T} \eta_{n,T}^{(p)}.$$

Then as per [35, Thm.2]):

**Theorem 5.** *For  $0 < p \leq \infty$ , the coefficient operator and the reconstruction operators are both bounded. This makes  $D(\mathcal{Q}, L^p, Y_v)$  a retract of  $d(\mathcal{Q}, \ell^p, Y_v)$ , that is,  $RC = Id_{D(\mathcal{Q}, L^p, Y_v)}$ . In particular:*

$$\|f\|_{D(\mathcal{Q}, L^p, Y_v)} \approx \inf \left\{ \|\{c_{n,T}\}_{n,T}\|_{d(\mathcal{Q}, \ell^p, Y_v)} : f = \sum_{n,T} c_{n,T} |T|^{\frac{1}{p}-\frac{1}{2}} \eta_{n,T} \right\}. \quad (2.2.5)$$

As a special case of decomposition spaces, let us consider the situation where  $\mathcal{T}$  is a structured family of affine transformations,  $Y_v = (\ell^q)_{v_{w,\beta}}$ ,  $w$  is a  $\mathcal{Q}$ -moderate function,  $\beta \in \mathbb{R}$  and  $v_{w,\beta} = \{(w(b_T))^\beta\}_{A_T \cdot + b_T \in \mathcal{T}}$ . In this case, the space is called a **smoothness space** and one use the notation:

$$S_{p,q}^\beta(\mathcal{T}, w) := D(\mathcal{Q}, L^p, (\ell^q)_{v_{w,\beta}}).$$

Let  $\{\eta_{n,T}\}$  be the Meyer-type Parseval frame associated with Meyer wavelet [69] and  $\mathcal{T}$  as given by (2.2.3). By the notation introduced in (2.2.4)

$$|\langle f, \eta_{n,T}^{(\tau)} \rangle| = |T|^{\frac{1}{p}-\frac{1}{\tau}} |\langle f, \eta_{n,T}^{(p)} \rangle|, \quad 0 < \tau, p \leq \infty.$$

Thus, if there is a  $\delta > 0$  such that  $w(b_T) = w(T) \approx |T|^{1/\delta}$ , for  $T \in \mathcal{T}$ , then

$$\begin{aligned} \|f\|_{S_{p,q}^\beta} &\approx \left( \sum_{T \in \mathcal{T}} |T|^{\frac{\beta q}{\delta}} \left( \sum_{n \in \mathbb{Z}^d} |\langle f, \eta_{n,T}^{(p)} \rangle|^p \right)^{q/p} \right)^{1/p} \\ &\approx \left( \sum_{T \in \mathcal{T}} \left( \sum_{n \in \mathbb{Z}^d} |\langle f, \eta_{n,T}^{(r)} \rangle|^p \right)^{q/p} \right)^{1/p}, \quad \frac{\beta}{\delta} = \frac{1}{p} - \frac{1}{r}. \end{aligned}$$

The spaces  $S_{p,q}^\beta(\mathcal{T}, w)$  provide a natural setting for the analysis of nonlinear approximations. For example, using (2.2.5), one obtains the Jackson-type inequality [70]:

$$\inf_{g \in \Sigma_n} \|f - g\|_{S_{p,p}^\beta} \leq C \|f\|_{S_{\tau,\tau}^\gamma} n^{-(\gamma-\beta)/\delta}, \quad \frac{1}{\tau} - \frac{1}{p} = \frac{\gamma - \beta}{\delta}, \quad (2.2.6)$$

where

$$\Sigma_n = \left\{ g = \sum_{n,T \in \Lambda} c_{n,T} \eta_{n,T} : \#\Lambda \leq n \right\}.$$

Notice that, using  $d$ -dimensional (separable) dyadic wavelets  $\{\eta_{n,j}\}$ , with  $\mathcal{T} = \{2^j I_d : j \in \mathbb{Z}\}$ , where  $I_d$  is the  $d$ -dimensional identity matrix and  $\mathcal{Q}$  is an appropriate structured admissible covering, one obtain that

$$\|f\|_{S_{p,q}^\beta} \approx \left( \sum_{j \in \mathbb{Z}} 2^{jq \frac{d}{2}(\beta/\delta + 1/2 - 1/p)} \left( \sum_{n \in \mathbb{Z}} |\langle f, \eta_{n,j} \rangle|^p \right)^{q/p} \right)^{1/p},$$

which can be identified with the Besov space norm of the Besov space  $B_{p,q}^{\frac{\beta}{\delta}}(\mathbb{R}^d)$ .

## 2.3 The shearlet representation

In this section, we recall the construction of the Parseval frames of shearlets in dimension  $d = 2$ .

This construction, which is a modification of the original approach in [1, 5], produces smooth Parseval frames of shearlets for  $L^2(\mathbb{R}^2)$  as appropriate combinations of shearlet systems defined in cone-shaped regions in the Fourier domain  $\widehat{\mathbb{R}}^2$ . Hence,  $\widehat{\mathbb{R}}^2$  is partitioned into the following cone-shaped regions:

$$\mathcal{P}_1 = \left\{ (\xi_1, \xi_2) \in \mathbb{R}^2 : \left| \frac{\xi_2}{\xi_1} \right| \leq 1 \right\}, \quad \mathcal{P}_2 = \left\{ (\xi_1, \xi_2) \in \mathbb{R}^2 : \left| \frac{\xi_2}{\xi_1} \right| > 1 \right\}.$$

To define the shearlet systems associated with these regions, for  $\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$ , let  $\phi \in C^\infty(\mathbb{R})$  be a function such that  $\hat{\phi}(\xi) \in [0, 1]$ ,  $\hat{\phi} \subset [-1/4, 1/4]$  and  $\hat{\phi}|_{[-1/8, 1/8]} = 1$  and

let also

$$\widehat{\Phi}(\xi) = \widehat{\Phi}(\xi_1, \xi_2) = \widehat{\phi}(\xi_1) \widehat{\phi}(\xi_2) \quad (2.3.7)$$

and

$$W(\xi) = W(\xi_1, \xi_2) = \sqrt{\widehat{\Phi}^2(2^{-1}\xi_1, 2^{-1}\xi_2) - \widehat{\Phi}^2(\xi_1, \xi_2)}.$$

It follows that

$$\widehat{\Phi}^2(\xi_1, \xi_2) + \sum_{j \geq 0} W^2(2^{-j}\xi_1, 2^{-j}\xi_2) = 1 \quad \text{for } (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2. \quad (2.3.8)$$

Notice that each function  $W_j^2 = W^2(2^{-j}\cdot)$  has support in the Cartesian corona

$$C_j = [-2^{j-1}, 2^{j-1}]^2 \setminus [-2^{j-3}, 2^{j-3}]^2$$

and that the functions  $W_j^2$ ,  $j \geq 0$ , produce a smooth tiling of the frequency plane into a Cartesian corona:

$$\sum_{j \geq 0} W^2(2^{-j}\xi) = 1 \quad \text{for } \xi \in \mathbb{R}^2 \setminus [-\frac{1}{4}, \frac{1}{4}]^2 \subset \widehat{\mathbb{R}}^2. \quad (2.3.9)$$

Next, let  $v \in C^\infty(\mathbb{R})$  be chosen so that  $\text{supp } v \subset [-1, 1]$  and

$$|v(u-1)|^2 + |v(u)|^2 + |v(u+1)|^2 = 1 \quad \text{for } |u| \leq 1. \quad (2.3.10)$$

In addition, we will assume that  $v(0) = 1$  and that  $v^{(n)}(0) = 0$  for all  $n \geq 1$ .

Hence, for  $V_{(1)}(\xi_1, \xi_2) = v(\frac{\xi_2}{\xi_1})$  and  $V_{(2)}(\xi_1, \xi_2) = v(\frac{\xi_1}{\xi_2})$ , the *shearlet systems associated with the cone-shaped regions*  $\mathcal{P}_h$ ,  $h = 1, 2$  are defined as the countable collection of functions

$$\{\psi_{j,\ell,k}^{(h)} : j \geq 0, -2^{\lfloor j/2 \rfloor} \leq \ell \leq 2^{\lfloor j/2 \rfloor}, k \in \mathbb{Z}^2\}, \quad (2.3.11)$$

where

$$\widehat{\psi}_{j,\ell,k}^{(h)}(\xi) = |\det A_{(h)}|^{-j/2} W(2^{-j}\xi) V_{(h)}(\xi A_{(h)}^{-j} B_{(h)}^{-\ell}) e^{2\pi i \xi A_{(h)}^{-j} B_{(h)}^{-\ell} k}, \quad (2.3.12)$$

and

$$A_{(1)} = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix}, \quad B_{(1)} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A_{(2)} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 2 \end{pmatrix}, \quad B_{(2)} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \quad (2.3.13)$$

Notice that the dilation matrices  $A_{(1)}, A_{(2)}$  are associated with anisotropic dilations and, more specifically, *parabolic scaling* dilations; by contrast, the *shearing* matrices  $B_{(1)}, B_{(2)}$  are non-expanding and their integer powers control the directional features of the shearlet system. Hence, the systems (2.3.11) form collections of well-localized functions defined at various scales, orientations and locations, controlled by the indices  $j, \ell, k$  respectively. In particular, the functions  $\hat{\psi}_{j,\ell,k}^{(1)}$ , given by (2.3.12), are supported inside the trapezoidal regions

$$\Sigma_{j,\ell} := \{(\xi_1, \xi_2) : \xi_1 \in [-2^{j-1}, -2^{j-3}] \cup [2^{j-3}, 2^{j-1}], |\frac{\xi_2}{\xi_1} - \ell 2^{-j/2}| \leq 2^{-j/2}\} \quad (2.3.14)$$

inside the Fourier plane, with a similar condition holding for the functions  $\hat{\psi}_{j,\ell,k}^{(2)}$ . This is illustrated in Fig. 2.1.

As shown in [19], a smooth Parseval frame for  $L^2(\mathbb{R}^2)$  is obtained by combining the two shearlet systems associated with the cone-based regions  $\mathcal{P}_1$  and  $\mathcal{P}_2$  together with a coarse scale system, which takes care of the low frequency region. To ensure that all elements of this combined shearlet system are  $C_c^\infty$  in the frequency domain, the elements whose supports overlap the boundaries of the cone regions in the frequency domain are appropriately modified. Namely one defines *shearlet system for  $L^2(\mathbb{R}^2)$*  as the collection

$$\{\tilde{\psi}_{-1,k} : k \in \mathbb{Z}^2\} \cup \{\tilde{\psi}_{j,\ell,k,h} : j \geq 0, |\ell| < 2^{[j/2]}, k \in \mathbb{Z}^2, h = 1, 2\} \cup \{\tilde{\psi}_{j,\ell,k} : j \geq 0, \ell = \pm 2^{[j/2]}, k \in \mathbb{Z}^2\}, \quad (2.3.15)$$

consisting of:

- the *coarse-scale shearlets*  $\{\tilde{\psi}_{-1,k} = \Phi(\cdot - k) : k \in \mathbb{Z}^2\}$ , where  $\Phi$  is given by (2.3.7);
- the *interior shearlets*  $\{\tilde{\psi}_{j,\ell,k,h} = \psi_{j,\ell,k}^{(h)} : j \geq 0, |\ell| < 2^{[j/2]}, k \in \mathbb{Z}^2, h = 1, 2\}$ , where the functions  $\psi_{j,\ell,k}^{(h)}$  are given by (2.3.12);

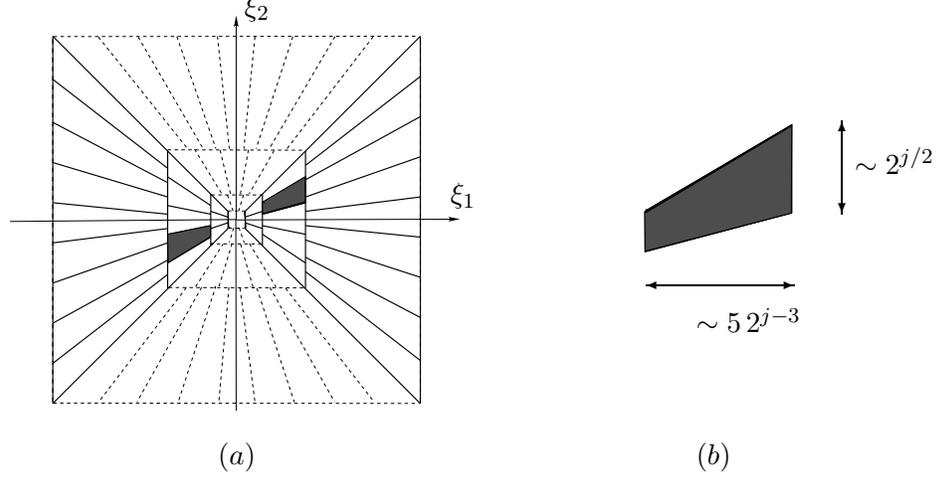


Figure 2.1: (a) The tiling of the frequency plane  $\widehat{\mathbb{R}}^2$  induced by the shearlets. (b) Frequency support  $\Sigma_{j,\ell}$  of a shearlet  $\psi_{j,\ell,k}$ , for  $\xi_1 > 0$ . The other half of the support, for  $\xi_1 < 0$ , is symmetrical.

- the *boundary shearlets*  $\{\tilde{\psi}_{j,\ell,k} : j \geq 0, \ell = \pm 2^{\lfloor j/2 \rfloor}, k \in \mathbb{Z}^2\}$ , obtained by joining together slightly modified versions of  $\psi_{j,\ell,k}^{(1)}$  and  $\psi_{j,\ell,k}^{(2)}$ , for  $\ell = \pm 2^{\lfloor j/2 \rfloor}$ , after that they have been restricted in the Fourier domain to the cones  $\mathcal{P}_1$  and  $\mathcal{P}_2$ , respectively. The precise definition is given below.

For  $j \geq 1$ ,  $\ell = \pm 2^{\lfloor j/2 \rfloor}$ ,  $k \in \mathbb{Z}^2$ , define

$$(\tilde{\psi}_{j,\ell,k})^\wedge(\xi) = \begin{cases} 2^{-\frac{3}{4}j - \frac{1}{2}} W(2^{-j}\xi_1, 2^{-j}\xi_2) v\left(2^{j/2}\frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi_2^{-1} A_{(1)}^{-j} B_{(1)}^{-\ell} k}, & \text{if } \xi \in \mathcal{P}_1 \\ 2^{-\frac{3}{4}j - \frac{1}{2}} W(2^{-j}\xi_1, 2^{-j}\xi_2) v\left(2^{j/2}\frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi_2^{-1} A_{(1)}^{-j} B_{(1)}^{-\ell} k}, & \text{if } \xi \in \mathcal{P}_2. \end{cases}$$

For  $j = 0$ ,  $k \in \mathbb{Z}^2$ ,  $\ell = \pm 1$ , define

$$(\tilde{\psi}_{0,\ell,k})^\wedge(\xi) = \begin{cases} W(\xi_1, \xi_2) v\left(\frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_1 \\ W(\xi_1, \xi_2) v\left(\frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_2. \end{cases}$$

### 2.3. THE SHEARLET REPRESENTATION

---

For brevity, let us denote the system (2.3.15) using the compact notation

$$\{\tilde{\psi}_\mu, \mu \in M\}, \quad (2.3.16)$$

where  $M = M_C \cup M_I \cup M_B$  are the indices associated with *coarse scale shearlets*, *interior shearlets*, and *boundary shearlets*, respectively, given by

- $M_C = \{\mu = (j, k) : j = -1, k \in \mathbb{Z}^2\}$  (coarse scale shearlets)
- $M_I = \{\mu = (j, \ell, k, h) : j \geq 0, |\ell| < 2^{\lfloor j/2 \rfloor}, k \in \mathbb{Z}^2, h = 1, 2\}$  (interior shearlets)
- $M_B = \{\mu = (j, \ell, k) : j \geq 0, \ell = \pm 2^{\lfloor j/2 \rfloor}, k \in \mathbb{Z}^2\}$  (boundary shearlets).

We have the following result from [19]:

**Theorem 6.** *The system of shearlets (2.3.15) is a Parseval frame for  $L^2(\mathbb{R}^2)$ . In addition, the elements of this system are  $C^\infty$  and compactly supported in the Fourier domain.*

*Proof.* We first show that the system of shearlets (2.3.11) is a Parseval frame for  $L^2(\mathcal{P}_1)^\vee$ .

Notice that

$$(\xi_1, \xi_2) A_{(1)}^{-j} B_{-\ell}^{(1)} = (2^{-2j} \xi_1, -\ell 2^{-2j} \xi_1 + 2^{-j} \xi_2).$$

Hence, we can write the elements of the system of shearlets (2.3.11) as

$$\hat{\psi}_{j,\ell,k}^{(1)}(\xi_1, \xi_2) = 2^{-\frac{3}{2}j} W(2^{-2j} \xi_1, 2^{-2j} \xi_2) v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi A_{(1)}^{-j} B_{-\ell}^{(1)} k}.$$

### 2.3. THE SHEARLET REPRESENTATION

Using the change of variable  $\eta = \xi A_{(1)}^{-j} B_{-\ell}^{(1)}$  and the notation  $Q = [-\frac{1}{2}, \frac{1}{2}]^2$ , we have:

$$\begin{aligned}
& \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{k \in \mathbb{Z}^2} |\langle \hat{f}, \hat{\psi}_{j,\ell,k}^{(1)} \rangle|^2 \\
&= \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{k \in \mathbb{Z}^2} \left| \int_{\mathbb{R}^2} 2^{-\frac{3}{2}j} \hat{f}(\xi) W(2^{-2j} \xi_1, 2^{-2j} \xi_2) v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi A_{(1)}^{-j} B_{-\ell}^{(1)} k} d\xi \right|^2 \\
&= \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{k \in \mathbb{Z}^2} \left| \int_Q 2^{\frac{3}{2}j} \hat{f}(\eta B_{\ell}^{(1)} A_{(1)}^j) W(\eta_1, 2^{-j}(\eta_2 + \ell \eta_1)) v\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i \eta k} d\eta \right|^2 \\
&= \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \int_Q 2^{3j} |\hat{f}(\eta B_{\ell}^{(1)} A_{(1)}^j)|^2 |W(\eta_1, 2^{-j}(\eta_2 + \ell \eta_1))|^2 |v\left(\frac{\eta_2}{\eta_1}\right)|^2 d\eta \\
&= \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 |W(2^{-2j} \xi_1, 2^{-2j} \xi_2)|^2 |v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right)|^2 d\xi \\
&= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} |W(2^{-2j} \xi_1, 2^{-2j} \xi_2)|^2 |v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right)|^2 d\xi.
\end{aligned}$$

In the computation above, we have used the fact that the function

$$W(\eta_1, 2^{-j}(\eta_2 + \ell \eta_1)) v\left(\frac{\eta_2}{\eta_1}\right)$$

is supported inside  $Q$  since  $v\left(\frac{\eta_2}{\eta_1}\right)$  is supported inside the cone  $|\frac{\eta_2}{\eta_1}| \leq 1$  and  $W(\eta_1, 2^{-j}(\eta_2 + \ell \eta_1))$  is supported inside the strip  $|\eta_1| \leq \frac{1}{2}$ .

Finally, using the properties of  $W$  and  $v$ , we observe that

$$\begin{aligned}
\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} |W(2^{-2j} \xi_1, 2^{-2j} \xi_2)|^2 |v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right)|^2 &= \sum_{j \geq 0} |W(2^{-2j} \xi_1, 2^{-2j} \xi_2)|^2 \sum_{\ell = -2^j}^{2^j} |v\left(2^j \frac{\xi_2}{\xi_1} - \ell\right)|^2 \\
&= \sum_{j \geq 0} |W(2^{-2j} \xi_1, 2^{-2j} \xi_2)|^2 = 1 \text{ for } (\xi_1, \xi_2) \in \mathcal{P}_1.
\end{aligned}$$

Thus, we conclude that for each  $f \in L^2(\mathcal{P}_1)^\vee$

$$\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, \hat{\psi}_{j,\ell,k}^{(1)}(\xi_1, \xi_2) \rangle|^2 = \|f\|^2.$$

A similar construction yields a Parseval frame for  $L^2(\mathcal{P}_2)^\vee$ . Namely, let

$$\{\psi_{j,\ell,k}^{(2)} : j \geq 0, -2^j \leq 2^j, k \in \mathbb{Z}^2\}, \quad (2.3.17)$$

### 2.3. THE SHEARLET REPRESENTATION

---

where

$$\hat{\psi}_{j,\ell,k}^{(2)}(\xi) = |\det A_{(2)}|^{-j/2} W(2^{-2j}\xi) V(\xi A_{(2)}^{-j} B_{-\ell}^{(2)}) e^{2\pi i \xi A_{(2)}^{-j} B_{-\ell}^{(2)} k},$$

$A_{(2)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$ ,  $B_{\ell}^{(2)} = \begin{pmatrix} 0 & 1 \\ 1 & \ell \end{pmatrix}$ . Noticing that

$$(\xi_1, \xi_2) A_{(2)}^{-j} B_{-\ell}^{(2)} = (2^{-2j}\xi_2, 2^{-j}\xi_1 - \ell 2^{-2j}\xi_2),$$

similar to the case above, we can write the elements of the system of vertical shearlets (2.3.17) as

$$\hat{\psi}_{j,\ell,k}^{(2)}(\xi_1, \xi_2) = 2^{-\frac{3}{2}j} W(2^{-2j}\xi_1, 2^{-2j}\xi_2) v\left(2^j \frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi A_{(2)}^{-j} B_{-\ell}^{(2)} k}.$$

Since

$$\begin{aligned} \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v\left(2^j \frac{\xi_1}{\xi_2} - \ell\right)|^2 &= \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 \sum_{\ell = -2^j}^{2^j} |v\left(2^j \frac{\xi_1}{\xi_2} - \ell\right)|^2 \\ &= \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 = 1 \text{ for } (\xi_1, \xi_2) \in \mathcal{P}_2, \end{aligned}$$

a computation essentially identical to the one above shows that for each  $f \in L^2(\mathcal{P}_2)^\vee$

$$\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, \hat{\psi}_{j,\ell,k}^{(2)}(\xi_1, \xi_2) \rangle|^2 = \|f\|^2.$$

To obtain a Parseval frame of shearlets for  $L^2(\mathbb{R}^2)$ , we will take a combination of horizontal and vertical shearlets, together with a coarse scale system which will account for the low frequency region. Notice that it is possible to build such a system in a way that all element are well localized. This cannot be achieved using the ‘‘standard’’ system of shearlets since, at the boundary of the cone regions, its shearlet elements are continuous but not differentiable.

The shearlet system for  $L^2(\mathbb{R}^2)$  is defined as the union of the *coarse scale shearlets*, the *interior shearlets*, and the *boundary shearlets*, given by

$$\{\tilde{\psi}_{-1,k} : k \in \mathbb{Z}^2\} \cup \{\tilde{\psi}_{j,\ell,k,d} : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^2, d = 1, 2\} \cup \{\tilde{\psi}_{j,\ell,k} : j \geq 0, \ell = \pm 2^j, k \in \mathbb{Z}^2\},$$

### 2.3. THE SHEARLET REPRESENTATION

---

where the coarse scale shearlets are the elements  $\tilde{\psi}_{-1,k} = \Phi(\cdot - k)$ , the interior shearlets are the elements  $\tilde{\psi}_{j,\ell,k,d} = \psi_{j,\ell,k}^{(d)}$  (defined for  $|\ell| < 2^j$ ) and the boundary shearlets  $\tilde{\psi}_{j,\ell,k}$ , where  $\ell = \pm 2^j$ , are defined by joining together  $\psi_{j,\ell,k}^{(1)}$  and  $\psi_{j,\ell,k}^{(2)}$  after that they have been restricted to the cones  $\mathcal{P}_1$  and  $\mathcal{P}_2$ , respectively, in the Fourier domain. That is, for  $j \geq 1$ , we define

$$(\tilde{\psi}_{j,2^j,k})^\wedge(\xi) = \begin{cases} 2^{-\frac{3}{2}j-\frac{1}{2}} W(2^{-2j}\xi_1, 2^{-2j}\xi_2) v\left(2^j\left(\frac{\xi_2}{\xi_1} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{-\ell}^{(1)} k}, & \text{if } \xi \in \mathcal{P}_1 \\ 2^{-\frac{3}{2}j-\frac{1}{2}} W(2^{-2j}\xi_1, 2^{-2j}\xi_2) v\left(2^j\left(\frac{\xi_1}{\xi_2} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{-\ell}^{(1)} k}, & \text{if } \xi \in \mathcal{P}_2, \end{cases}$$

with a similar definition for  $\ell = -2^j$ . For  $j = 0$ ,  $\ell = \pm 1$ , we define

$$(\tilde{\psi}_{0,\ell,k,d})^\wedge(\xi) = \begin{cases} 2^{-\frac{1}{2}} W(\xi_1, \xi_2) v\left(\frac{\xi_2}{\xi_1} - \ell\right) e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_1 \\ 2^{-\frac{1}{2}} W(\xi_1, \xi_2) v\left(\frac{\xi_1}{\xi_2} - \ell\right) e^{2\pi i \xi k}, & \text{if } \xi \in \mathcal{P}_2, \end{cases}$$

It turns out that the boundary elements  $\tilde{\psi}_{j,2^j,k}$  are smooth compactly supported functions in the frequency domain. The support condition follows trivially from the definition. It is also easy to verify that  $(\tilde{\psi}_{j,2^j,k})^\wedge(\xi_1, \xi_2)$  is continuous since the two parts of the piecewise defined function are equal for  $\xi_1 = \xi_2$ . To verify the smoothness, notice that

$$\begin{aligned} \frac{\partial}{\partial \xi_1} W(2^{-2j}\xi) v\left(2^j\left(\frac{\xi_2}{\xi_1} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{-\ell}^{(1)} k} \Big|_{\xi_2=\xi_1} &= 2^{-2j} W'(2^{-2j}\xi_1, 2^{-2j}\xi_1) v(0) e^{2\pi i 2^{-2j-1}\xi_1 k_1} \\ &+ 2^j \xi_1 W(2^{-2j}\xi_1, 2^{-2j}\xi_1) v'(0) e^{2\pi i 2^{-2j-1}\xi_1 k_1} \\ &+ 2\pi i (2^{-2j-1}k_1 - 2^{-j-1}k_2) W(2^{-2j}\xi_1, 2^{-2j}\xi_1) \times \\ &v(0) e^{2\pi i 2^{-2j-1}\xi_1 k_1}, \\ \frac{\partial}{\partial \xi_1} W(2^{-2j}\xi) v\left(2^j\left(\frac{\xi_1}{\xi_2} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{-\ell}^{(1)} k} \Big|_{\xi_2=\xi_1} &= 2^{-2j} W'(2^{-2j}\xi_1, 2^{-2j}\xi_1) v(0) e^{2\pi i 2^{-2j-1}\xi_1 k_1} \\ &+ \frac{2^j}{\xi_1} 2^{-\frac{3}{2}j} W(2^{-2j}\xi_1, 2^{-2j}\xi_1) v'(0) e^{2\pi i 2^{-2j-1}\xi_1 k_1} \\ &+ 2\pi i (2^{-2j-1}k_1 - 2^{-j-1}k_2) W(2^{-2j}\xi_1, 2^{-2j}\xi_1) \times \\ &v(0) e^{2\pi i 2^{-2j-1}\xi_1 k_1}. \end{aligned}$$

Since  $v'(0) = 0$ , the two partial derivatives agree for  $\xi_1 = \xi_2$ . A very similar calculation shows that also the partial derivatives with respect to  $\xi_2$  agree for  $\xi_1 = \xi_2$ . It follows from

### 2.3. THE SHEARLET REPRESENTATION

---

these calculations that the boundary elements  $\tilde{\psi}_{j,2^j,k}$  are smooth functions in the frequency domain.

We can now show that we have a Parseval frame for  $L^2(\mathbb{R}^2)$ .

First, we will examine the tiling properties of the boundary elements. We have:

$$\begin{aligned}
& \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} |\langle \hat{f}, (\tilde{\psi}_{j,2^j,k})^\wedge \rangle|^2 \\
&= \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} \left| \int_{\mathcal{P}_1} 2^{-\frac{3}{2}j - \frac{1}{2}} \hat{f}(\xi) W(2^{-2j} \xi_1, 2^{-2j} \xi_2) v\left(2^j \left(\frac{\xi_2}{\xi_1} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(-2^j)}^{(1)} k} d\xi \right|^2 \\
&+ \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} \left| \int_{\mathcal{P}_2} 2^{-\frac{3}{2}j - \frac{1}{2}} \hat{f}(\xi) W(2^{-2j} \xi_1, 2^{-2j} \xi_2) v\left(2^j \left(\frac{\xi_1}{\xi_2} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(-2^j)}^{(1)} k} d\xi \right|^2
\end{aligned} \tag{2.3.18}$$

We will use the change of variable  $\eta = \xi 2^{-1} A_{(1)}^{-j} B_{-\ell}^{(1)}$ . Notice that

$$v\left(2^j \frac{\xi_1}{\xi_2} - \ell\right) = V(\xi A_{(2)}^{-j} B_{-\ell}^{(2)}) = V(\eta 2 B_\ell^{(1)} A_{(1)}^j A_{(2)}^{-j} B_{-\ell}^{(2)}).$$

Since  $B_\ell^{(1)} A_{(1)}^j A_{(2)}^{-j} B_{-\ell}^{(2)} = \begin{pmatrix} \ell 2^{-j} & 2^j - \ell 2^{-j} \\ 2^{-j} & -\ell 2^{-j} \end{pmatrix}$ , then

$$V(\eta 2 B_{(2^j)}^{(1)} A_{(1)}^j A_{(2)}^{-j} B_{(-2^j)}^{(2)}) = v\left(\frac{-\eta_2}{\eta_1 + 2^{-j} \eta_2}\right) \tag{2.3.19}$$

The support condition of  $v$  implies that

$$\left| \frac{\eta_2}{\eta_1 + 2^{-j} \eta_2} \right| \leq 1$$

This implies that

$$\left| \frac{\eta_2}{\eta_1} \right| \leq \left| 1 + 2^{-j} \frac{\eta_2}{\eta_1} \right| \leq 1 + 2^{-j} \left| \frac{\eta_2}{\eta_1} \right|$$

and, thus,

$$(1 - 2^{-j}) \left| \frac{\eta_2}{\eta_1} \right| \leq 1$$

$$\left| \frac{\eta_2}{\eta_1} \right| \leq (1 - 2^{-j})^{-1} \leq 2 \quad \text{for } j \geq 1.$$

### 2.3. THE SHEARLET REPRESENTATION

---

This shows that, if  $|\eta_1| \leq \frac{1}{4}$ , then  $|\eta_2| \leq 2|\eta_1| \leq \frac{1}{2}$  and, thus, the function (2.3.19) is supported inside  $Q$ . Using these observations, we have:

$$\begin{aligned}
& \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} |\langle \hat{f}, (\tilde{\psi}_{j,2^j,k})^\wedge \rangle|^2 \\
&= \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} \left| \int_{\mathcal{P}_1} 2^{-\frac{3}{2}j - \frac{1}{2}} \hat{f}(\xi) W(2^{-2j}\xi_1, 2^{-2j}\xi_2) v\left(2^j\left(\frac{\xi_2}{\xi_1} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(-2^j)}^{(1)} k} d\xi \right|^2 \\
&+ \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} \left| \int_{\mathcal{P}_2} 2^{-\frac{3}{2}j - \frac{1}{2}} \hat{f}(\xi) W(2^{-2j}\xi_1, 2^{-2j}\xi_2) v\left(2^j\left(\frac{\xi_1}{\xi_2} - 1\right)\right) e^{2\pi i \xi 2^{-1} A_{(1)}^{-j} B_{(-2^j)}^{(1)} k} d\xi \right|^2 \\
&= \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} \left| \int_Q 2^{\frac{3}{2}j + \frac{1}{2}} \hat{f}(2\eta B_{2^j}^{(1)} A_{(1)}^j) W(2\eta_1, 2^{-j+1}(\eta_2 + 2^j\eta_1)) v\left(\frac{\eta_2}{\eta_1}\right) e^{2\pi i \eta k} d\eta \right|^2 \\
&+ \sum_{j \geq 0} \sum_{k \in \mathbb{Z}^2} \left| \int_Q 2^{\frac{3}{2}j + \frac{1}{2}} \hat{f}(2\eta B_{2^j}^{(1)} A_{(1)}^j) W(2\eta_1, 2^{-j+1}(\eta_2 + 2^j\eta_1)) v\left(\frac{-\eta_2}{\eta_1 + 2^{-j}\eta_2}\right) e^{2\pi i \eta k} d\eta \right|^2 \\
&= \sum_{j \geq 0} \int_Q 2^{3j+1} |\hat{f}(2\eta B_{2^j}^{(1)} A_{(1)}^j)|^2 |W(2\eta_1, 2^{-j+1}(\eta_2 + 2^j\eta_1))|^2 |v\left(\frac{\eta_2}{\eta_1}\right)|^2 d\eta \\
&+ \sum_{j \geq 0} \int_Q 2^{3j+1} |\hat{f}(2\eta B_{2^j}^{(1)} A_{(1)}^j)|^2 |W(2\eta_1, 2^{-j+1}(\eta_2 + 2^j\eta_1))|^2 |v\left(\frac{-\eta_2}{\eta_1 + 2^{-j}\eta_2}\right)|^2 d\eta \\
&= \sum_{j \geq 0} \int_{\mathcal{P}_1} |\hat{f}(\xi)|^2 |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v\left(2^j\left(\frac{\xi_2}{\xi_1} - 1\right)\right)|^2 d\xi \\
&+ \sum_{j \geq 0} \int_{\mathcal{P}_2} |\hat{f}(\xi)|^2 |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v\left(2^j\left(\frac{\xi_1}{\xi_2} - 1\right)\right)|^2 d\xi.
\end{aligned}$$

An analogous result holds for the boundary elements along the line  $\xi_1 = -\xi_2$ , corresponding to  $\ell = -2^j$ .

Combining all the results above, we have:

$$\begin{aligned}
 & \sum_{d=1}^2 \sum_{j \geq 0} \sum_{|\ell| < 2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, \tilde{\psi}_{j,\ell,k,d} \rangle|^2 + \sum_{j \geq 0} \sum_{\ell = \pm 2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, \tilde{\psi}_{j,\ell,k} \rangle|^2 \\
 &= \sum_{j \geq 0} \sum_{|\ell| < 2^j} \sum_{k \in \mathbb{Z}^2} |\langle \hat{f}, \hat{\psi}_{j,\ell,k}^{(1)} \rangle|^2 + \sum_{j \geq 0} \sum_{|\ell| < 2^j} \sum_{k \in \mathbb{Z}^2} |\langle \hat{f}, \hat{\psi}_{j,\ell,k}^{(2)} \rangle|^2 + \sum_{j \geq 0} \sum_{\ell = \pm 2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, (\tilde{\psi}_{j,\ell,k})^\wedge \rangle|^2 \\
 &= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 \left( \sum_{|\ell| < 2^j} |v(2^j \frac{\xi_2}{\xi_1} - \ell)|^2 + \sum_{|\ell| < 2^j} |v(2^j \frac{\xi_1}{\xi_2} - \ell)|^2 \right) d\xi \\
 &+ \int_{\mathcal{P}_1} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v(2^j(\frac{\xi_1}{\xi_2} - 1))|^2 d\xi \\
 &+ \int_{\mathcal{P}_2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v(2^j(\frac{\xi_1}{\xi_2} - 1))|^2 d\xi \\
 &+ \int_{\mathcal{P}_1} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v(2^j(\frac{\xi_1}{\xi_2} + 1))|^2 d\xi \\
 &+ \int_{\mathcal{P}_2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi_1, 2^{-2j}\xi_2)|^2 |v(2^j(\frac{\xi_1}{\xi_2} + 1))|^2 d\xi \\
 &= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi)|^2 \left( \sum_{|\ell| \leq 2^j} |v(2^j \frac{\xi_2}{\xi_1} - \ell)|^2 \chi_{\mathcal{P}_1}(\xi) + \sum_{|\ell| \leq 2^j} |v(2^j \frac{\xi_1}{\xi_2} - \ell)|^2 \chi_{\mathcal{P}_2}(\xi) \right) d\xi \\
 &= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi)|^2 d\xi
 \end{aligned}$$

Finally:

$$\begin{aligned}
 & \sum_{k \in \mathbb{Z}^2} |\langle f, \tilde{\psi}_{-1,k} \rangle|^2 + \sum_{d=1}^2 \sum_{j \geq 0} \sum_{|\ell| < 2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, \tilde{\psi}_{j,\ell,k,d} \rangle|^2 + \sum_{j \geq 0} \sum_{\ell = \pm 2^j} \sum_{k \in \mathbb{Z}^2} |\langle f, \tilde{\psi}_{j,\ell,k} \rangle|^2 \\
 &= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 |\Phi(\xi)|^2 d\xi + \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 \sum_{j \geq 0} |W(2^{-2j}\xi)|^2 d\xi \\
 &= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 \left( |\Phi(\xi)|^2 + \sum_{j \geq 0} |W(2^{-2j}\xi)|^2 \right) d\xi = \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 d\xi. \tag{2.3.20}
 \end{aligned}$$

□

Due to the symmetry of the construction, in the following it will be usually sufficient

to specialize our to the shearlets in  $\mathcal{P}_1$ . In that case, we will indicate the matrices  $A_{(1)}$  and  $B_{(1)}$  with  $A$  and  $B$ , and the shearlets in  $\mathcal{P}_1$  simply with  $\psi_{j,\ell,k}$ .

## 2.4 Shearlet-type decomposition

In this section, we define a class of smoothness spaces associated with the shearlet-type decomposition of the frequency plane  $\widehat{\mathbb{R}}^2$  presented in Section 2.3.

### 2.4.1 Shearlet-type covering

We start by constructing a structured admissible covering of  $\widehat{\mathbb{R}}^2$  associated with the structured family of affine transformations generating the shearlet systems of Section 2.3.

For  $A = A_1$  and  $B = B_1$  given by (2.3.13), consider the family of affine transformations  $\{T_{(j,\ell)} : (j,\ell) \in \mathcal{M}\}$  on  $\widehat{\mathbb{R}}^2$  by

$$\xi T_{(j,\ell)} = \xi B^\ell A^j, \quad (j,\ell) \in \mathcal{M}, \quad (2.4.21)$$

where  $\mathcal{M} = \{(j,\ell) : j \geq 0, -2^{\lfloor j/2 \rfloor} \leq \ell \leq 2^{\lfloor j/2 \rfloor}\}$ . Next, we choose two bounded sets  $P$  and  $Q$  in  $\widehat{\mathbb{R}}^2$  defined by  $V \cup V^-$  and  $U \cup U^-$ , respectively, where  $V$  is the trapezoid with vertices  $(1/4, 1/4)$ ,  $(1/2, 1/2)$ ,  $(1/2, -1/2)$ ,  $(1/4, -1/4)$ ,  $V^- = \{\xi \in \mathbb{R}^2 : -\xi \in V\}$ ,  $U$  is the trapezoid with vertices  $(1/8, 3/8)$ ,  $(5/8, 5/8)$ ,  $(1/8, -3/8)$ ,  $(5/8, -5/8)$  and  $U^- = \{\xi \in \mathbb{R}^2 : -\xi \in U\}$ . Also, let  $U_0$  be the cube  $[-1/2, 1/2]^2$ ,  $T_0$  to be the affine transformation such that  $U_0 = T_0(V \cup V^-)$  and  $R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Hence, let us consider the structured family of affine transformations:

$$\mathcal{T}_{\mathcal{M}} = \{T_0, T_{(j,\ell)}, RT_{(j,\ell)} : (j,\ell) \in \mathcal{M}\}. \quad (2.4.22)$$

We have the following:

**Proposition 7.** *The set  $\mathcal{Q} = \{QT : T \in \mathcal{T}_{\mathcal{M}}\}$ , where  $\mathcal{T}_{\mathcal{M}}$  is given by (2.4.22), is a structured admissible covering of  $\widehat{\mathbb{R}}^2$ .*

**Proof.** It is easy to verify that:

$$\widehat{\mathbb{R}}^2 = U_0 \cup \left( \bigcup_{(j,\ell) \in \mathcal{M}} PT_{(j,\ell)} \right) \cup \left( \bigcup_{(j,\ell) \in \mathcal{M}} PT_{(j,\ell)} R \right).$$

In fact, the right hand side of the above expression describes the shearlet tiling of the frequency plane illustrated in Fig. 2.1(a). Obviously, the family  $\{U_0, QT_{(j,\ell)}, QT_{(j,\ell)}R : (j,\ell) \in \mathcal{M}\}$  is also a covering of  $\widehat{\mathbb{R}}^2$ . To conclude that  $\{QT : T \in \mathcal{T}_{\mathcal{M}}\}$  is a structured admissible covering of  $\widehat{\mathbb{R}}^2$  it remains to show that the condition (2.2.2) is satisfied. Due to the symmetry of the construction, it is sufficient to specialize our argument to the cone-shaped region  $\mathcal{P}_1$ .

Let us examine the action of a linear mapping  $T_{j,\ell} \in \mathcal{T}_{\mathcal{M}}$  on the trapezoid  $U$ . We have that

$$U_{(j,\ell)} := UT_{(j,\ell)} = U \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix} = U \begin{pmatrix} 2^j & \ell 2^{j/2} \\ 0 & 2^{j/2} \end{pmatrix}.$$

Hence  $T_{(j,\ell)}$  maps the trapezoid  $U$  into another trapezoids where  $(\xi_1, \xi_2) \mapsto (2^j \xi_1, 2^{j/2}(\ell \xi_1 + \xi_2))$ .

Since the first coordinate  $\xi_1 \in U$  ranges over  $[1/8, 5/8]$ , under the action of  $T_{(j,\ell)}$  this interval is mapped into  $[2^{j-3}, 5 \cdot 2^{j-3}]$ . Hence, since  $2^{j-3} < 5 \cdot 2^{(j-1)-3} < 5 \cdot 2^{j-3}$ , each trapezoid  $U_{(j,\ell)}$  can intersect horizontally only the similar sets corresponding to the scales indices  $j-1$  and  $j+1$ .

To estimate the number of vertical intersections, observe that, for  $j$  fixed, all trapezoids  $U_{(j,\ell)}$  have the same vertical extension, independently of the parameter  $\ell$ . Specifically, for  $j$

fixed, the trapezoids  $U_{(j,\ell)}$  (which are defined for  $\xi_1 \in [2^{j-3}, 52^{j-3}]$ ) has vertically extension equal to  $2^{\frac{3}{8}} 2^{j/2} = 3 \cdot 2^{j/2-2}$  on the left side and  $2^{\frac{5}{8}} 2^{j/2} = 5 \cdot 2^{j/2-2}$  right side. The shearing matrix  $B$  produce a vertical displacement by  $2^{j/2} \xi_1$ . This implies that the left side of a trapezoid  $U_{(j,\ell)}$  is displaced by  $2^{j/2-3}$  and the right side by  $5 \cdot 2^{j/2-3}$ . It follows that, inside the strip  $[2^{j-3}, 5 \cdot 2^{j-3}]$ , the maximum number of trapezoids intersecting the trapezoid  $U_{(j,\ell)}$  is less than or equal to the maximum between  $2 \left( \left\lceil \frac{3 \cdot 2^{j/2-2}}{2^{j/2-3}} \right\rceil + 1 \right)$  and  $2 \left( \left\lceil \frac{5 \cdot 2^{j/2-2}}{5 \cdot 2^{j/2-3}} \right\rceil + 1 \right)$ ; that is,  $2(\lceil 3 \cdot 2 \rceil + 1) = 14$ .

We can now estimate the number of trapezoids  $U_{(j',\ell')}$  intersecting the trapezoid  $U_{(j,\ell)}$ , where  $j$  and  $\ell$  are fixed. As we observed, it must be  $j' - 1 \leq j' \leq j + 1$ . Also, as we already know, the right side of the trapezoid  $U_{(j,\ell)}$ , for  $\xi_1 = 5 \cdot 2^{j-3} = 2^{j+1} \cdot 5 \cdot 2^{-4}$ , extends vertically by  $5 \cdot 2^{j/2-2}$ . For the same value of  $\xi_1$ , a generic trapezoid of the form  $U_{(j+1,\ell')}$  will be displaced vertically by  $2^{(j+1)/2} \cdot 5 \cdot 2^{-4}$  under the action of the shearing matrix  $B$ . It follows that the number of trapezoids  $U_{(j+1,\ell')}$  intersecting  $U_{(j,\ell)}$  is less than or equal to  $2 \left( \left\lceil \frac{5 \cdot 2^{j/2-2}}{5 \cdot 2^{(j+1)/2-4}} \right\rceil + 1 \right) = 2 \lceil 2^{3/2} \rceil + 2 = 6$ . On the other side, a trapezoid of the form  $U_{(j-1,\ell')}$  at  $\xi_1 = 2^{j-3} = 2^{j-1} \cdot 2^{-2}$  is displaced by  $2^{j/2-5/2}$  under the action of a shearing matrix  $B$ . Thus, the number of trapezoids  $U_{(j-1,\ell')}$  intersecting  $U_{(j,\ell)}$  is less than or equal  $2 \left( \left\lceil \frac{3 \cdot 2^{j/2-2}}{2^{j/2-5/2}} \right\rceil + 1 \right) = 2 \lceil 3 \cdot 2^{1/2} \rceil + 2 = 10$ .

Combining the above observations, it follows that the total number of trapezoids intersecting the trapezoid  $U_{(j,\ell)}$  is less than or equal to  $14 + 6 + 10 = 30$ .  $\square$

In the following, we will refer to the structured admissible covering of Proposition 7 as the **shearlet-type covering**. By Proposition 2, there is at least a BAPU associated with this admissible covering. In Section 2.4.3 we will analyze the relation between the (Fourier transform of the) Parseval frame of shearlets introduced in Section 3 and one such BAPU.

### 2.4.2 Minimal admissible covering

Given  $\mathcal{T} := \{T_i\}$  a set of invertible affine transformations, we say that  $\{T_i Q\}$  is a **minimal admissible covering** if there is no  $Q'$  s.t.  $Q'$  compactly contained in  $Q$  and  $\{T_i Q'\}$  is an admissible covering.

Let us consider the trapezoid

$$P' = \{(\xi_1, \xi_2) : |\xi_2| \leq \xi_1/2, \xi_1 \in [1/4, 1/2]\}$$

and

$$P'' := P' \cup P'^-.$$

**Theorem 8.** *The tiling  $\{P''T : T \in \mathcal{T}_M\}$ , where  $\mathcal{T}_M$  is given by (2.4.22), satisfies the property of minimality. In addition, it is equivalent to any possible structured admissible covering with respect to the family  $\mathcal{T}_M$ .*

**Proof.** As above, it is sufficient to specialize our argument to the cone-shaped region  $\mathcal{P}_1$ .

Let  $C$  be a generic set in the cone  $\mathcal{P}_1$ . Recall that our family of transformation acts on a generic element of  $\mathbb{R}^2$  in the following way:

$$(\xi_1, \xi_2)B^\ell A^j = (2^j \xi_1, 2^{j/2}(\ell \xi_1 + \xi_2)).$$

It follows that the first coordinate is moved only by a dyadic dilation of  $2^j$ . Since  $j \geq 0$ , we are bound to consider the first coordinate varying in  $[1/4, 1/2]$ . That means that  $C$  is contained in the strip  $\{(\xi_1, \xi_2) : \xi_1 \in [1/4, 1/2], |\xi_2| \leq \xi_1\}$ .

We can also observe that the actions of the shearing matrix  $B^\ell$  consist in lifting up by  $\ell \xi_1$  the second coordinate of each point  $(\xi_1, \xi_2)$ . Since the set  $\{P' B^\ell A^j : |\ell| \leq 2^{j/2}\}$ , with  $j$

fixed, must cover the whole strip  $\{(\xi_1, \xi_2) : \xi_1 \in [2^{j-2}, 2^{j-1}], |\xi_2| \leq \xi_1\}$ , for each fixed value  $\xi_{1,0}$  of the first coordinate the following equation must be satisfied:  $2^{j/2}\xi_{1,0} + 2^{j/2}\xi_{2,b} \leq 2^{j/2}\xi_{2,a}$ , where  $\xi_{2,a}$  and  $\xi_{2,b}$  are respectively the lower and the highest values of the second coordinate in the set  $C$  corresponding to the value  $\xi_1 = \xi_{1,0}$ . That means that  $\xi_{2,b}$  and  $\xi_{2,a}$  must satisfy  $|\xi_{2,b} - \xi_{2,a}| \geq \xi_{1,0}$ . Hence, the minimal choice we can get is  $|\xi_{2,b} - \xi_{2,a}| = \xi_{1,0}$ . This condition is clearly satisfied by taking  $\xi_{2,b} = \xi_{1,0}/2$  and  $\xi_{2,a} = -\xi_{1,0}/2$ .

Next, we can show that any structured admissible covering with respect to the family of transformation  $\mathcal{T}_M$  is equivalent to the covering  $\mathcal{P}' := \{P'T : T \in \mathcal{T}_M\}$ . That means that, if  $\mathcal{Q}' := \{Q'T : T \in \mathcal{T}_M\}$  is an admissible structured covering for  $\hat{\mathbb{R}}^2$ , with  $Q'$  compact in  $\hat{\mathbb{R}}^2$ , then  $\mathcal{Q}'$  and  $\mathcal{P}'$  are equivalent. Thanks to Theorem 1 and our observations above, it is sufficient to prove that  $\#\{T \in \mathcal{T}_M : P'T \cap Q' \neq \emptyset\} < \infty$  and  $\#\{T \in \mathcal{T}_M : Q'T \cap P' \neq \emptyset\} < \infty$ .

Again, we only need to examine the cone-shaped region  $\mathcal{P}_1$ . Let  $\mathcal{Q}' := \{Q'T : T \in \mathcal{T}_M\}$  be an admissible structured covering, with  $Q'$  a generic compact subset of  $\mathcal{P}_1$ . As observed above, for  $\mathcal{Q}'$  to be an admissible structured covering,  $Q'$  must cover at least the strip domain  $\{(\xi_1, \xi_2) : \xi_1 \in [1/4, 1/2], |\xi_2| \leq \xi_1\}$ . Furthermore, we notice that  $Q'$  cannot contain the origin. Indeed, if  $0 \in Q'$ , then  $\forall j \geq 0$  we have that  $0 \in Q'A^j$ , hence  $\#\{j \in \mathbb{N} : Q' \cap Q'A^j \neq \emptyset\} = \infty$ , which makes  $\mathcal{Q}'$  non admissible. Thus, we have that  $Q'$  is contained or equal (in the worst case) to  $\{(\xi_1, \xi_2) : \xi_1 \in [m, M], |\xi_2| \leq \xi_1\}$ , with  $0 < m \leq 1/4 < 1/2 \leq M$ . In the worst case, there exists  $k \in \mathbb{N}$  so that  $M \leq 2^k$ ; that means that  $Q'$  is contained in  $k + 2$  strips of the form  $\{(\xi_1, \xi_2) : \xi_1 \in [2^j, 2^{j+1}], |\xi_2| \leq \xi_1\}$ . In each of these strips we have at most  $2^{j/2+1}$  trapezoids. It follows that  $Q'$  is contained in  $U_0 \cup \bigcup_{\ell=-1}^1 P'B^\ell \cup \bigcup_{j=1}^{k+1} \bigcup_{|\ell| \leq 2^{j/2}} P'B^\ell A^j$ , that is, the number of overlapping sets is at most  $1 + 3 + \sum_{j=1}^{k+1} 2^{j/2+1} < \infty$ . Otherwise, we know that  $P'$  must be contained in  $\tilde{Q}'$ , since

$P'$  is minimal.  $\square$

### 2.4.3 Shearlet smoothness spaces

Having established the existence of a structured admissible covering associated with the shearlet decomposition, we can now define the associated smoothness spaces. Specifically, letting  $\mathcal{Q}$  be the shearlet-type covering with  $\mathcal{T}_{\mathcal{M}}$  the corresponding family of affine transformations (given in Proposition 7) and choosing  $w(j, \ell) = 2^j$ ,  $(j, \ell) \in \mathcal{M}$ , to be the  $\mathcal{Q}$ -moderate weight<sup>1</sup>, the **shearlet smoothness spaces** are defined by

$$S_{p,q}^{\beta}(\mathcal{T}_{\mathcal{M}}, w) := D(\mathcal{Q}, L^p, (\ell^q)_{2^{\beta}}).$$

As observed above, these spaces are independent from the choice of a particular BAPU.

Not surprisingly, the systems of shearlets  $\{\tilde{\psi}_{\mu} : \mu \in M\}$  introduced in Section 2.3 are closely associated with the shearlet smoothness spaces. Specifically, let us write the elements of the shearlet system in  $\mathcal{P}_1$ , in the Fourier domain, as

$$\hat{\psi}_{j,\ell,k}(\xi) = \hat{\psi}(\xi A^{-j} B^{-\ell}) u_{j,\ell,k}(\xi),$$

where  $u_{j,\ell,k}(\xi) = |\det A|^{-j/2} e^{2\pi i \xi A^{-j} B^{-\ell} k}$ . We can proceed similarly for the other elements of the shearlet system (2.3.16). Hence, using the notation of Section 2.4.1 we can write each element of the shearlet system, in the Fourier domain, in the form  $\hat{\psi}_T(\xi) u_{j,\ell,k,h}(\xi)$ , where  $\hat{\psi}_T(\xi) = \hat{\psi}(\xi T^{-1})$ , for  $T \in \mathcal{T}_{\mathcal{M}}$  and  $\mathcal{T}_{\mathcal{M}}$  given by (2.4.22).

We have the following observation.

**Proposition 9.** *The functions  $\{\hat{\psi}_T : T \in \mathcal{T}_{\mathcal{M}}\}$  given above form a squared BAPU with respect to the shearlet-type covering  $\mathcal{Q}$ .*

---

<sup>1</sup>Here the  $\mathcal{Q}$ -moderate function  $w$  is the first axis projection, and the sequence of points  $x_i \in Q_i$  is given by  $x_i = (2^j, 0)$ .

**Proof.**

It will be sufficient to consider the elements in the cone region  $\mathcal{P}_1$ ; a similar argument holds for the other elements. In this case, we consider the set of transformations  $T_{(j,\ell)} = A^{-j}B^{-\ell}$ ,  $(j, \ell) \in \mathcal{M}$  and the corresponding functions  $\hat{\psi}_{T_{(j,\ell)}}(\xi) = \hat{\psi}(\xi A^{-j}B^{-\ell})$ .

Notice that  $\hat{\psi}(\xi A^{-j}B^{-\ell})$  is supported in the set  $\Sigma_{j,\ell}$  given by (2.3.14). Also, using equations (2.3.7), (2.3.8) and the definition of  $\psi$  itself, we have that, for  $\xi \in \mathcal{P}_1$ ,

$$\begin{aligned} \sum_{\substack{j \geq 0 \\ |\ell| \leq 2^{\lfloor j/2 \rfloor}}} |\hat{\psi}(\xi A^{-j}B^{-\ell})|^2 &= \sum_{\substack{j \geq 0 \\ |\ell| \leq 2^{\lfloor j/2 \rfloor}}} |W^2(\xi A^{-j})| |V^2(\xi A^{-j}B^{-\ell})| \\ &= \sum_{j \geq 0} |W^2(\xi A^{-j})| \sum_{|\ell| \leq 2^{\lfloor j/2 \rfloor}} |V^2(\xi A^{-j}B^{-\ell})| = 1. \end{aligned}$$

A direct computations shows that, for  $p \in [0, \infty]$  and  $T \in GL_2(\mathbb{R})$ ,

$$\|\psi_T\|_p = |\det T|^{1-\frac{1}{p}} \|\psi\|_p.$$

Hence,

$$\|\psi_{T_{(j,\ell)}}\|_p = |\det A|^{j(1-\frac{1}{p})} \|\psi\|_p = 2^{\frac{3j}{2}(1-\frac{1}{p})} \|\psi\|_p,$$

and we observe that, whenever we consider  $0 \leq p \leq 1$ , this quantity is uniformly bounded for all  $j, \ell$ . In other words,  $\hat{\psi}(D)$  is a bounded operator on  $L^p(\mathbb{R}^2)$  for  $0 < p \leq 1$ .  $\square$

The result presented above shows that the family of shearlets  $\{\tilde{\psi}_\mu : \mu \in M\}$  is a system of form (2.2.3). Hence, thanks to Proposition 4, we have that

$$\|f\|_{S_{p,q}^\beta} \approx \left( \sum_{j,\ell,d} 2^{jq(\beta + \frac{3}{2}(\frac{1}{2} - \frac{1}{p}))} \left( \sum_{k \in \mathbb{Z}^2} |\langle f, \psi_{j,\ell,k}^{(d)} \rangle|^p \right)^{q/p} \right)^{1/q}.$$

### 2.4.4 Embedding results

As mentioned at the end of Sec. 2.2, the dyadic covering of the Fourier space is associated with the Besov spaces. In dimensions  $D = 2$ , let us consider the dyadic partition of the Fourier plane into the Cartesian coronae  $\widehat{\mathbb{R}}^2 = \bigcup_{j \in \mathbb{Z}} C_j$ , where

$$C_j = [-2^{j+1}, 2^{j+1}]^2 \setminus [-2^j, 2^j]^2. \quad (2.4.23)$$

It is intuitive that the shearlet-type covering can be considered as refinement of the dyadic covering of  $\mathbb{R}^2$ , suggesting the existence of a close relationship between Besov spaces and shearlet smoothness spaces. Indeed, we have the following observation which is similar to Lemma 7.4 in [34].

**Proposition 10.** *For  $0 < p \leq \infty$ ,  $0 < q < \infty$  and  $\beta \in \mathbb{R}$  we have*

$$B_{p,q}^{\beta + \frac{1}{2q}}(\mathbb{R}^2) \hookrightarrow S_{p,q}^{\beta}(\mathbb{R}^2).$$

*Likewise:*

$$S_{p,q}^{\beta-s}(\mathbb{R}^2) \hookrightarrow B_{p,q}^{\beta}(\mathbb{R}^2),$$

where  $s = \frac{1}{2} (\max(1, 1/p) - \min(1, 1/q))$ .

**Proof.** Let  $\{\phi_{j,\ell}\}$  be a BAPU corresponding to the shearlet-type covering  $\mathcal{T}_{\mathcal{M}}$  given in Proposition 7, and  $\Omega = \{\omega_j\}_{j \in \mathbb{N}}$  be a partition of unity with support on the dyadic frequency bands (2.4.23), satisfying

$$\|f\|_{B_{p,q}^{\beta}} \approx \left( \sum_{j \in \mathbb{N}} (2^{\beta j} \|\omega_j(D) f\|_p)^q \right)^{1/q}.$$

By the properties of the shearlet-like covering, it is clear that we can choose  $\Omega$  such that

$$\text{supp}(\phi_{j,\ell}) \subset \text{supp}(\tilde{\omega}_j), \ell \in L_j, \text{ and } \text{supp}(\omega_j) \subset \cup_{\ell \in L_j} \text{supp}(\tilde{\phi}_{j,\ell}),$$

for all  $j \in \mathbb{N}$ , where  $L_j = \{\ell : -2^{\lfloor j/2 \rfloor} \leq \ell \leq 2^{\lfloor j/2 \rfloor}\}$ . Observe that the cardinality of  $L_j$  is  $2(2^{\lfloor j/2 \rfloor + 1} + 1)$  (recall that there are 2 sets of shearlets: horizontal and vertical). Thus, for each level  $j$  there are about  $C 2^{j/2}$  trapezoids in the covering of  $\{\phi_{j,\ell}\}$ , for some constant  $C > 0$ . It follows that:

$$\begin{aligned} \sum_{j \in \mathbb{N}} \sum_{\ell \in L_j} 2^{\beta q j} \|\phi_{j,\ell}(D) f\|_p^q &= \sum_{j \in \mathbb{N}} \sum_{\ell \in L_j} 2^{\beta q j} \|\phi_{j,\ell}(D) \tilde{\omega}_j(D) f\|_p^q \\ &\leq C \sum_{j \in \mathbb{N}} \sum_{\ell \in L_j} 2^{\beta q j} \|\tilde{\omega}_j(D) f\|_p^q \\ &\leq C \sum_{j \in \mathbb{N}} 2^{j/2} 2^{\beta q j} \|\tilde{\omega}_j(D) f\|_p^q. \end{aligned}$$

Using a similar calculation, for  $p \geq 1$  and  $q < 1$  we have:

$$\begin{aligned} \sum_{j \in \mathbb{N}} (2^{\beta j} \|\omega_j(D) f\|_p)^q &= \sum_{j \in \mathbb{N}} \left( 2^{\beta j} \|\omega_j(D) \sum_{\ell \in L_j} \tilde{\phi}_{j,\ell}(D) f\|_p \right)^q \\ &\leq C \sum_{j \in \mathbb{N}} \left( 2^{\beta j} \left\| \sum_{\ell \in L_j} \tilde{\phi}_{j,\ell}(D) f \right\|_p \right)^q \\ &\leq C \sum_{j \in \mathbb{N}} \left( 2^{\beta j} \sum_{\ell \in L_j} \|\tilde{\phi}_{j,\ell}(D) f\|_p \right)^q \\ &\leq C \sum_{j \in \mathbb{N}} \sum_{\ell \in L_j} \left( 2^{\beta j} \|\tilde{\phi}_{j,\ell}(D) f\|_p \right)^q. \end{aligned}$$

The other cases follow similarly by using Hölder inequality and the bound on the sums over  $\ell$ .  $\square$

### 2.4.5 Equivalence with curvelet spaces

As mentioned in the introduction, curvelets provide an approach alternative to shearlets for the construction of sparse multidimensional representations. Using the same approach adopted in this work, a notion of curvelet spaces is introduced in [34] which is associated

with a structured family of affine transformations including rotations and dilations. We can show that the shearlet smoothness spaces defined in Section 2.4.3 are equivalent to the curvelet smoothness spaces with equivalent norm.

In order to state our result, let us recall the definition of **curvelet covering** in dimension two. This is defined as the collection of the sets

$$\begin{aligned} \{S_{j,l} := \{(\rho, \theta) \in \mathbb{R}^* \times \mathbb{R} : 2^{j-3} \leq \rho \leq 2^{j-2}, \theta \in [l\frac{2\pi}{2^{\lfloor j/2 \rfloor}}, (l+1)\frac{2\pi}{2^{\lfloor j/2 \rfloor}}]\}, j \in \mathbb{N}, l = 0, \dots, 2^{\lfloor j/2 \rfloor} - 1\} \cup \\ \cup \{S_0 := C_{(0,0)}(1/8)\}, \end{aligned}$$

where  $C_{(0,0)}(1/8)$  is the circumference of radius  $1/8$  centered in  $(0,0)$ . This covering can be obtained from the family of affine transformations

$$\mathcal{T}_C := \{D_{j,l} : j \in \mathbb{N}, l = 0, \dots, 2^{\lfloor j/2 \rfloor} - 1\} \cup \{D_0\}$$

acting on  $S_{0,0}$ , where  $D_{j,l}$  is the affine transformation that brings the element  $(\rho, \theta)$  into the element  $(\rho, \theta)\text{diag}(2^j, 2^{-\lfloor j/2 \rfloor}) + c_{j,l}$ , with  $c_{j,l} = (0, l\frac{2\pi}{2^{\lfloor j/2 \rfloor}})$  and  $D_0$  denoting the affine transformation that maps  $S_0$  in  $S_{1,0}$ . In fact,  $\{D_{j,l}S_{0,0} : j \in \mathbb{N}, l = 0, \dots, 2^{\lfloor j/2 \rfloor} - 1\} \cup \{D_0\tilde{S}_{1,0}\}$  is a covering for  $\mathbb{R}^2$ . By construction, the set  $S_{0,0}$  is compactly contained in  $\tilde{S}_{0,0}$ , which is also a covering of  $\mathbb{R}^2$ . Thus we have an admissible structured covering of  $\mathbb{R}^2$ .

As in [34], the curvelet spaces are defined as the decomposition spaces which are associated with their curvelet-type covering of  $\mathbb{R}^2$ . We can now state the following observation.

**Proposition 11.** *The shearlet and curvelet spaces are identical as decomposition space with equivalent norms.*

**Proof.** By Theorem 1 it is sufficient to show that the curvelet-type covering and the shearlet-type covering are equivalent. For that, we need to show that each trapezoid of the form  $Q_{j,\ell} = QB^\ell A^j$  intersects a finite number of curvelet type tiles and, vice versa, that

each set  $S_{j,l}$  is covered by a finite number of shearlet type tiles. As usual, owing to the symmetry of the construction, it is sufficient to consider the cone-shaped region  $\mathcal{P}_1$ .

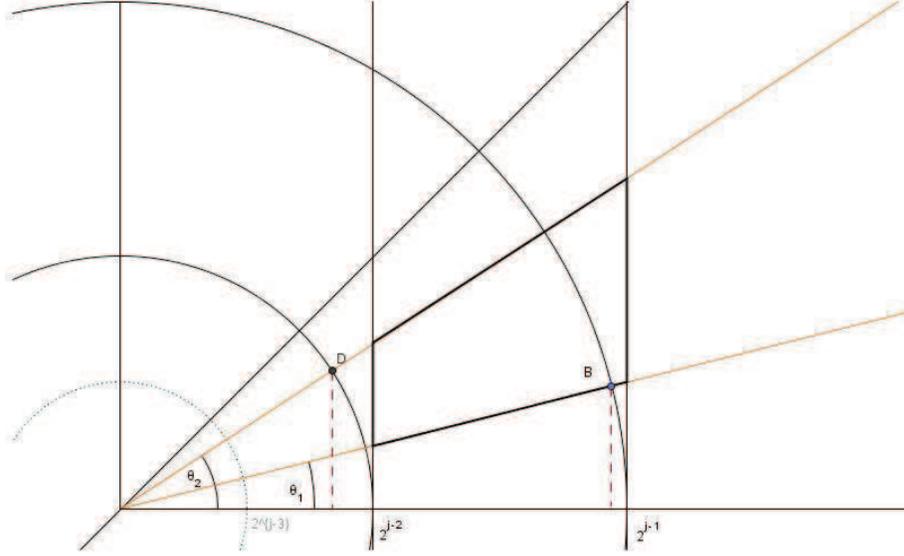


Figure 2.2: Equivalence of shearlet and curvelet coverings.

As illustrated in Figure 2.2, any set  $S_{j,l}$ , at level  $j$ , extends horizontally from the value  $\cos(\theta_1)2^{j-1} \leq 2^{j-1}$  (i.e., the  $\xi_1$ -coordinate of the point  $B$  in the figure) up to  $\cos(\theta_2)2^{j-2} \geq 2^{j-2} \cos(\pi/4) = 2^{j-3+1/2}$  (i.e., the  $\xi_1$ -coordinate of the point  $D$  in the figure). Hence, to cover this curvelet-type tile, we need shearlet-type tiles associated with scale parameters ranging from  $j-1$  to  $j$  (so that the  $\xi_1$  axis is covered between  $2^{j-3}$  and  $2^{j-1}$ ). Next, observe that if a generic angle  $\theta \in [0, \pi/4]$  is split into two equal size angles  $\theta_1 = \theta/2$ , the following inequality is satisfied:  $1/2 < \tan(\theta) - \tan(\theta_1) < 3/4$ . As we repeat the subdivision of the angle into equal size angles we use same observations so that  $\forall j \geq 0, N = 1, \dots, 2^{j/2} - 1$ , we have that

$$(1/2)^{\lfloor j/2 \rfloor} \leq \left[ \tan\left(N \frac{2\pi}{2^{\lfloor j/2 \rfloor}}\right) - \tan\left((N-1) \frac{2\pi}{2^{\lfloor j/2 \rfloor}}\right) \right] < (3/4)^{\lfloor j/2 \rfloor} \leq \left(\frac{2}{4}\right)^{\lfloor j/2 \rfloor} = 2^{-\lfloor j/2 \rfloor} \leq 2^{-j/2+1}.$$

We can now estimate the number of shearlet-type tiles needed to cover a given curvelet-type

tile. By the above observation, we have that

$$2^{j-1} \left[ \tan\left(N \frac{2\pi}{2^{\lfloor j/2 \rfloor}}\right) - \tan\left((N-1) \frac{2\pi}{2^{\lfloor j/2 \rfloor}}\right) \right] < 2^{j-1} 2^{-j/2+1} = 2^{j/2}.$$

This implies that, for  $\xi_1 \in [2^{j-2}, 2^{j-1}]$ , we need less than  $\lfloor \frac{2^{j/2}}{2^{j/2-1}} \rfloor + 2 \leq 4$  shearlet-type tiles to cover the curvelet-type tile  $S_{j,l}$ . Similarly, for  $\xi_1 \in [2^{j-3}, 2^{j-2}]$ , we need less than  $\lfloor \frac{2^{j/2-3/2}}{2^{j/2-3/2}} \rfloor + 2 \leq 3$  shearlet type tiles to cover  $S_{j,l}$ . Finally, since the boundary of the set  $S_{j,k}$  contains points with  $\xi_1 = 2^{j-1}$ , we include two more shearlet-type tiles to cover those points. In conclusion, we need 9 shearlet-type tiles to cover the set  $S_{j,l}$ .

Vice versa, the argument is similar. A fixed shearlet-type tile  $Q_{j,\ell}$  is supported in the strip-domain  $\xi_1 \in [2^{j-2}, 2^{j-1}]$ , and has left height  $2^{j/2-1}$  and right height  $2^{j/2}$ . Inside the corona  $2^{j-2} \leq |\xi| \leq 2^{j-1}$  there are less than  $2 \lfloor \frac{2^{j/2-1}}{2^{j-1} 2^{-j/2}} \rfloor + 2 = 4$  curvelet-type tiles intersecting  $Q_{j,\ell}$ . In fact, we have that

$$2^{j-1} \left[ \tan\left(N \frac{2\pi}{2^{\lfloor j/2 \rfloor}}\right) - \tan\left((N-1) \frac{2\pi}{2^{\lfloor j/2 \rfloor}}\right) \right] \geq 2^{j-1} (1/2)^{\lfloor j/2 \rfloor} \geq 2^{j/2-1}.$$

Using the same argument, we have that inside the corona  $2^{j-1} \leq |\xi| \leq 2^j$  at most  $2 \lfloor \frac{2^{j/2}}{2^{j-1} 2^{-(j+1)/2}} \rfloor + 2 \geq 2 \lfloor \frac{2^{j/2}}{2^{j-2} 2^{-(j+1)/2}} \rfloor + 2 = 4$  curvelet-type tiles intersect the set  $Q_{j,\ell}$ . Finally, the trapezoid  $Q_{j,\ell}$  can intersect the corona  $2^{j-3} \leq |\xi| \leq 2^{j-2}$  at most in one point, so it is sufficient to include at most 2 additional curvelet-type tiles. In conclusion, we need  $4 + 4 + 2 = 10$  curvelet-type tiles to cover one fixed shearlet-type tile.  $\square$

# Chapter 3

## 3D Shearlet Representations

The shearlets were originally introduced in [1, 2] within a larger class of affine-like systems called wavelets with composite dilations. Additionally, unlike curvelets and other directional systems recently introduced in the literature, the elements of the shearlet system form an affine-like system whose elements are generated from the action of translation and dilation operators on a finite set of generators. This property provides additional simplicity of construction and a connection with the theory of square integrable group representations of the affine group [39, 48]. The shearlet approach provides a general method for the construction of function systems ranging at various scales, locations and orientations according to various orthogonal transformations controlled by shearing matrices. 3D shearlets are particularly useful in video denoising and the processing of many types of scientific data, such as biological data, where it is important to process data in native resolution.

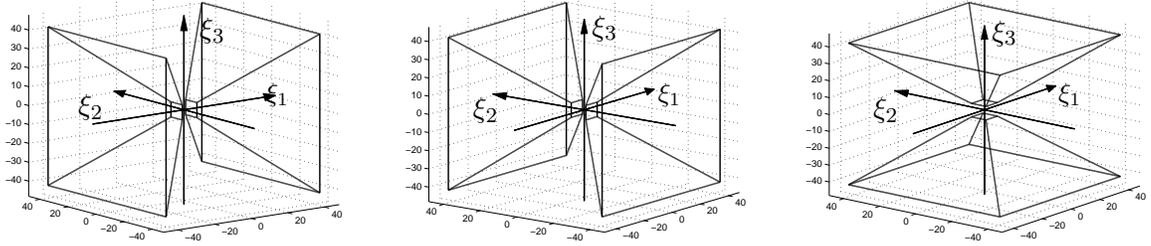


Figure 3.1: From left to right, the figure illustrates the pyramidal regions  $\mathcal{P}_1$ ,  $\mathcal{P}_2$ , and  $\mathcal{P}_3$  in the frequency space  $\widehat{\mathbb{R}}^3$ .

### 3.1 Shearlet in 3D

In dimension  $D = 3$ , a shearlet system is obtained by appropriately combining 3 systems of functions associated with the pyramidal regions

$$\mathcal{P}_1 = \left\{ (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : \left| \frac{\xi_2}{\xi_1} \right| \leq 1, \left| \frac{\xi_3}{\xi_1} \right| \leq 1 \right\},$$

$$\mathcal{P}_2 = \left\{ (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : \left| \frac{\xi_1}{\xi_2} \right| < 1, \left| \frac{\xi_3}{\xi_2} \right| \leq 1 \right\},$$

$$\mathcal{P}_3 = \left\{ (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : \left| \frac{\xi_1}{\xi_3} \right| < 1, \left| \frac{\xi_2}{\xi_3} \right| < 1 \right\},$$

in which the Fourier space  $\widehat{\mathbb{R}}^3$  is partitioned (see Fig. 3.1).

To define such systems, let  $\phi$  be a  $C^\infty$  univariate function such that  $0 \leq \hat{\phi} \leq 1$ ,  $\hat{\phi} = 1$  on  $[-\frac{1}{16}, \frac{1}{16}]$  and  $\hat{\phi} = 0$  outside the interval  $[-\frac{1}{8}, \frac{1}{8}]$ . That is,  $\phi$  is the scaling function of a Meyer wavelet, rescaled so that its frequency support is contained the interval  $[-\frac{1}{8}, \frac{1}{8}]$ . For  $\xi = (\xi_1, \xi_2, \xi_3) \in \widehat{\mathbb{R}}^3$ , define

$$\widehat{\Phi}(\xi) = \widehat{\Phi}(\xi_1, \xi_2, \xi_3) = \hat{\phi}(\xi_1) \hat{\phi}(\xi_2) \hat{\phi}(\xi_3) \tag{3.1.1}$$

and let  $W(\xi) = \sqrt{\widehat{\Phi}^2(2^{-2}\xi) - \widehat{\Phi}^2(\xi)}$ . It follows that

$$\widehat{\Phi}^2(\xi) + \sum_{j \geq 0} W^2(2^{-2j}\xi) = 1 \quad \text{for } \xi \in \mathbb{R}^3. \quad (3.1.2)$$

Notice that each function  $W_j = W(2^{-2j} \cdot)$ ,  $j \geq 0$ , is supported inside the Cartesian corona

$$[-2^{2j-1}, 2^{2j-1}]^3 \setminus [-2^{2j-4}, 2^{2j-4}]^3 \subset \widehat{\mathbb{R}}^3,$$

and the functions  $W_j^2$ ,  $j \geq 0$ , produce a smooth tiling of  $\widehat{\mathbb{R}}^3$ . Next, let  $V \in C^\infty(\mathbb{R})$  be such that  $\text{supp } V \subset [-1, 1]$  and

$$|V(u-1)|^2 + |V(u)|^2 + |V(u+1)|^2 = 1 \quad \text{for } |u| \leq 1. \quad (3.1.3)$$

In addition, we will assume that  $V(0) = 1$  and that  $V^{(n)}(0) = 0$  for all  $n \geq 1$ . It was shown in [5] that there are several examples of functions satisfying these properties. It follows from equation (3.1.3) that, for any  $j \geq 0$ ,

$$\sum_{m=-2^j}^{2^j} |V(2^j u - m)|^2 = 1, \quad \text{for } |u| \leq 1. \quad (3.1.4)$$

For  $d = 1, 2, 3$ ,  $\ell = (\ell_1, \ell_2) \in \mathbb{Z}^2$ , the 3D *shearlet systems associated with the pyramidal regions*  $\mathcal{P}_d$  are defined as the collections

$$\{\psi_{j,\ell,k}^{(d)} : j \geq 0, -2^j \leq \ell_1, \ell_2 \leq 2^j, k \in \mathbb{Z}^3\}, \quad (3.1.5)$$

where

$$\widehat{\psi}_{j,\ell,k}^{(d)}(\xi) = |\det A_{(d)}|^{-j/2} W(2^{-2j}\xi) F_{(d)}(\xi A_{(d)}^{-j} B_{(d)}^{[-\ell]}) e^{2\pi i \xi A_{(d)}^{-j} B_{(d)}^{[-\ell]} k}, \quad (3.1.6)$$

$$F_{(1)}(\xi_1, \xi_2, \xi_3) = V\left(\frac{\xi_2}{\xi_1}\right)V\left(\frac{\xi_3}{\xi_1}\right), \quad F_{(2)}(\xi_1, \xi_2, \xi_3) = V\left(\frac{\xi_1}{\xi_2}\right)V\left(\frac{\xi_3}{\xi_2}\right), \quad F_{(3)}(\xi_1, \xi_2, \xi_3) = V\left(\frac{\xi_1}{\xi_3}\right)V\left(\frac{\xi_2}{\xi_3}\right),$$

the anisotropic dilation matrices  $A_{(d)}$  are given by

$$A_{(1)} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_{(2)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_{(3)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

and the *shear matrices* are defined by

$$B_{(1)}^{[\ell]} = \begin{pmatrix} 1 & \ell_1 & \ell_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, B_{(2)}^{[\ell]} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_1 & 1 & \ell_2 \\ 0 & 0 & 1 \end{pmatrix}, B_{(3)}^{[\ell]} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \ell_1 & \ell_2 & 1 \end{pmatrix}.$$

Due to the assumptions on  $W$  and  $v$ , the elements of the system of shearlets (3.1.5) are well localized and bandlimited. In particular, the shearlets  $\hat{\psi}_{j,\ell,k}^{(1)}(\xi)$  can be written more explicitly as

$$\hat{\psi}_{j,\ell_1,\ell_2,k}^{(1)}(\xi) = 2^{-2j} W(2^{-2j}\xi) V\left(2^j \frac{\xi_2}{\xi_1} - \ell_1\right) V\left(2^j \frac{\xi_3}{\xi_1} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-\ell_1, -\ell_2]} k}, \quad (3.1.7)$$

showing that their supports are contained inside the trapezoidal regions

$$\{(\xi_1, \xi_2, \xi_3) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], |\frac{\xi_2}{\xi_1} - \ell_1 2^{-j}| \leq 2^{-j}, |\frac{\xi_3}{\xi_1} - \ell_2 2^{-j}| \leq 2^{-j}\}.$$

This expression shows that these support regions become increasingly more elongated at fine scales, due to the action of the anisotropic dilation matrices  $A_{(1)}^j$ , with the orientations of these regions controlled by the shearing parameters  $\ell_1, \ell_2$ . A typical support region is illustrated in Fig. 3.2. Similar properties hold for the elements associated with the regions  $\mathcal{P}_2$  and  $\mathcal{P}_3$ .

A Parseval frame of shearlets for  $L^2(\mathbb{R}^3)$  is obtained by using an appropriate combination of the systems of shearlets associated with the 3 pyramidal regions  $\mathcal{P}_d$ ,  $d = 1, 2, 3$ , together with a coarse scale system, which will take care of the low frequency region. In order to build such system in a way that all its elements are smooth in the Fourier domain, one has to appropriately define the elements of the shearlet systems overlapping the boundaries of the pyramidal regions  $\mathcal{P}_d$  in the Fourier domain.

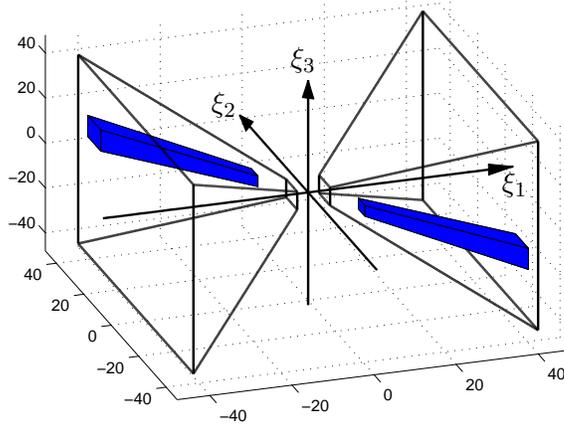


Figure 3.2: Frequency support of a representative shearlet function  $\psi_{j,\ell,k}$ , inside the pyramidal region  $\mathcal{P}_1$ . The orientation of the support region is controlled by  $\ell = (\ell_1, \ell_2)$ ; its shape is becoming more elongated as  $j$  increases ( $j = 4$  in this plot).

Hence, we define the *3D shearlet systems for  $L^2(\mathbb{R}^3)$*  as the collections

$$\begin{aligned} & \left\{ \tilde{\psi}_{-1,k} : k \in \mathbb{Z}^3 \right\} \cup \left\{ \tilde{\psi}_{j,\ell,k,d} : j \geq 0, |\ell_1| < 2^j, |\ell_2| \leq 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3 \right\} \\ & \cup \left\{ \tilde{\psi}_{j,\ell,k} : j \geq 0, \ell_1, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3 \right\} \end{aligned} \quad (3.1.8)$$

consisting of:

- the *coarse scale shearlets*  $\{\tilde{\psi}_{-1,k} = \Phi(\cdot - k) : k \in \mathbb{Z}^3\}$ , where  $\Phi$  is given by (3.1.1);
- the *interior shearlets*  $\{\tilde{\psi}_{j,\ell,k,d} = \psi_{j,\ell,k}^{(d)} : j \geq 0, |\ell_1||\ell_2| < 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$ , where  $\psi_{j,\ell,k}^{(d)}$  are given by (3.1.6);
- the *boundary shearlets*  $\{\tilde{\psi}_{j,\ell,k} : j \geq 0, |\ell_1| \leq 2^j, \ell_2 = \pm 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$ .

These boundary shearlets are obtained by adding together the functions  $\psi_{j,\ell,k}^{(1)}$ ,  $\psi_{j,\ell,k}^{(2)}$  and  $\psi_{j,\ell,k}^{(3)}$ , for  $\ell_1 = \pm 2^j$  or  $\ell_2 = \pm 2^j$ , after they have been restricted to their respective pyramidal

regions. For example (see [19] for all cases and additional detail), when  $\ell_1 = \pm 2^j$ ,  $|\ell_2| < 2^j$ , we define

$$(\tilde{\psi}_{j,\ell_1,\ell_2,k,1})^\wedge(\xi) = \begin{cases} 2^{-2j} W(2^{-2j}\xi) V\left(2^j \frac{\xi_2}{\xi_1} - \ell_1\right) V\left(2^j \frac{\xi_3}{\xi_1} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_1, \\ 2^{-2j} W(2^{-2j}\xi) V\left(2^j \frac{\xi_1}{\xi_2} - \ell_1\right) V\left(2^j \frac{\xi_3}{\xi_2} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_2; \end{cases}$$

when  $\ell_1, \ell_2 = \pm 2^j$ , we define

$$(\tilde{\psi}_{j,\ell_1,\ell_2,k})^\wedge(\xi) = \begin{cases} 2^{-2j} W(2^{-2j}\xi) V\left(2^j \frac{\xi_2}{\xi_1} - \ell_1\right) V\left(2^j \frac{\xi_3}{\xi_1} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_1, \\ 2^{-2j} W(2^{-2j}\xi) V\left(2^j \frac{\xi_1}{\xi_2} - \ell_1\right) V\left(2^j \frac{\xi_3}{\xi_2} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_2, \\ 2^{-2j} W(2^{-2j}\xi) V\left(2^j \frac{\xi_1}{\xi_3} - \ell_1\right) V\left(2^j \frac{\xi_2}{\xi_3} - \ell_2\right) e^{2\pi i \xi A_{(1)}^{-j} B_{(1)}^{[-(\ell_1,\ell_2)]} k}, & \text{if } \xi \in \mathcal{P}_3. \end{cases}$$

Notice that, thanks on the assumptions on  $W$  and  $V$ , the piecewise defined boundary shearlet functions are smooth and compactly supported in the Fourier domain (see [16, 19] for additional detail). In addition, the system of shearlets (3.1.8) is a Parseval frame. To state this result, let us introduce the following compact notation to write the 3D shearlet system (3.1.8) as

$$\{\tilde{\psi}_\mu, \mu \in \mathcal{M}\}, \quad (3.1.9)$$

where  $\mathcal{M} = \mathcal{M}_C \cup \mathcal{M}_I \cup \mathcal{M}_B$  are the indices associated with the *coarse scale shearlets*, the *interior shearlets* and the *boundary shearlets* given by

- $\mathcal{M}_C = \{\mu = (j, k) : j = -1, k \in \mathbb{Z}^3\}$ ;
- $\mathcal{M}_I = \{\mu = (j, \ell_1, \ell_2, k, d) : j \geq 0, |\ell_1|, |\ell_2| < 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$ ;
- $\mathcal{M}_B = \{\mu = (j, \ell_1, \ell_2, k, d) : j \geq 0, |\ell_1| \leq 2^j, \ell_2 \pm 2^j, k \in \mathbb{Z}^3, d = 1, 2, 3\}$ .

Hence we have the following result whose proof is found in [19]:

**Theorem 12.** *The 3D system of shearlets (3.1.9) is a Parseval frame of  $L^2(\mathbb{R}^3)$ . That is, for any  $f \in L^2(\mathbb{R}^3)$ ,*

$$\sum_{\mu \in \mathcal{M}} |\langle f, \tilde{\psi}_\mu \rangle|^2 = \|f\|^2.$$

The mapping from  $f \in L^2(\mathbb{R}^3)$  into the elements  $\langle f, \tilde{\psi}_\mu \rangle$ ,  $\mu \in \mathcal{M}$ , is called the *3D shearlet transform*.

As mentioned above, it is proved in [15, 16] that the 3D Parseval frame of shearlets  $\{\tilde{\psi}_\mu, \mu \in \mathcal{M}\}$  achieves the essentially optimal approximation rate (1.0.1) for functions of 3 variables which are  $C^2$  regular away from discontinuities along  $C^2$  surfaces.

# Chapter 4

## 3D Discrete Shearlet Transform(3D DShT)

In this section, we present a digital implementation of the 3D shearlet transform introduced above. Following essentially the same architecture as the algorithm of the 2D Discrete Shearlet Transform in [8], this new implementation can be described as the cascade of a multiscale decomposition, based on a version of the Laplacian pyramid filter, followed by a stage of directional filtering. The main novelty of the 3D approach consists in the design of the directional filtering stage, which attempts to reproduce the frequency decomposition faithfully provided by the corresponding mathematical transform by using a method based on the pseudo-spherical Fourier transform.

Let us start by expressing the elements of the shearlet system in a form that is more convenient for deriving an algorithmic implementation of the shearlet transform. For  $\xi = (\xi_1, \xi_2, \xi_3)$  in  $\widehat{\mathbb{R}}^3$ ,  $j \geq 0$ , and  $-2^j \leq \ell_1, \ell_2 \leq 2^j$ , we define the directional windowing functions

$$U_{j,\ell}^{(1)}(\xi) =$$

---


$$= \begin{cases} V(2^j \frac{\xi_2}{\xi_1} - \ell_1) V(2^j \frac{\xi_3}{\xi_1} - \ell_2) & \text{if } |\ell_1|, |\ell_2| < 2^j, \\ V(2^j \frac{\xi_2}{\xi_1} - \ell_1) V(2^j \frac{\xi_3}{\xi_1} - \ell_2) \mathcal{X}_{\mathcal{P}_1}(\xi) + V(2^j \frac{\xi_1}{\xi_2} - \ell_1) V(2^j \frac{\xi_3}{\xi_2} - \ell_2) \mathcal{X}_{\mathcal{P}_2}(\xi) & \text{if } \ell_1 = \pm 2^j, |\ell_2| < 2^j; \\ V(2^j \frac{\xi_2}{\xi_1} - \ell_1) V(2^j \frac{\xi_3}{\xi_1} - \ell_2) \mathcal{X}_{\mathcal{P}_1}(\xi) + V(2^j \frac{\xi_1}{\xi_2} - \ell_1) V(2^j \frac{\xi_3}{\xi_2} - \ell_2) \mathcal{X}_{\mathcal{P}_2}(\xi) \\ + V(2^j \frac{\xi_1}{\xi_3} - \ell_1) V(2^j \frac{\xi_2}{\xi_3} - \ell_2) \mathcal{X}_{\mathcal{P}_3}(\xi) & \text{if } \ell_1, \ell_2 = \pm 2^j. \end{cases}$$

Notice that only the elements  $U_{j,\ell}^{(1)}$  with indices  $|\ell_1|, |\ell_2| < 2^j$  are strictly contained inside the region  $\mathcal{P}_1$ ; the elements with indices  $\ell_1 = \pm 2^j$  or  $\ell_2 = \pm 2^j$  are supported across  $\mathcal{P}_1$  and some other pyramidal region. However, it is convenient to associate this family of functions with the index 1. We define the functions  $U_{j,\ell}^{(2)}$  and  $U_{j,\ell}^{(3)}$  associated with the pyramidal regions  $\mathcal{P}_2$  and  $\mathcal{P}_3$  in a similar way<sup>1</sup>. Using this notation, we can write each element of the 3D shearlet system as

$$\hat{\psi}_{j,\ell,k}^{(d)} = 2^{-2j} W(2^{-2j}\xi) U_{j,\ell}^{(d)}(\xi) e^{-2\pi i \xi A_{(d)}^{-j} B_{(d)}^{[-\ell]} k}.$$

It follows from the properties of the shearlet construction that

$$\sum_{d=1}^3 \sum_{j \geq 0} \sum_{\ell_1 = -2^j}^{2^j} \sum_{\ell_2 = -2^j}^{2^j} |W(2^{-2j}(\xi))|^2 |U_{j,\ell}^{(d)}(\xi)|^2 = 1, \quad \text{for } |\xi_1|, |\xi_2|, |\xi_3| \geq \frac{1}{8}. \quad (4.0.1)$$

The (fine scale) 3D *shearlet transform* of  $f \in L(\mathbb{R}^3)$  can be expressed as the mapping from  $f$  into the *shearlet coefficients*

$$\langle f, \psi_{j,\ell,k}^{(d)} \rangle = \int_{\mathbb{R}^3} \hat{f}(\xi) W(2^{-2j}\xi) U_{j,\ell}^{(d)}(\xi) e^{2\pi i \xi A_{(d)}^{-j} B_{(d)}^{[-\ell]} k} d\xi, \quad (4.0.2)$$

where  $j \geq 0$ ,  $\ell = (\ell_1, \ell_2)$  with  $|\ell_1|, |\ell_2| \leq 2^j$ ,  $k \in \mathbb{Z}^3$  and  $d = 1, 2, 3$ .

This expression shows that the shearlet transform of  $f$ , for  $j, \ell, k$  and  $d$  fixed, can be computed using the following steps:

1. In the frequency domain, compute the  $j$ -th subband decomposition of  $f$  as  $\hat{f}_j(\xi) = \hat{f}(\xi) W(2^{-2j}\xi)$ .

---

<sup>1</sup>Notice however that they do not contain the boundary term for  $\ell_1, \ell_2 = \pm 2^j$ , which only needs to be included once.

2. Next (still in the frequency domain), compute the  $(j, \ell, d)$ -th directional subband decomposition of  $f$  as  $\hat{f}_{j,\ell,d}(\xi) = \hat{f}_j(\xi) U_{j,\ell}^{(d)}(\xi)$ .
3. Compute the inverse Fourier transform. This step can be represented as a convolution of the  $j$ -th subband decomposition of  $f$  and the directional filter  $\check{U}_{j,\ell}^{(d)}$ , that is,  $\langle f, \psi_{j,\ell,k}^{(d)} \rangle = f_j * \check{U}_{j,\ell}^{(d)}(A_d^{-j} B_d^{-\ell} k)$ .

Hence, the shearlet transform of  $f$  can be described as a cascade of a subband decomposition with a directional filtering stage.

## 4.1 3D DShT algorithm

The new numerical algorithm for computing the digital values of the 3D shearlet transform, which is called *3D DShT algorithm*, will follow closely the 3 steps indicated above.

Before describing the numerical algorithm, let us recall that a digital 3D function  $f$  is an element of  $\ell^2(\mathbb{Z}_N^3)$ , where  $N \in \mathbb{N}$ , that is, it consists of a finite array of values  $\{f[n_1, n_2, n_3] : n_1, n_2, n_3 = 0, 1, 2, \dots, N-1\}$ . Here and in the following, we adopt the convention that a bracket  $[\cdot, \cdot, \cdot]$  denotes an array of indices whereas the standard parenthesis  $(\cdot, \cdot, \cdot)$  denotes a function evaluation. Given a 3D discrete function  $f \in \ell^2(\mathbb{Z}_N^3)$ , its Discrete Fourier Transform is given by:

$$\hat{f}[k_1, k_2, k_3] = \frac{1}{N^{\frac{3}{2}}} \sum_{n_1, n_2, n_3=0}^{N-1} f[n_1, n_2, n_3] e^{(-2\pi i(\frac{n_1}{N}k_1 + \frac{n_2}{N}k_2 + \frac{n_3}{N}k_3))}, -\frac{N}{2} \leq k_1, k_2, k_3 < \frac{N}{2}.$$

We shall interpret the numbers  $\hat{f}[k_1, k_2, k_3]$  as samples  $\hat{f}[k_1, k_2, k_3] = \hat{f}(k_1, k_2, k_3)$  from the trigonometric polynomial

$$\hat{f}(\xi_1, \xi_2, \xi_3) = \frac{1}{N^{3/2}} \sum_{n_1, n_2, n_3=0}^{N-1} f[n_1, n_2, n_3] e^{(-2\pi i(\frac{n_1}{N}\xi_1 + \frac{n_2}{N}\xi_2 + \frac{n_3}{N}\xi_3))}.$$

We can now proceed with the description of the implementation of the 3D DShT algorithm.

First, to calculate  $\hat{f}_j(\xi)$  in the digital domain, we perform the computation in the DFT domain as the product of the DFT of  $f$  and the DFT of the filters  $w_j$  corresponding to the bandpass functions  $W(2^{-2j}\cdot)$ . This step can be implemented using the Laplacian pyramid algorithm [20], which results in the decomposition of the input signal  $f \in \ell^2(\mathbb{Z}_N^3)$  into a low-pass and high-pass components. After extensive testing, we found that a very satisfactory performance is achieved using the modified version of the Laplacian pyramid algorithm developed in [14]. For the first level of the decomposition, this algorithm downsamples the low-pass output by a non-integer factor of 1.5 (upsampling by 2 followed by downsampling by 3) along each dimension; the high-pass output is not downsampled. In the subsequent decomposition stages, the low-pass output is downsampled by 2 along each dimension and the high-pass output is not downsampled. Although the fractional sampling factor in the first stage makes the algorithm slightly more redundant than the traditional Laplacian pyramid, it was found that the added redundancy is very useful in reducing the frequency domain aliasing (see [14] for more detail).

Next, one possible approach for computing the directional components  $\hat{f}_{j,\ell,d}$  of  $\hat{f}$  consists in resampling the  $j$ -th subband component of  $f$  into a pseudo-spherical grid and applying a two-dimensional band-pass filter. Even though this is not the approach we will use for our numerical experiments, the method that we will use is conceptually derived from this one.

Recall that the *pseudo-spherical grid* is the 3D extension of the 2D pseudo-polar grid and is parameterized by the planes going through the origin and their slopes. That is, the

pseudo-spherical coordinates  $(u, v, w) \in \mathbb{R}^3$  are given by

$$(u, v, w) = \begin{cases} (\xi_1, \frac{\xi_2}{\xi_1}, \frac{\xi_3}{\xi_1}) & \text{if } (\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_1}, \\ (\xi_2, \frac{\xi_1}{\xi_2}, \frac{\xi_3}{\xi_2}) & \text{if } (\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_2}, \\ (\xi_3, \frac{\xi_1}{\xi_3}, \frac{\xi_2}{\xi_3}) & \text{if } (\xi_1, \xi_2, \xi_3) \in \mathcal{D}_{C_3}. \end{cases}$$

Using this change of variables, it follows that  $\hat{f}_{j,\ell,d}(\xi)$  can be written as

$$\hat{g}_j(u, v, w) U^{(d)}(u, 2^j v - \ell_1, 2^j w - \ell_2), \quad (4.1.3)$$

where  $\hat{g}_j(u, v, w)$  is the function  $\hat{f}_j(\xi)$ , after the change of variables, and  $U^{(d)} = U_{0,0}^{(d)}$ . Notice that  $U^{(d)}$  does not depend on  $u$ . For example, when  $d = 1$ , the expression (4.1.3) can be written as

$$\hat{g}_j(u, v, w) V(2^j v - \ell_1) V(2^j w - \ell_2),$$

showing that the different directional components of  $\hat{f}_j$  are obtained by simply translating the window function  $V$  in the pseudo-spherical domain. In fact, this is a direct consequence of using shearing matrices to control orientations and is its main advantage with respect to rotations. As a result, the discrete samples  $g_j[n_1, n_2, n_3] = g_j(n_1, n_2, n_3)$  are the values of the DFT of  $f_j[n_1, n_2, n_3]$  on the pseudo-spherical grid and they can be computed by direct reassignment or by adapting the pseudo-polar DFT algorithm [21, 22] to the 3D setting. The 3D pseudo-polar DFT evaluates the Fourier transform of the data on the pseudo-polar grid and is formally defined as

$$\begin{aligned} \hat{P}_1(f)(k, l, j) &:= \hat{f}(k, -\frac{2l}{N}k, -2\frac{2j}{N}k), \\ \hat{P}_2(f)(k, l, j) &:= \hat{f}(-\frac{2l}{N}k, k, -2\frac{2j}{N}k), \\ \hat{P}_3(f)(k, l, j) &:= \hat{f}(-\frac{2l}{N}k, -2\frac{2j}{N}k, k), \end{aligned}$$

for  $k = -\frac{2N}{2}, \dots, \frac{2N}{2}$  and  $l, k = -\frac{N}{2}, \dots, \frac{N}{2}$ .

Let  $\{u_{j,\ell_1,\ell_2}^{(d)}[n_2, n_3] : n_2, n_3 \in \mathbb{Z}\}$  be the sequence whose DFT gives the discrete samples of the window functions  $U^{(d)}(2^j v - \ell_1, 2^j w - \ell_2)$ . For example, when  $d = 1$ , we have that  $u_{j,\ell_1,\ell_2}^{(1)}[k_2, k_3] = V(2^j k_2 - \ell_1) V(2^j k_3 - \ell_2)$ . Then, for fixed  $k_1 \in \mathbb{Z}$ , we have

$$\mathcal{F}_2 \left( \mathcal{F}_2^{-1}(\hat{g}_j) * \check{u}_{j,\ell_1,\ell_2}^{(d)}[n_2, n_3] \right) [k_1, k_2, k_3] = \hat{g}_j[k_1, k_2, k_3] u_{j,\ell_1,\ell_2}^{(d)}[k_2, k_3] \quad (4.1.4)$$

where  $\mathcal{F}_2$  is the two dimensional DFT, defined as

$$\mathcal{F}_2(f)[k_2, k_3] = \frac{1}{N} \sum_{n_2, n_3=0}^{N-1} f[n_2, n_3] e^{(-2\pi i(\frac{n_2}{N} k_2 + \frac{n_3}{N} k_3))}, \quad -\frac{N}{2} \leq k_2, k_3 < \frac{N}{2}. \quad (4.1.5)$$

Equation (4.1.4) gives the algorithmic procedure for computing the discrete samples of the right hand side of (4.1.3). That is, the 3D shearlet coefficients (4.0.2) can be calculated from equation (4.1.4) by computing the inverse pseudo-spherical DFT by directly re-assembling the Cartesian sampled values and applying the inverse 3-dimensional DFT.

In fact, the last observation suggests an alternative approach for computing the directional components  $\hat{f}_{j,\ell,d}$  of  $\hat{f}$ . This approach was found to perform better and it was used to produce the numerical results below. The main idea consists in mapping the filters from the pseudo-spherical domain back into the Cartesian domain and then performing a convolution with band-passed data, similar to one of the methods used for the 2D setting in [8]. Specifically if  $\phi_P$  is the mapping from Cartesian domain into the pseudo-spherical domain then the 3D shearlet coefficients in the Fourier domain can be expressed as

$$\phi_P^{-1} \left( \hat{g}_j[k_1, k_2, k_3] u_{j,\ell_1,\ell_2}^{(d)}[k_2, k_3] \right).$$

Following the approach in [8], this can be expressed as

$$\phi_P^{-1}(\hat{g}_j[k_1, k_2, k_3]) \phi_P^{-1} \left( \hat{\delta}_P[k_1, k_2, k_3] u_{j,\ell_1,\ell_2}^{(d)}[k_2, k_3] \right),$$

where  $\hat{\delta}_P$  is the DFT of the (discrete) delta distribution in the pseudo-spherical grid. Thus the 3D discrete shearlet coefficients in the Fourier domain can be expressed as

$$\hat{f}_j[k_1, k_2, k_3] \hat{h}_{j,\ell_1,\ell_2}^{(d)}[k_1, k_2, k_3],$$

where  $\hat{h}_{j,\ell_1,\ell_2}^{(d)}[k_1, k_2, k_3] = \phi_P^{-1} \left( \hat{\delta}_P[k_1, k_2, k_3] u_{j,\ell_1,\ell_2}^{(d)}[k_2, k_3] \right)$ . Notice that the new filters  $h_{j,\ell_1,\ell_2}^{(d)}$  are not obtained by a simple change of variables, but by applying a resampling which converts the pseudo-spherical grid to a Cartesian grid. This resampling is done using a linear map where possibly several points from the polar grid are mapped to the same point on the rectangular grid. Although these filters are not compactly supported, they can be implemented with a matrix representation that is smaller than the size of the data  $f$ , hence allowing to implement the computation of the 3D DShT using a convolution in space domain. One benefit of this approach is that one does not need to resample the DFT of the data into a pseudo-spherical grid, as required using the first method.

Since the computational effort is essentially determined by the FFT which is used to transform data and compute convolutions, it follows that the 3D DShT algorithm runs in  $O(N^3 \log(N))$  operations.

## 4.2 Implementation issues

In principle, for the implementation of the 3D DShT algorithm one can choose any collection of filters  $U_{j,\ell}^{(d)}$  as long as the tiling condition (4.0.1) is satisfied. The simplest solution is to choose functions  $U_{j,\ell}^{(d)}$  which are characteristic functions of appropriate trapezoidal regions in the frequency domain, but this type of filters are poorly localized in space domain. To be faithful to the continuous construction and also to ensure well localized filters in the space domain, our implementation uses filters of Meyer type. A similar choice was also found effective in [8] for the 2D setting. As mentioned above, by taking the inverse DFT, it is possible to implement these filters using matrix representations of size  $L^3$  with  $L \ll N$ , where  $N^3$  is the data size. In the numerical experiments considered below, we have chosen  $L = 24$ , which was found to be a very good compromise between localization

and computation times. Finally, for the number of directional bands, our algorithm allows us to choose a different number of directional bands in each pyramidal region. The theory prescribes to choose a number  $n$  of directional bands which, in each pyramidal region, grows like  $2^{2j}$ , hence giving  $n = 4, 16, 64, \dots$  directional bands, as the scale is becoming finer. As will discuss below, we found it convenient to slightly modify this canonical choice in the video denoising applications.

As a first illustration of the new 3D shearlet decomposition, we have run the 3D DShT algorithm using the *Tempete* video, of size  $192^3$  voxels. Fig. 4.1 shows some representative 2D frames reconstructed from the 3-level 3D DShT decomposition of the *Tempete* video sequence. In particular, the figure shows a frame reconstructed from the approximation levels and some frames reconstructed from some representative directional subbands. The reconstruction from the directional subbands reported in this figure indicates that the shearlet decomposition is very sensitive to directional features.

### 4.3 Correlation with theory

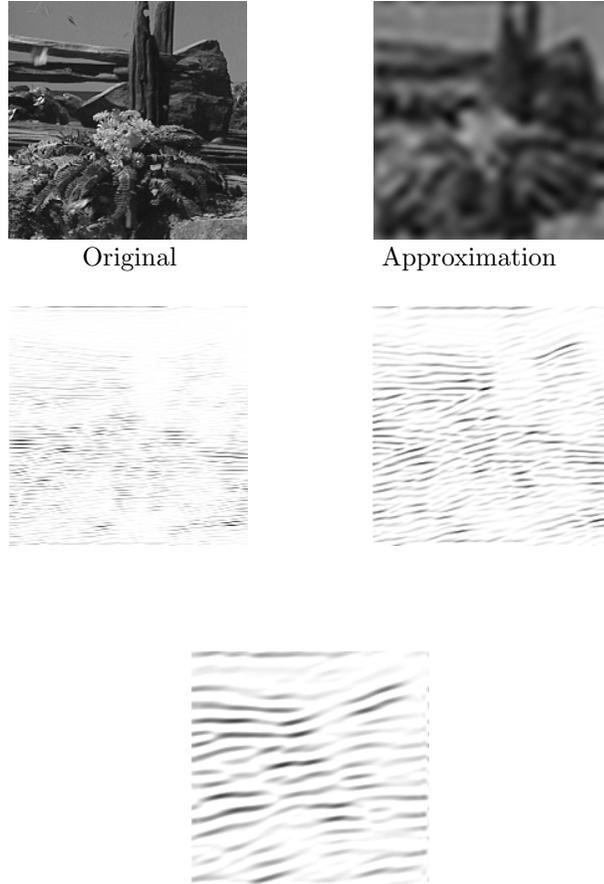
The numerical implementation of the 3D shearlet transform attempts to faithfully reproduce the frequency footprint associated with the 3D shearlet decomposition. Hence, it is natural to ask how this numerical implementation behaves with respect to the theoretical estimate (1.0.1).

To demonstrate that the approximation properties predicted by the theory are reflected in the approximation properties of the digital implementation, we have run some numerical experiments using a piece-wise constant radial function  $f$  with jump discontinuities of the

form

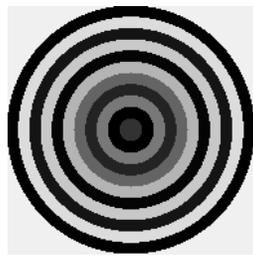
$$f(x, y, z) = c_i \text{ if } r_i \leq x^2 + y^2 + z^2 < r_{i+1} \text{ for given vector } c = (c_i) \in \mathbb{R}^n, r = (r_i) \in \mathbb{R}^{n+1}.$$

For example, by choosing  $r = (1.0, 9.6, 18.3, 26.9, 35.5, 44.2, 52.8, 61.5, 70.1, 78.7, 87.4, 96.0)$ , and  $c = (50, 0, 120, 35, 100, 180, 5, 200, 20, 220, 1, 240)$ , we found that the error  $\|f - f_M^S\|$  decays like  $M^{-0.6192}$  for our test image, as compared to a theoretical rate which is of the order  $(\log M) M^{-0.5}$ . Here  $f_M$  is the nonlinear approximation of  $f$  obtained using the  $M$  largest shearlet coefficient in its 3D DShT expansion. The results of this test are plotted in Fig. 4.2 showing the nonlinear approximation error  $\|f - f_M^S\|$  and comparing this plot to the theoretical curve  $(\log M) M^{-0.5}$ .

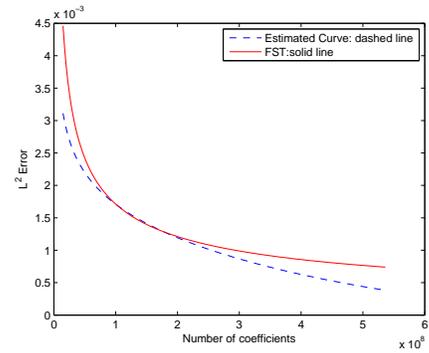


1 level Detail. Band  $(\ell_1 = 4, \ell_2 = 4)$  2 level Detail. Band  $(\ell_1 = 2, \ell_2 = 2)$  3 rd level Detail. Band  $(\ell_1 = 2, \ell_2 = 2)$

Figure 4.1: 3D DShT Decomposition of *Tempete* movie. The figure illustrates some representative 2D frames reconstructed from the 3D DShT decomposition of the movie. All detail frames are extracted from directional subbands contained in the pyramidal region  $D_{C_1}$ . Detail frames, which show highly directional features, are shown in inverted gray scale.



(a)



(b)

Figure 4.2: Analysis of the nonlinear approximation error using the 3D DShT algorithm. (a) Cross section of the piecewise constant radial function  $f$  (on  $\mathbb{R}^3$ ). (b) Approximation error  $\|f - f_M\|_2$ .

# Chapter 5

## Application and Numerical Experiments

As in the 2D setting, the ability of the 3D shearlet transform to deal with geometric information efficiently and its sparsity properties have the potential to produce significant improvement in many 3D data processing applications. As examples of these applications, we have developed algorithms for video denoising, enhancement, and mixed dictionary for denoising which are based on the new 3D Discrete Shearlet Transform presented above.

### 5.1 Video denoising

The denoising of video is highly desirable for enhanced perceptibility, better compression and pattern recognition applications. While noise can have different distributions like Poisson, Laplacian, or Gaussian distribution, we only considered the situation of zero-mean additive white Gaussian noise, which offers a good model for many practical situations. Hence, we assume that, for a given video  $f$ , we observe

$$y = f + n,$$

where  $n$  is Gaussian white noise with zero mean and standard deviation  $\sigma$ .

It is well known that the ability to sparsely represent data is very useful in decorrelating the signal from the noise. This notion has been precisely formalized in the classical *wavelet shrinkage* approach by Donoho and Johnstone [23, 24], which has led to many successful denoising algorithms. In the following, we adapt this idea to design For the choice of the threshold parameter, we adopt the same criterion which was found successful in the 2D setting, based on the classical Bayes Shrink method [25]. This consists in choosing

$$T_{j,\ell} = \frac{\sigma^2}{\sigma_{j,\ell}},$$

where  $\sigma_{j,\ell}$  is the standard deviation of the shearlet coefficients in the  $(j, \ell)$ -th subband. Although hard thresholding is a rather crude form of thresholding and more sophisticated methods are available, still this method is a good indication of the potential of a transform in denoising applications. Also notice that hard thresholding performs better when dealing with data where it is important to preserve edges and sharp discontinuities (cf. [10, 26]).

For the 3D discrete shearlet decomposition, in all our tests we have applied a 3-level decomposition according to the algorithm described above. For the number of directional bands, we have chosen  $n = 16, 16, 64$  (from the coarsest to the finest level) in each of the pyramidal region. Even though this does not exactly respect the rule canonical choice ( $n = 4, 16, 64$ ) prescribed by the continuous model, we found that increasing the number of directional subbands at the coarser level produces some improvement in the denoising performance. Recall that, as indicated above, in our numerical implementation, down-sampling occurs only at the bandpass level, and there is no anisotropic down-sampling. Thus, the numerical implementation of the 3D DShT which we found most effective in the denoising algorithm is highly redundant. Specifically, for data set of size  $N^3$ , a 3-level 3D DShT decomposition produces  $3 * (64 * N^3 + 16 * (\frac{2}{3}N)^3 + 16 * (\frac{2}{6}N)^3) + (\frac{2}{6}N)^3 \approx 208 * N^3$

coefficients. As we will see below (Table II), this requires a higher computational cost than less redundant algorithms.

The 3D shearlet-based thresholding algorithm was tested on 3 video sequences, called *mobile*, *coastguard*, and *tempeste*, for various values of the standard deviation  $\sigma$  of the noise (values  $\sigma = 30, 40, 50$  were considered). All these video sequences, which have been resized to  $192 \times 192 \times 192$ , can be downloaded from the website <http://www.cipr.rpi.edu>. For a baseline comparison, we tested the performance of the shearlet-based denoising algorithm (denoted by 3DSHEAR) against the following state-of-the-art algorithms: 3D Curvelets (denoted by 3DCURV, cf. [13]), Undecimated Discrete Wavelet Transform (denoted by UDWT, based on *symlet* of length 16), Dual Tree Wavelet Transform (denoted by DTWT, cf. [27]) and Surfacelets (denoted by SURF, cf. [14]). We also compared against the 2D discrete shearlet transform (denoted by 2DSHEAR), which was applied frame by frame, in order to illustrate the benefit of using a 3D transform, rather than a 2D transform acting on each frame.

As a performance measure, we used the standard *peak signal-to-noise ratio* (PSNR), measured in decibel(dB), which is defined by

$$\text{PSNR} = 20 \log_{10} \frac{255N}{\|f - \tilde{f}\|_F},$$

where  $\|\cdot\|_F$  is the Frobenius norm and  $f$  is an array of size  $N \times N \times N$ .

The performance of the shearlet-based denoising algorithm 3DSHEAR relative to the other algorithms is shown in Table I, with the numbers in bold indicating the best performance. Notice that performance values for the algorithms 3DCURV, UDWT, and DTWT are taken from [14].

The data in Table I show that the 3D Discrete Shearlet Denoising Algorithm 3DSHEAR

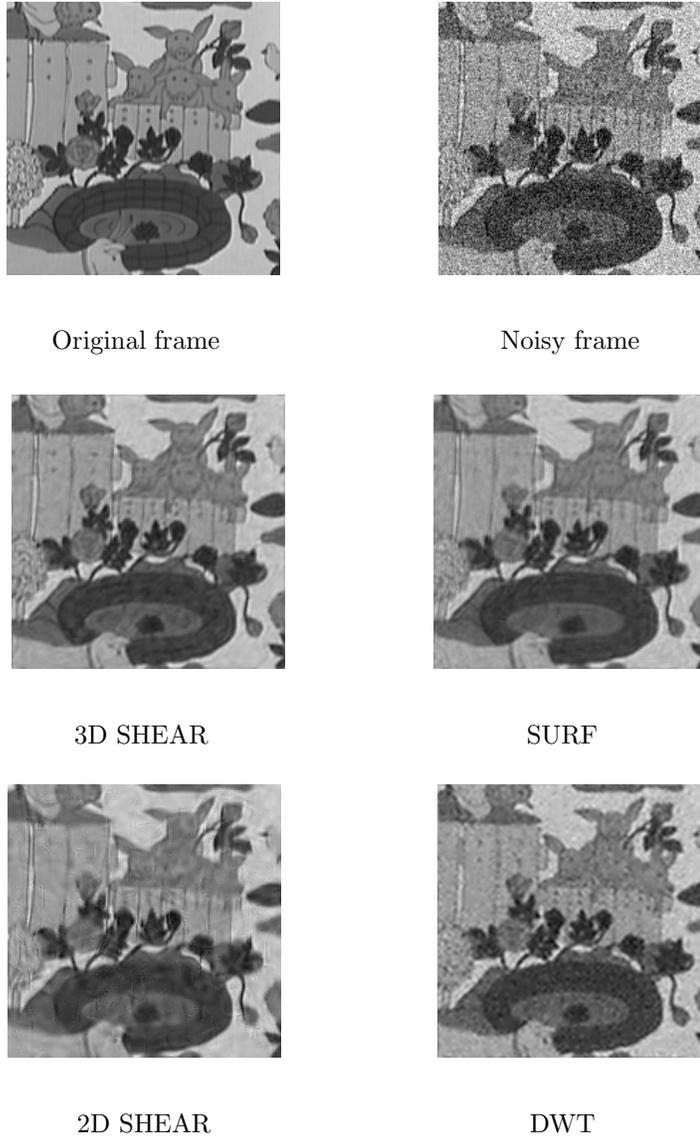


Figure 5.1: Video Denoising of Mobile Video Sequence. The figure compares the denoising performance of the denoising algorithm based on the 3D DShT, denoted as 3DSHEAR, on a representative frame of the video sequence *Mobile* against various video denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB, corresponding to  $\sigma = 30$ ), denoised frame using 3DSHEAR (PSNR=**28.68 dB**), SURF (PSNR=28.39 dB), 2DSHEAR (PSNR=25.97 dB), and DWT (PSNR=24.93 dB).

Table 5.1: Table I: Video denoising performance using different video sequences.

PSNR (dB)	Mobile			Coastguard			Tempete		
Noise $\sigma$	30	40	50	30	40	50	30	40	50
3DCURV	23.54	23.19	22.86	25.05	24.64	24.29			
UDWT	24.02	22.99	22.23	25.95	24.95	24.2			
DTWT	24.56	23.43	22.58	26.06	25.01	24.22			
SURF	28.39	<b>27.18</b>	<b>26.27</b>	26.82	25.87	<b>25.15</b>	24.2	23.26	22.61
3DSHEAR	<b>28.68</b>	27.15	25.97	<b>27.36</b>	<b>26.10</b>	25.12	<b>25.24</b>	<b>23.97</b>	<b>22.81</b>
2DSHEAR	25.97	24.40	23.20	25.20	23.82	22.74	22.89	21.63	20.75
DWT	24.93	23.94	23.03	24.34	23.44	22.57	22.09	21.5	20.92

is highly competitive against both traditional and other state-of-the-art video denoising algorithm. In particular, 3DSHEAR consistently outperforms the curvelet-based routine 3DCURV, the wavelet-based routines UDWT and DTWT and the 2D shearlet-based algorithm. 3DSHEAR also outperforms or is essentially equivalent to the surfacelets-based denoising algorithm in all cases we tested, except for one case, namely the mobile video sequence for low noise, with standard deviation  $\sigma = 50$ . Notice that for higher noise level 3DSHEAR always provide the best performance in all tests that were run.

Table 5.2: Table II: Comparison of running times for different 3D transforms.

Algorithm	Running time (data size: $192^3$ )
SURF	34 sec
3DSHEAR	263 sec
2DSHEAR	154 sec
3D DWT	7.5 sec

The superior performance of the 3DSHEAR algorithm depends in part on its excellent directional selectivity; but it also benefits from the redundancy of the transform, since high redundancy usually produces a better performance in denoising applications. The drawback

is that the higher redundancy requires higher computational effort, which explains the worse performance of 3DSHEAR with respect to 3D DWT and SURF in terms of running times. This is reported in the Table II, which compares the running times for these different 3D transforms, applied to a data set of size  $193^3$ ; all routines were run using the same system which is based on an Intel CPU 2.93GHz.

In Fig. 5.1 and 5.2, we illustrate the performance of the various video denoising routines on a typical frame extracted from the denoised video sequences *Mobile* and *Coast Guard*. Although this type of comparison is more subjective in nature, the figures show that the visual quality of the shearlet-denoised frame is also superior.

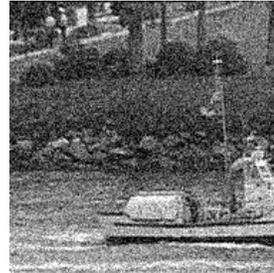
## 5.2 Video enhancement

In several imaging applications, it is important to enhance the visual appearance of certain features that carry useful information. For example, in ultrasound medical images weak edges are usually related to important physical or structural properties so that it is desirable to make weak edges more prominent while keeping the strong features intact. A classical application is mammography, where image enhancement can be useful to improve the visibility of small tumors for early detection [28].

Several techniques have been proposed to enhance the contrast level of an image. For example, since weak edges are mostly associated with the high frequency components of an image, many traditional enhancement methods consist in amplifying the highpass subbands of an image which has been decomposed into different frequency subbands. Unfortunately, these methods are not very efficient to preserve the geometrical features of the data and, as a result, when they are applied to enhance weak edges, they also amplify noise and



Original frame



Noisy frame



3D SHEAR



SURF



2D SHEAR



DWT

Figure 5.2: Video Denoising of Coast Guard Video Sequence. The figure illustrates the denoising performance on a representative frame of the video sequence using various denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB), denoised frame using 3DSHEAR (PSNR=**27.36 dB**), SURF (PSNR=26.82 dB), 2DSHEAR (PSNR=25.20 dB), DWT (PSNR=24.34 dB).

produce visual artifacts. By contrast, multiscale techniques are much more effective in enhancing weak edges without blowing up the noise [28, 29]. The advantage of the shearlet framework, in particular, is to provide a unique ability to control the geometric information associated with multidimensional data. Thus, the shearlet transform appears to be particularly promising as a tool for selectively enhancing the component of the data associated with the weak edges, as was recently observed in [30] for 2D images (see also [31, 32] for other results concerning the application of directional multiscale transforms in image enhancement). In this section, we present an algorithm which extends this approach to the 3D setting and applies the 3D Discrete Shearlet Transform to decompose data into several directional subbands and to selectively amplify some of the shearlet coefficients.

In fact, by the properties of the shearlet decomposition, the shearlet coefficients which are large in magnitude, at fine scales, are closely associated with the singularities of the data. More precisely, strong surfaces of discontinuity will produce large or significant coefficients in all directional subbands, whereas weak surfaces of discontinuity will produce large or significant coefficients only in very few directional subbands. On the other hand, no significantly large coefficients are produced by the noise (provided, of course, SNR is “reasonable”).

Based on these observations, each voxel  $k$  of a data set  $f \in \ell^2(\mathbb{Z}_N^3)$  can be classified into one of three distinct categories by analysing the magnitude of the corresponding shearlet coefficients  $\langle f, \psi_{j,\ell,k}^{(d)} \rangle$ . Notice that heuristic observations have shown that it is sufficient to consider only the shearlet coefficients at the finest scale, so that the parameter  $j$  is fixed in the procedure described below. Hence, in our enhancement algorithm, for each voxel  $k$ , we compute the average and the maximum of the magnitude of the shearlet coefficients taken over all the directional subbands, which we denoted by the functions *mean* and *max*, respectively; next, we compute the *enhanced coefficient* as follows (cf. a similar

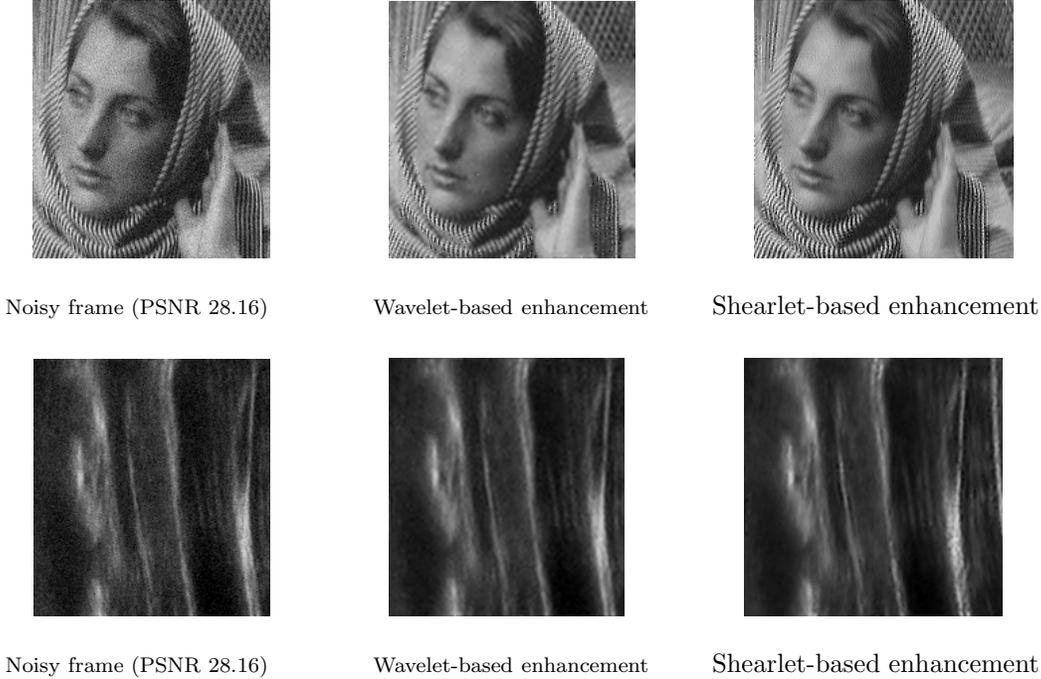


Figure 5.3: Video Enhancement. Representative frames from the Barbara video sequence (above) and from the Anterior ultrasound video sequence (below) illustrate the performance of the shearlet-based enhancement algorithm. This is compared against a similar wavelet-based enhancement algorithm.

enhancement is defined in [33]):

$$E(x) = \begin{cases} x & \text{if } \text{mean}(x) \geq c\sigma, \\ \max\left\{\left(\frac{c\sigma}{|x|}\right)^p, 1\right\} x & \text{if } \text{mean}(x) < c\sigma \ \& \ \max(x) \geq c\sigma, \\ 0 & \text{if } \text{mean}(x) < c\sigma \ \& \ \max(x) < c\sigma, \end{cases}$$

where  $x$  is the input coefficient,  $\sigma$  is the standard deviation of the noise in the subband associated with the finest resolution level,  $0 < p < 1$  is a parameter controlling the portion of edges to be treated as considered “weak edges” and  $c$  is a tuning parameter, determining the enhancement factor ( $c \in [1, 5]$ ).

To illustrate its performance, the shearlet-based enhancement algorithm was tested to

enhance two noisy video sequences, the *Barbara movie*, obtained from the Barbara picture by moving a window frame around the picture) and the *Anterior ultrasound movie*, showing an ultrasound movie sequence of the Anterior triangle, a muscular region near the neck. In both cases, the noise is additive white Gaussian noise with zero mean and standard deviation  $\sigma = 10$ . Also, in both case the enhancement algorithm was run using parameters  $c = 1, p = 1$ , and the performance of the algorithm was compared against an undecimated wavelet-based enhancement routine, which uses the same enhancement function  $E$ . Representative 2D frames from the enhanced video sequences are illustrated in Figure 5.3, showing that the shearlet-based routine performs significantly better both in terms of contrast improvement and noise suppression. For comparison with the shearlet-based enhancement, we also run a similar routine based on the surfacelets, but its performance was not better than the wavelet-based routine. This is due in most part to the fact that the surfacelets algorithm has low redundancy, unlike the shearlet-based and wavelet-based algorithms.

Since the performance of the enhancement algorithm is only partially illustrated by the video frames in Figure 5.3, to better convince the reader, the complete enhancement videos are available at

<http://www.math.uh.edu/~dlabate/software>.

### 5.3 Denoising with mixed dictionary

There is no single representation which is optimal for all the feature in a natural data set. It is well known that wavelets are good to sparsify isolated singularities and that the cosine transform is good for globally oscillating texture. Hence basic denoising via thresholding using a single dictionary is not the right approach for denoising algorithms . In literature

this idea has been improved further by requiring not to use fix dictionary but to learn one which sparsify a set of images known as finding sparse dictionary factorization. In current work we restrict our attention to two fixed dictionaries scenario where data is superposition of two component.

We will start with single dictionary scenario under sparsity assumption before proceeding toward mixed dictionary denoising. To be more precise, given noisy observations

$$y = x + n,$$

where  $n$  is zero-mean white Gaussian noise with variance  $\sigma^2$ . The objective is to estimate  $x$  from observation  $y$ , under the assumption that it has a sparse representation in an over-complete dictionary (typically a frame), i.e.,  $x = \mathcal{F}\alpha$ . In this case we can set up the minimization problem

$$\hat{\alpha} = \min \|\alpha\|_0 \quad \text{subject to } \|y - \mathcal{F}\alpha\|_2 \leq \sigma, \quad (5.3.1)$$

where the  $\ell_0$  norm<sup>1</sup> is counting the number of nonzero entries of  $\alpha = (\alpha_k)$ , that is,  $\|(\alpha_k)\|_0 = \#\{k : \alpha_k \neq 0\}$ . Since the algorithmic solution of (5.3.1) is NP-hard, this problem is usually modified by relaxing  $\ell_0$  to an  $\ell_1$ -norm, hence defining the *basis pursuit denoising* (BPDN) problem [62]

$$\hat{\alpha} = \min \|\alpha\|_1 \quad \text{subject to } \|y - \mathcal{F}\alpha\|_2 \leq \sigma, \quad (5.3.2)$$

and leading to the estimator  $\hat{x} = \mathcal{F}\hat{\alpha}$ . Due to the convexity of the  $\ell_1$  norm, there are many ways to solve this problem rather efficiently including interior point methods and gradient projections (cf. [63]).

When the dictionary is redundant, then the solution of minimization problem (5.3.2) is not necessarily the solution of the classical shrinkage method. Yet, also in this case,

---

<sup>1</sup>This is not technically a norm, but this abuse of notation is customary in the literature.

appropriate iterative-shrinkage algorithms have been introduced which extend the classical Donoho-Johnstone wavelet shrinkage method, starting with the algorithm introduced by Starck, Murtagh and Bijaoui in 1995 [64] and including the celebrated algorithm of Daubechies, Defrise, and De Mol [56].

An interesting variant of problem (5.3.2) is the situation where the data to be recovered are known to be a superposition of several components, each one having a sparse representation with respect to a certain dictionary. In many situations the data may contain textured components along with piecewise smooth components. Hence, we can model  $x$  as a superposition

$$x = x_p + x_t,$$

where  $x_p$  and  $x_t$  are the piecewise smooth component and textured component of a video, respectively. In this situation it was shown that taking a mix dictionary approach to denoise generally result in improved estimates with respect to the standard thresholding approach. This approach of seperating signal into different components is known as Morphological Component Analysis(MCA) [59]. Following the idea used in the 2D setting, discrete cosine (DCT) representation is used for the texture component of  $x$  and a shearlet representation for the piecewise smooth component of the data. Let  $x_p = \mathcal{SH}\alpha_p$  in shearlet dictionary  $\mathcal{SH}$  and  $x_t = \mathcal{D}\alpha_t$  in DCT dictionary  $\mathcal{D}$ . Hence  $x$  can be estimated by solving the following optimization problem:

$$\hat{\alpha}_p, \hat{\alpha}_t = \arg \min_{\alpha_p, \alpha_t} \lambda \|\alpha_p\|_1 + \lambda \|\alpha_t\|_1 + \gamma TV(\alpha_p) + \frac{1}{2} \|y - \mathcal{SH}\alpha_p - \mathcal{D}\alpha_t\|_2^2, \quad (5.3.3)$$

where  $\mathcal{SH}^T$  and  $\mathcal{D}^T$  denote the Moore-Penrose pseudo inverse of  $\mathcal{SH}$  and  $\mathcal{D}$ , respectively, and  $TV$  is the Total Variation. It is well known that shrinkage based algorithm used sparsity of representation for denoising. Hence a fast iterative shrinkage algorithm known as Separable Surrogate Functionals(SSF) [56–58] is used to solve 5.3.3. Once the denoised

components are obtained,  $x$  is estimated as  $\hat{x} = \mathcal{S}\mathcal{H}\hat{\alpha}_p + \mathcal{D}\hat{\alpha}_t$ .

### 5.3.1 Iterative shrinkage algorithm

Let  $\mathcal{S}_\lambda(x) = \text{sign}(x)(|x| - \lambda)_+$  be the element-wise soft thresholding operator and  $H$  be the undecimated Haar wavelet dictionary. Then the solution of optimization problem 5.3.3 is given by following SSF algorithm 5.4. TV correction term is replaced by the Haar Transform as done in 2D case [58]. This is applied to smooth component to control the ringing effect near the edge caused by oscillation of atoms in dictionary  $\mathcal{S}\mathcal{H}$ . Similar adjustment was used in [59] and was motivated by observing connection between TV and the Haar Wavelet as given in [60].

**Input:**  $x$   
**Initialization:** Initialize  $k = 1, \alpha_p^0 = 0, \alpha_t^0 = 0$  and  $r^0 = x - \mathcal{S}\mathcal{H}\alpha_p^0 - \mathcal{D}\alpha_t^0$   
**Do:**  
 1. Update estimate as

$$\begin{aligned}\tilde{\alpha}_p^k &= \mathcal{S}_\lambda\left(\frac{1}{c}\mathcal{S}\mathcal{H}^T(r^{k-1}) + \alpha_p^{k-1}\right) \\ \alpha_p^k &= \mathcal{S}\mathcal{H}^T H S_\gamma(H^T \mathcal{S}\mathcal{H}_p \tilde{\alpha}_p^k) \\ \alpha_t^k &= \mathcal{S}_\lambda\left(\frac{1}{c}\mathcal{D}_t^T(r^{k-1} + \alpha_t^{k-1})\right)\end{aligned}$$

2. Update the residual  
 $r^k = x - \mathcal{S}\mathcal{H}\alpha_p^k - \mathcal{D}\alpha_t^k$   
**Until:**  $\lambda^k = 2.1\sigma$ .  
**Output:**  $\hat{\alpha}_p = \alpha_p^k$  and  $\hat{\alpha}_t = \alpha_t^k$

Figure 5.4: Separable Surrogate Functional (SSF) iterative shrinkage algorithm to solve 5.3.3

### 5.3.2 Experiments

This section contains several numerical experiments on video denoising, based on the approach described in the previous section. The shearlet dictionary is used for smooth piecewise part and DCT for texture part. Parameter  $\gamma$  was tuned to produce lowest error. The thresholding parameter  $\lambda$  decays like  $\lambda = \lambda^k$  during each iteration and stop the iteration when  $\lambda^k = T\sigma$ ,  $T \approx 2.1$  [61]. Thus we stop when the residual is at noise level. As the stop criterion require knowlege of  $\sigma$ , one can use the median estimator on finest scale wavelet coefficient.

In the first experiment we use Tempete video and denoise it using the standard thresholding rule for different algorithm as well as their combination with DCT. Although this movie is not so rich in a texture component but still amalgamated Shearlet and DCT dictionary, it performs very good compared to any other dictionary combination as shown in Table 5.3. Figure 5.5 contains some denosied Tempete movie frames in a different scenario.

Table 5.3: Table III: Mix Dictionary Denoising results (PSNR) using *Tempete video*.

$\sigma$	Noisy	DWT	LP	shear	curv	shear/DCT	DWT/DCT	LP/DCT	curv/DCT
20	22.14	22.61	23.10	25.87	22.60	27.47	24.09	24.45	25.29
30	18.62	21.10	22.04	24.63	22.27	25.61	22.38	22.61	23.02
40	16.12	20.47	21.30	23.69	22.00	24.34	21.40	21.51	21.97

In the second experiment the oil painting video is used, which is rich in texture component as well as in a piece-wise smooth component. Clearly in this scenario Shearlet and DCT perform very good compared to other competing approaches like Curvelet and DCT, as shown in Table 5.4. Figure 5.6 shows typical extracted frame of denoised oil painting.

### 5.3. DENOISING WITH MIXED DICTIONARY

---

Table 5.4: Table IV: Mix Dictionary Denoising results (PSNR) using *Oil painting video*.

$\sigma$	Noisy	DWT	LP	shear	curv	shear/DCT	DWT/DCT	LP/DCT	curv/DCT
20	22.14	26.34	27.01	28.04	27.32	31.01	27.74	28.32	27.66
30	18.62	24.81	25.52	27.12	26.86	29.07	26.03	26.37	25.94
40	16.12	23.87	24.26	26.33	26.44	27.68	24.89	25.02	24.67

All the routines were running on a cluster of 12 Intel(R) Xeon(R) CPUS at 2.93GHz. Most of the code was written using matlab except the 3D curvelet interface which used mex files for calling C language based routines. Table 5.5 shows running time for a combination of different dictionaries.

Table 5.5: Table V: Comparison of running times for different routines.

Algorithm	Running time (data size: $192^3$ )
DWT	9 sec
DWT/DCT	208 sec
LP	4 sec
LP/DCT	308 sec
Curv	23 sec
Curv/DCT	425 sec
Shear	664 sec
Shear/DCT	10769 sec



Original frame



Noisy frame



DWT



DWT/DCT



LP



LP/DCT

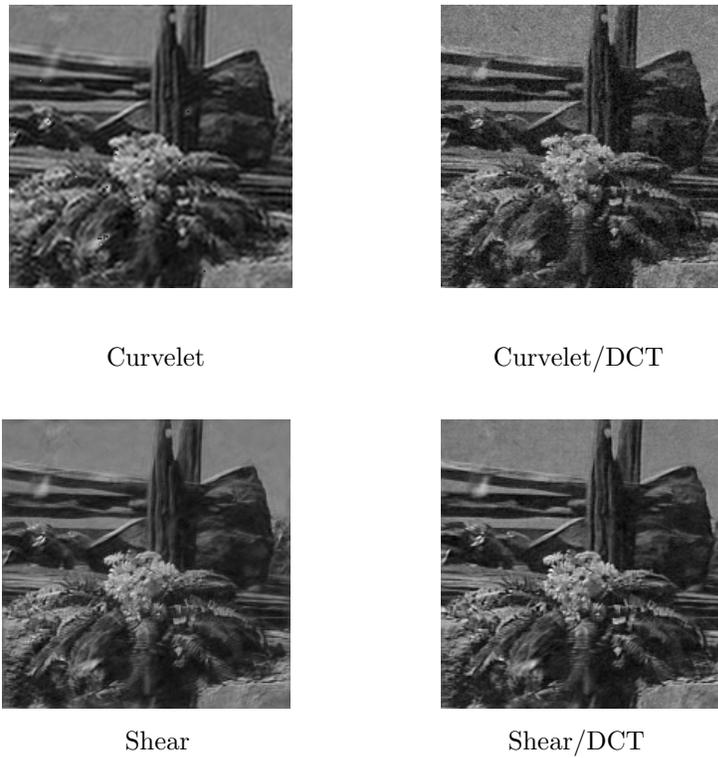


Figure 5.5: Video Denoising of Tempete Video Sequence. The figure compares the denoising performance of the denoising algorithm based on the 3D DShT, denoted as Shear, on a representative frame of the video sequence *Tempete* against various video denoising routines. Starting from the top left: original frame, noisy frame (PSNR=22.14 dB, corresponding to  $\sigma = 20$ ), denoised frame using DWT (PSNR= 22.16 dB), DWT/DCT (PSNR=24.09 dB), LP (PSNR=23.10 dB), LP/DCT (PSNR=24.45 dB), Shear (PSNR=25.87 dB), Shear/DCT (PSNR=**27.47 dB**), Curvelet (PSNR= 22.60 dB) and Curvelet/DCT (PSNR=25.29 dB).



Original frame



Noisy frame



DWT



DWT/DCT



LP



LP/DCT



Curvelet



Curvelet/DCT



Shear



Shear/DCT

Figure 5.6: Video Denoising of Oil Painting Video Sequence. The figure compares the denoising performance of the denoising algorithm based on the 3D DShT, denoted as Shear, on a representative frame of the video sequence *Oil Painting* against various video denoising routines. Starting from the top left: original frame, noisy frame (PSNR=18.62 dB, corresponding to  $\sigma = 30$ ), denoised frame using DWT (PSNR= 24.81 dB), DWT/DCT (PSNR=26.03 dB), LP (PSNR=25.52 dB), LP/DCT (PSNR=26.37 dB), Shear (PSNR=27.12 dB), Shear/DCT (PSNR=**29.07 dB**), Curvelet (PSNR= 26.86 dB) and Curvelet/DCT (PSNR=25.94 dB).

# Bibliography

- [1] K. Guo, G. Kutyniok, and D. Labate, Sparse multidimensional representations using anisotropic dilation and shear operators, in: *Wavelets and Splines*, G. Chen, M. Lai (eds.), Nashboro Press, 2006, pp. 189-201.
- [2] K. Guo, W.-Q Lim, D. Labate, G. Weiss and E. Wilson, Wavelets with composite dilations, *Electron. Res. Announc. Amer. Math. Soc.*, vol. 10, pp. 78–87, 2004.
- [3] E. J. Candès and D. L. Donoho, New tight frames of curvelets and optimal representations of objects with  $C^2$  singularities, *Comm. Pure Appl. Math.*, vol. 57, no 2, pp. 219-266, 2004.
- [4] M. N. Do and M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091-2106, 2005.
- [5] K. Guo, D. Labate, Optimally sparse multidimensional representation using shearlets, *SIAM J. Math. Anal.*, vol 39, no. 1, pp. 298-318, 2007.
- [6] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, Data compression and harmonic analysis, *IEEE Trans. Inform. Th.*, vol. 44, pp. 2435–2476, 1998.
- [7] G. R. Easley, D. Labate, F. Colonna, Shearlet-based total variation diffusion for denoising, *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 260-268, 2009.
- [8] G. R. Easley, D. Labate, and W. Lim, Sparse directional image representations using the discrete shearlet transform, *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 25-46, 2008.
- [9] S. Yi, D. Labate, G. R. Easley, and H. Krim, A shearlet approach to edge analysis and detection, *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 929-941, 2009.

- [10] F. Colonna, G. Easley, K. Guo, and D. Labate, Radon transform inversion using the shearlet representation, *Appl. Comput. Harmon. Anal.*, vol. 29, no. 2, pp. 232-250, 2010.
- [11] P. Kittipoom, G. Kutyniok, and W.-Q Lim, Construction of compactly supported shearlet frames, *Constr. Approx.*, vol. 35, pp. 21-72, 2012.
- [12] G. Kutyniok and W.-Q Lim, Compactly supported shearlets are optimally sparse, *J. Approx. Theory*, vol. 63, pp. 1564-1589, 2011.
- [13] E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, Fast discrete curvelet transforms, *SIAM Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861-899, 2006.
- [14] Y. Lu and M. N. Do, Multidimensional directional filter banks and surfacelets, *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 918-931, 2007.
- [15] K. Guo, and D. Labate, Optimally sparse 3D approximations using shearlet representations, *Electron. Res. Announc. Amer. Math. Soc.*, vol. 17, pp. 126-138, 2010.
- [16] K. Guo, and D. Labate, Optimally sparse representations of 3D data with  $C^2$  surface singularities using Parseval frames of shearlets, preprint 2010.
- [17] E. L. Pennec and S. Mallat, Sparse geometric image representations with bandelets, *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 423-438, 2005.
- [18] S. Mallat, Geometrical grouplets, *Appl. Comput. Harmon. Anal.*, vol. 26, no. 2, pp. 161-180, 2009.
- [19] K. Guo, and D. Labate, The Construction of smooth Parseval frames of shearlets, preprint 2011.
- [20] P. J. Burt, E. H. Adelson, The Laplacian pyramid as a compact image code, *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532-540, 1983.
- [21] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Israeli, and Y. Shkolnisky, A framework for discrete integral transformations I - The pseudo-polar Fourier transform, *SIAM Journal on Scientific Computing*, vol. 30, no. 2, pp. 764-784, 2008.
- [22] Y. Keller, A. Averbuch, Y. Shkolnisky, Algebraically accurate volume registration using Euler's theorem and the 3D pseudopolar FFT, *IEEE Comput. Soc. Conf. on Comput. Vision and Pattern Rec.*, vol. 2, pp. 795-800, 2005
- [23] D. L. Donoho and I. M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, *Biometrika* vol. 81, no. 3, pp. 425-455, 1994.
- [24] D. L. Donoho, Denoising via soft thresholding, *IEEE Trans. Inform. Th.*, vol. 41, no. 3, pp. 613-627, 1995.

- [25] G. Chang, B. Yu and M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [26] J. L. Starck, F. Murtagh, J. Fadili, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*, Cambridge U. Press, 2010.
- [27] I. W. Selesnick, The double-density dual-tree DWT, *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1304-1314, 2004.
- [28] C. Chang, A. F. Laine, Coherence of multiscale features for contrast enhancement of digital mammograms, *IEEE Trans. Info. Tech. in Biomedicine*, vol. 3, no. 1, pp. 32-46, 1999.
- [29] S. Dippel, M. Stahl, R. Wiemker, and T. Blaffert, Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform, *IEEE Trans. Medical Imaging*, vol. 21, no. 4, pp. 343-353, 2002.
- [30] V. M. Patel, G. R. Easley, and D. M. Healy, A new multiresolution generalized directional filter bank design and application in image enhancement, *Proc. IEEE Int. Conf. on Image Process.*, San Diego, pp. 2816-2819, 2008.
- [31] J. L. Starck, F. Murtagh, E. J. Candès, and D. L. Donoho, Gray and color image contrast enhancement by the curvelet transform, *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 706-717, 2003.
- [32] J. Zhou, A. L. Cunha, and M. N. Do, Non-subsampled Contourlet transform: construction and application in Enhancement, *IEEE Int. Conf. on Image Process.*, Genoa, 2005
- [33] K. V. Velde, Multi-scale color image enhancement, *Proc. IEEE Int. Conf. on Image Process.*, vol. 3, pp. 584-587, 1999.
- [34] L. Borup and M. Nielsen. Nonlinear approximation in  $\alpha$ -modulation spaces. *Math. Nachr.*, vol 279, pp 101-120, 2006.
- [35] L. Borup and M. Nielsen. Frame decomposition of decomposition spaces. *J. Fourier Anal. Appl.*, vol. 13, pp. 39-70, 2007.
- [36] P. G. Casazza. The art of frame theory. *Taiwanese J. Math.*, vol. 4, pp. 129-201, 2000.
- [37] C. Chesneau, M.J. Fadili and J.-L. Starck. Stein block thresholding for image denoising. *Appl. Comput. Harmon. Anal.*, vol. 28, no.1, pp. 67-88, 2010.
- [38] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhäuser, Boston, 2003.
- [39] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark and G. Teschke. The uncertainty principle associated with the continuous shearlet transform. *Int. J. Wavelets Multiresolut. Inf. Process.*, vol. 6, pp. 157-181, 2008.

- [40] S. Dahlke, G. Steidl and G. Teschke. Multivariate shearlet transform, shearlet coorbit spaces and their structural properties. In: *Shearlets: Multiscale Analysis for Multivariate Data*. G. Kutyniok and D. Labate, Birkhäuser, Boston, 2012.
- [41] H. G. Feichtinger. Banach spaces of distributions defined by decomposition methods, II. *Math. Nachr.*, vol. 132, pp. 207-237, 1987.
- [42] H. G. Feichtinger and P. Gröbner. Banach spaces of distributions defined by decomposition methods, I. *Math. Nachr.*, vol. 123, pp. 97-120, 1985.
- [43] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decomposition I. *J. Funct. Anal.*, vol. 86, pp. 307-340, 1989.
- [44] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decomposition II. *Monatsh. Math.*, vol. 108, pp. 129-148, 1989.
- [45] K. Gröchenig and S. Samarah. Nonlinear approximation with local Fourier bases. *Constr. Approx.*, vol. 16, no. 3, pp. 317-331, 2000.
- [46] K. Guo and D. Labate. Representation of Fourier integral operators using shearlets. *J. Fourier Anal. Appl.*, vol. 14, pp. 327-371, 2008.
- [47] G. Kutyniok. Sparsity equivalence of anisotropic decompositions. (preprint) 2012.
- [48] G. Kutyniok and D. Labate. Resolution of the wavefront set using continuous shearlets. *Trans. Amer. Math. Soc.*, vol. 361, pp. 2719-2754, 2009.
- [49] G. Kutyniok and D. Labate. Shearlets: Multiscale analysis for multivariate data. Birkhäuser, Boston, 2012.
- [50] Y. Meyer. Wavelets and operators. Cambridge Stud. Adv. Math. vol. 37, Cambridge Univ. Press, Cambridge, UK, 1992.
- [51] V. M. Patel, G. Easley and D.M. Healy. Shearlet-based deconvolution. *IEEE Transactions on Image Processing*, vol. 18, pp. 2673-2685, 2009.
- [52] K. N. Rasmussen and M. Nielsen. Compactly supported frames for decomposition spaces. *J. Fourier Anal. Appl.*, vol. 18, pp. 87-117, 2012.
- [53] S. Dahlke, G. Kutyniok, G. Steidl and G. Teschke. Shearlet coorbit spaces and associated Banach frames. *Appl. Comput. Harmon. Anal.*, vol. 27, no. 2, pp. 195-214, 2009.
- [54] S. Dahlke, G. Steidl and G. Teschke. Shearlet coorbit spaces: compactly supported analyzing shearlets, traces and embeddings. *J. Fourier Anal. Appl.*, vol. 17, no. 6, pp. 1232-1255, 2011.

- [55] H. Triebel. Theory of function spaces. Birkhäuser, Basel, 1983.
- [56] I. Daubechies, M. Defrise, C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, Vol. 57, No. 11. 1413-1457, 2004
- [57] Michael Zibulevsky, Michael Elad. L1-L2 Optimization in Signal and Image Processing. *IEEE Signal Processing Magazine.*, Vol. 27, No. 3, 76-88 , may 2010
- [58] V. M. Patel, G. R. Easley, and R. Chellappa. Component-based restoration of speckled images. *IEEE International Conference on Image Processing.*, Brussels, Belgium, 2011
- [59] J.-L. Starck, M. Elad, and D. L. Donoho. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Transactions on Image Processing.*, Vol. 14, No. 10, 15701582, 2005
- [60] G. Steidl, J. Weickert, T. Brox, P. Mrazek, and M. Welk, On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides, *SIAM J. Numer. Anal.*, vol. 42, 686713, 2004.
- [61] J. Bobin, J.-L. Starck, J. Fadili, Y. Moudden, and D. L. Donoho, Morphological component analysis: An adaptive thresholding strategy, *IEEE Transactions on Image Processing*, vol. 16, no. 11, 26752681, 2007.
- [62] S. S. Chen, D. L. Donoho, and M. A. Saunders, Atomic decomposition by basis pursuit, *SIAM Rev.*, vol. 43, no. 1, pp. 129-159, 2001.
- [63] M. Elad. *Sparse and redundant representations: From theory to applications in signal and image Processing*, Springer, New York, NY, 2010.
- [64] J. L. Starck, F. Murtagh, A. Bijaoui. Multiresolution support applied to image filtering and restoration, *Graphic. Models Image Process.*, vol. 57, pp. 420-431, 1995.
- [65] D. Labate, L. Mantovani, and P. S. Negi, Shearlet smoothness spaces, submitted 2012
- [66] D. Labate and P. S. Negi, 3D discrete shearlet transform and video denoising, *Wavelets XIV (San Diego, CA)*, SPIE Proc. 8138, 2011.
- [67] P. S. Negi and D. Labate, 3D discrete shearlet transform and video processing, *IEEE Transaction on Image Processing.*, vol. 21, no. 6, pp. 2944-2954, 2012.
- [68] G. Easley, D. Labate, and P. S. Negi, 3D data denoising using combined sparse dictionaries, submitted 2012
- [69] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998, 1999, Elsevier.

*BIBLIOGRAPHY*

---

- [70] G. Garrigós and E. Hernández. Sharp Jackson and Bernstein inequalities for N-term approximation in sequence spaces with applications, *Indiana Univ. Math. J.*, vol. 53, no. 6, pp. 1739-1762, 2004.