# Continuous Parabolic Molecules

P. Grohs and Z. Kereta

Research Report No. 2015-17
June 2015

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

# Continuous Parabolic Molecules

Philipp Grohs*[1] and Željko Kereta†[1]

[1]*ETH-Zürich.*

**Abstract**

Decomposition systems based on parabolic scaling have in the last years garnered attention for their ability to answer questions regarding curvilinear singularities of functions. Well known examples of these systems are curvelets and shearlets. In recent years there has been a sufficient body of evidence to suggest that these systems are able to answer the same fundamental questions and it should thus be possible to consider them as parts of a broader framework. Thus far each such system required proofs of their properties that are tailored to their specific constructions, which is a predicament that can be avoided by focusing on the fundamental features they share.

Another incentive is that while these systems exhibit same or similar properties, the specifics of their constructions might make a difference. For example, some systems are good for theoretical considerations whereas other systems might be better suited for implementations.

In this paper we will construct a framework for parabolic molecules in the continuous setting, and show that it is wide enough to contain both the curvelet- and shearlet-type systems. Using almost-orthogonality we will show that some results of note (resolution of the wavefront set, microlocal Sobolev regularity) are universal for all suitable continuous parabolic molecules. The main tool we will use is that molecules are *almost-orthogonal* in a certain sense.

*Email: `philipp.grohs@sam.math.ethz.ch`
†Email: `zeljko.kereta@math.ethz.ch`

# 1 Introduction

## 1.1 Parabolic molecules and previous work

For a long time now wavelets have been the go-to transform in applied harmonic analysis, combining powerful features with a wide spectrum of applications. Recently however, while still widely used, it became apparent that wavelets come with limitations which are unavoidable and cannot be circumvented. These limitations are apparent in, for example, image processing. This stems from the fact that edges, a fundamental feature of natural images, are by and large anistropic constructs. Therefore, since wavelets are inherently isotropic objects it should come as no surprise that they are not perfectly adequate to deal with natural images. While there have been a number of attempts to salvage this situation by adjusting wavelets, the real breakthrough came with the advent of curvelets in 2004 [1]. Curvelets were the first system fully adapted for dealing with anisotropic phenomena, providing optimally sparse approximations for bivariate functions. As opposed to wavelets, curvelets are defined not only for a range of locations and scales, but also orientations. A number of other directional representation systems have since been introduced. Among those are contourlets and shearlets, all addressing the various questions posed in the multivariate setting in a unique way.

Up until recently, if we wanted to establish that a given system exhibits a certain feature, such as with regards to the resolution of the wavefront set, we had to produce a proof specific to that system. These proofs all follow along very similar lines. Furthermore, since these decomposition systems exhibit equal or similar approximation properties it seems reasonable to assume that they could be seen as parts of some general framework. Another reason why having a general framework might be heplful is that some systems (for example curvelets) are better suited to address theoretical questions, while other systems (for example shearlets) are better for implementations. Therefore, we would prefer to do the *proofs* in systems for which they would be easiest to procure, and then use the means and tools of our framework to infer that same properties also hold for all other systems of parabolic molecules.

The notion of molecules associated with anisotropic scaling came in the work of Candés and Demanet [2], where they used curvelet molecules to deal with wave propagators. Some recent papers, for example [3], introduced the notion of parabolic molecules in an attempt to unite the existing discrete transforms based on parabolic scaling. There the authors showed that curvelet and shearlet-like constructions are both members of this class. More importantly, it was shown that we can control the Gramian of two systems of parabolic molecules, that is, that the Gramian exhibits strong off-diagonal decay. From there it is possible to make inferences on the various properties these systems share.

The framework of molecules for directional representation systems we are about to introduce uses *parabolic* as its keyword. This reflects the fact that all interesting systems (ones that provide optimally sparse representations of cartoon images) obey a law of parabolic scaling, which is a type of scaling leaving the parabola $y = x^2$ invariant.

## 1.2 Results and Contributions

The goal of this paper is to further the ethos of parabolic molecules and extend it to the continuous setting. In other words, we will introduce a framework for decomposition-reconstruction systems that are defined in the analogue domain. This framework should allow for an easy transfer of various results from one system to another, without having to know specific details regardingh each individual system, but rather adhering to underlying properties they exhibit.

In order for all of this to work we first ought to establish which systems fit into our framework. Then we will show that a certain result is universal for all such systems (assuming it holds for any one specific system) by using our main tool, almost orthogonality of parabolic molecules. In this paper we shall show that we can apply this procedure to curvelets and shearlets. Hence, it will follow that these two systems are equally suited to address questions regarding the resolution of the wavefront set and microlocal Sobolev regularity.

## 1.3 Contents

We begin in Section 2 with a formal definition of Continuous Parabolic Molecules (CPMs) and other related concepts. Section 3 is here to show that analysing systems of interest based on parabolic scaling, namely curvelets and shearlets, are CPMs and that their parametrisations are admissible. In Section 4 we will prove the quintessential result of this paper; almost orthogonality of two families of continuous parabolic molecules. In Section 5 we will show some select applications and consequences of almost orthogonality concerning microlocal Sobolev regularity and the resolution of the wavefront set. Notably, the $L^2$ condition regarding microlocal Sobolev regularity has previously been shown to hold only for curvelets, and it is in this paper extended to shearlets as well. In the appendix we will try to elucidate some, mostly technical, elements of various proofs in the paper, and we will also include a construction of a new shearlet family which admits a useful representation formula.

## 1.4 Notation

We denote by $L^p(\mathbb{R}^d)$ the Lebesgue space with the usual norm $\|\cdot\|_p$. The Fourier transform of an $L^1(\mathbb{R}^d)$ function $f$ is defined as

$$\hat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} f(\mathbf{x}) e^{-2\pi\imath \mathbf{x}\cdot\boldsymbol{\xi}} d\mathbf{x}$$

This definition can by density be extended to tempered distributions, using the standard arguments.

Chevrons are used for two purposes, depending on the number of arguments. If there is only one argument then $\langle x \rangle = (1+x^2)^{1/2}$. Otherwise, if there are two arguments then $\langle \cdot, \cdot \rangle$ will denote the inner product in a given Hilbert space. We use $A \lesssim B$ to indicate that $A \leqslant CB$ with a uniform constant $C$.

Throughout this paper we will work in $\mathbb{R}^2$, with a spatial variable $\mathbf{x}$ and a frequency domain variable $\boldsymbol{\xi}$. When we will be talking about parametrisations, we will use $\mathbf{b}$ to denote a location parameter in $\mathbb{R}^2$, to distinguish it from an $\mathbf{x}$ which is a general element of $\mathbb{R}^2$, not associated with CPMs or their parametrisations. Letter $a$ will denote the scaling parameter and $\theta$ will be reserved for angles. Norm of a vector $\mathbf{x}$ will be denoted by $|\mathbf{x}|$, which is notation we will also use to denote the absolute value of a real number.

# 2 Continuous Parabolic Molecules

Continuous parabolic molecules are continuous analogues of discrete parabolic molecules, which were introduced in [3]. Roughly speaking, CPMs are families of $L^2(\mathbb{R}^2)$ functions whose members obey certain smoothness and decay conditions, and are associated with a unique scale, position and angle.

Let us start by setting up some notation and definitions. We define the parameter space as

$$\mathbb{P} = \mathbb{R} \times [0, 2\pi) \times \mathbb{R}^2,$$

where a point $p = (a, \theta, \mathbf{b}) \in \mathbb{P}$ describes a scale $a$, an orientation $\theta$, and a location $\mathbf{b}$.

Let $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ be a rotation matrix associated with an angle $\theta$, and let $D_a = \operatorname{diag}(a, a^{1/2})$ be the (anisotropic) scaling matrix, with a scaling parameter $a \in \mathbb{R}^+$. Alternatively we could replace $a^{1/2}$ with $a^\alpha$, where $\alpha \in [0, 1)$. Such constructions have been considered in [4]. Still, parabolic scaling plays a fundamental role in the analysis and seems to be the best choice.

Members of a CPM family are associated with a scale, orientation, location triplet through a parametrisation, which is, loosely speaking, a subset of the parameter space $\mathbb{P}$.

**Definition 2.1.** A parametrisation is a pair $(\Lambda, \Phi)$ where $\Lambda$ is an index set and $\Phi$ is a mapping $\Phi \colon \Lambda \to \mathbb{P}$. A parametrisation family is a family of parametrisations $(\Lambda_i, \Phi_i)_{i \in I}$, where $I$ is an index set.

The set $\Lambda$ serves to index the members of a given family of functions, while $\Phi$ then associates those indices to a specific scale, orientation, location triplet in $\mathbb{P}$. Definition 2.1 introduces the possibility of having a family of parametrisations, instead of having just one parametrisation. The idea here is that a family of functions might be composed of several parts, each part for example dealing with a different part of the frequency domain. One example would be the cone-adapted shearlets where the frequency domain is split up in 4 cones, and the shearlets are then defined separately on the vertical and on the horizontal cones. This allows for the parametrisation functions $\Phi$ to have better properties. In addition to that, we ought to also have a special parametrisation dealing with low frequency regimes, since its construction most often does not go by the same rules as that of the high-frequency regimes. Since this would only serve to further complicate the notation, make the proofs lengthier with no real conceptual changes, and since most of the systems we look at address these low frequency regimes in pretty much the same way, we will not give the low frequency regimes a special treatment in the proofs, nor the statements, of our claims.

Having said that, in the rest of the paper we will only be concerned with the case $|I| = 1$. This is because all of our proofs are to do with inequalities and establishing bounds on norms. Hence, provided the index set $I$ is finite, the general case follows from the case $|I| = 1$ by using the triangle inequality.

We are now ready to define continuous parabolic molecules.

**Definition 2.2.** Let $(\Lambda, \Phi)$ be a parametrisation. A family of functions $\{m_\lambda : \lambda \in \Lambda\}$ is called a family of *continuous parabolic molecules* of order $(R, M, N_1, N_2)$ if it can be written as

$$m_\lambda(\mathbf{x}) = a_\lambda^{-3/4}\varphi^{(\lambda)}\left(D_{1/a_\lambda}R_{\theta_\lambda}(\mathbf{x}-\mathbf{b}_\lambda)\right),$$

where $(a_\lambda, \theta_\lambda, \mathbf{b}_\lambda) = \Phi(\lambda) \in \mathbb{P}$ and $\varphi^{(\lambda)}$ satisfies

$$|\partial^\beta \hat{\varphi}^{(\lambda)}(\xi)| \lesssim \min\left(1, a_\lambda + |\xi_1| + a_\lambda^{1/2}|\xi_2|\right)^M \langle|\xi|\rangle^{-N_1}\langle\xi_2\rangle^{-N_2} \tag{1}$$

for all multi-indices $|\beta| \leqslant R$. The implicit constants are uniform over $\lambda$.

Definition 2.2 implies a number of useful consequences. Firstly, it implies

$$|\hat{m}_\lambda(\xi)| \lesssim a_\lambda^{1/2}\min(1, a_\lambda(1+|\xi|))^M\langle a\xi\rangle^{-N_1}\langle a^{1/2}(R_{\theta_\lambda}\xi)_2\rangle^{-N_2}.$$

Similar estimates hold for it's derivatives. Therefore, the definition implies a (somewhat biased) directional decay as the coordinates tend to infinity and $M$ almost vanishing moments. Furthermore, $R$ describes spatial localisation of the molecule while $N_1$ and $N_2$ are statements about its smoothness.

We will now introduce a pseudo-distance function, which is just a continuous setting analogue of the one used in [2], which in turn is a variation of the pseudo-distance introduced in [5].

**Definition 2.3.** The pseudo-distance function $w\colon \mathbb{P} \times \mathbb{P} \to \mathbb{R}$ is for a pair of triplets $\lambda = (a_\lambda, \theta_\lambda, \mathbf{b}_\lambda)$, $\nu = (a_\nu, \theta_\nu, \mathbf{b}_\nu) \in \mathbb{P}$ defined by

$$w(\lambda, \nu) = \frac{a_M}{a_m}\left(1 + a_M^{-1}d(\lambda, \nu)\right),$$

where

$$\begin{aligned} a_m &= \min(a_\lambda, a_\nu), \\ a_M &= \max(a_\lambda, a_\nu), \\ d(\lambda, \nu) &= |\theta_\lambda - \theta_\nu|^2 + |\mathbf{b}_\lambda - \mathbf{b}_\nu|^2 + |\langle \mathbf{e}_\lambda, \mathbf{b}_\lambda - \mathbf{b}_\nu\rangle|, \\ \mathbf{e}_\lambda &= (\cos(\theta_\lambda), \sin(\theta_\lambda))^\tau. \end{aligned}$$

The function $w$ is not a proper distance function but is not too far off. A detailed list of properties of $w$ can be found in [2]. We should note that (real-valued) curvelets and shearlets are associated with a ray, hence, angle differences $\theta_\lambda - \theta_\nu$ are understood modulo $\pi$.

In the forthcoming text, when applying this pseudo-distance to parametrisations of some two families of continuous parabolic molecules, we will write $w(\lambda, \nu)$ for $\lambda \in \Lambda_1, \nu \in \Lambda_2$ when, if written properly, it should read $w(\Phi_1(\lambda), \Phi_2(\nu))$. This implicit notation is intended to make the notation a little less cumbersome, hopefully with no losses to the clarity of exposition.

The fundamental property of continuous parabolic molecules, which we will establish in this paper is the notion that any two families of continuous parabolic molecules are almost orthogonal, in the sense that they exhibit strong off-diagonal decay. This decay will be given in terms of controlling the inner product of two given families of CPMs with the decay of the pseudo-distance function $w$ between their indices.

We will now introduce the notion of admissibility of a parametrisation. First we need to define the canonical parametrisation.

**Definition 2.4.** The parametrisation pair $(\Lambda_0, \Phi)$ where $\Lambda_0 = \mathbb{P}$, and $\Phi$ is the identity is called the *canonical parametrisation*.

The following notion will be essential in the upcoming proofs.

**Definition 2.5.** Index set $\Lambda$ is said to be $k$-*admissible* if

$$\sup_{\lambda \in \Lambda} \int_{\Lambda_0} w(\lambda, \nu)^{-k} d\mu(\nu) < \infty \quad \text{and} \quad \sup_{\lambda \in \Lambda_0} \int_{\Lambda} w(\lambda, \nu)^{-k} d\mu(\lambda) < \infty \tag{2}$$

where $d\mu(\lambda) = \frac{d\lambda}{a^3}$.

# 3 Examples of Continuous Parabolic Molecules

In recent years a number of representation systems based on parabolic scaling have been introduced. Some examples of those are Hart Smith's transform [5], curvelets [1], shearlets [6], and others. The goal of this section is to show that the CPM framework encompasses both the curvelet and the shearlet-type systems.

The approach of using abstract, nondescript molecules has historical precedence. Most notably for the present paper, in [2] the authors introduced the notion of curvelet molecules, and similarly, the authors of [3] did the same for shearlet molecules. It is important to note that both of notions of molecules were based and influenced by the vaguelettes [7]. We will show that CPMs provide a generalisation of both of these concepts. In the remainder of this section we will briefly introduce curvelets and curvelet molecules, and show that curvelet molecules form a family of continuous parabolic molecules. Then we will do the same for shearlets.

## 3.1 Curvelets

Denote by $r$ and $\omega$ the polar coordinates in the frequency domain. Take a pair of smooth, non-negative and real-valued windows $W(r)$ and $V(\omega)$, which are called the *radial window* and the *angular window*, respectively. Furthermore, assume that $W$ takes positive real arguments, and is supported on $[1/2, 2]$, while $V$ takes real arguments and is supported on $[-1, 1]$. These windows must also satisfy the following admissibility conditions

$$\int_0^\infty W(ar)^2 \frac{da}{a} = 1, \ \forall r > 0,$$
$$\int_{-1}^{1} V(u)^2 du = 1.$$

At scale $a$ the generating element $\gamma_{a\mathbf{0}0}$ is defined via polar coordinates in the frequency domain as

$$\hat{\gamma}_{a\mathbf{0}0}(r, \alpha) = a^{3/4} W(ar) V(\alpha/\sqrt{a}).$$

The scale parameter $a$ has to satisfy $a \leqslant a_0$, where $a_0$ represents the coarsest scale and must obey $a_0 \leqslant \pi^2$ for the construction to work, but we will take it to be 1. The remaining curvelets at scale $a$ are defined via rotations and translations of the generating element $\gamma_{a\mathbf{0}0}$

$$\gamma_\lambda(\mathbf{x}) = \gamma_{a_\lambda \mathbf{0}0}\left(R_{\theta_\lambda}(\mathbf{x} - \mathbf{b}_\lambda)\right), \text{ where } \quad \lambda = (a_\lambda, \theta_\lambda, \mathbf{b}_\lambda) \in \Lambda := [0, a_0] \times [0, 2\pi) \times \mathbb{R}^2.$$

The family $\Gamma = \{\gamma_\lambda : \lambda \in \Lambda\}$ is called the *family of second generation curvelets*.

Let us now define continuous curvelet molecules.

**Definition 3.1.** Take an index set $\Lambda_0 = [0, a_0] \times [0, 2\pi) \times \mathbb{R}^2$. A family $\{m_\lambda : \lambda \in \Lambda_0\}$ of functions is called a family of *curvelet molecules* of regularity $R$ if it can be expressed as

$$m_\lambda(\mathbf{x}) = a_\lambda^{-3/4} \varphi^{(\lambda)}\left(D_{1/a_\lambda} R_{\theta_\lambda}(\mathbf{x} - \mathbf{b}_\lambda)\right)$$

such that

$$|\partial^\beta \varphi^{(\lambda)}(\mathbf{x})| \lesssim \langle|\mathbf{x}|\rangle^{-N}$$

and

$$|\hat{\varphi}^{(\lambda)}(\boldsymbol{\xi})| \lesssim \min\left(1, a_\lambda + |\xi_1| + a_\lambda^{1/2}|\xi_2|\right)^M \tag{3}$$

hold $|\beta| \leqslant R$, $N = 0, 1, 2, \ldots$, and all $M = 0, 1, \ldots, R$. The implicit constants are uniform over $\lambda \in \Lambda_0$.

Looking at Definitions 2.2 and 3.1 side by side, it should come as no surprise that second generation curvelets constitute a family of curvelet molecules for an arbitrary degree of regularity $R$ [1]. We begin by showing that every family of curvelet molecules is also a family of CPMs. The proof can be found in Appendix A.

**Proposition 3.1.** *A system of curvelet molecules of regularity* $3R/2$ *constitutes a system of CPMs of order* $(R, R, R/2, R/2)$*, with canonical parametrisation.*

The first step in establishing that curvelet molecules are a subtype of continuous parabolic molecules is defining the relevant parametrisation. The canonical parametrisation, which was defined in Definition 2.4, has been constructed with exactly curvelets in mind. As we mentioned before, admissibility of parametrisations will play a crucial role later on. We will show now that the curvelet parametrisation is admissible for all $k > 2$.

**Lemma 3.2.** *Canonical parametrisation* $(\Lambda_0, \Phi)$ *is* $k$*-admissible for all* $k > 2$.

*Proof.* We want to show that

$$\sup_{\nu \in \Lambda_0} \int_{\Lambda_0} w(\lambda, \nu)^{-k} d\mu(\lambda) < \infty$$

holds for all $k > 2$. We have

$$\int_{\Lambda_0} w(\lambda, \nu)^{-k} d\mu(\lambda) = \int_{[0, a_0]} \frac{a_m^k}{a_M^k} \int_{[0,2\pi) \times \mathbb{R}^2} \left[1 + a_M^{-1} d(\lambda, \nu)\right]^{-k} d\mu(\lambda).$$

It can be shown (Lemma A.3) that

$$\int_{[0,2\pi) \times \mathbb{R}^2} \left[1 + q^{-1} d(\lambda, \nu)\right]^{-k} d\theta d\mathbf{b} \lesssim q^2 \tag{4}$$

holds for all $q \in \mathbb{R}^+$. Hence, it follows

$$\begin{aligned} \int_{\Lambda_0} w(\lambda, \nu)^{-k} d\mu(\lambda) &\lesssim \int_0^{a_0} \frac{a_m^k}{a_M^k} a_M^2 \frac{da}{a^3} = \tilde{a}^{-k+2} \int_0^{a_\nu} a^{k-3} da + a_\nu{}^k \int_{a_\nu}^{a_0} a^{-k-1} da \\ &\lesssim \frac{1}{k-2} + \frac{1}{k} - \frac{1}{k} a_\nu{}^k a_0^{-k} \leqslant \frac{1}{k-2} + \frac{1}{k} < \infty \end{aligned}$$

which is true as long as $k > 2$. In other words, $\Lambda_0$ is $k$-admissible for $k > 2$. □

Finally, we are ready to show that second generation curvelets constitute a family of continuous parabolic molecules.

**Proposition 3.3.** *Second generation curvelets are a family of CPMs of order* $(R, R, R/2, R/2)$, *for an arbitrary* $R \in \mathbb{N}$, *whose parametrisations is admissible for all* $k > 2$.

*Proof.* By [2], second generation curvelets constitute a family of curvelet molecules of arbitrary degree of regularity $R$. Hence, the statement follows by Proposition 3.1. □

## 3.2 Shearlets

In order to cover all the possible orientations in $\mathbb{R}^2$ curvelets use rotations, that is, we consider actions of rotation operators $R_\theta$, for $\theta \in [0, 2\pi)$, on generating elements. Shearlets, on the other hand, handle directions through the shearing operator, given by the shearing matrix $S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$, for $s \in \mathbb{R}$. The difference between the two is that the rotation operators rotate both coordinates by a given angle, whereas shearing changes the slope of a given point by displacing its $y$-coordinate with respect to the shearing variable $s$. The rationale behind using shears is that rotations destroy the integer lattice $\mathbb{Z}^2$, unless the rotation angle is $k\pi/2$ for an integer $k$, while shearing leaves $\mathbb{Z}^2$ invariant as long as $s$ is an integer, thus allowing for a unified treatment of the continuous and discrete settings.

We can now define the continuous shearlet system.

**Definition 3.2.** For $\psi \in L^2(\mathbb{R}^2)$ satisfying the admissibility condition

$$\int_{\mathbb{R}^2} \frac{|\hat{\psi}(\boldsymbol{\xi})|^2}{\xi_1^2} d\boldsymbol{\xi} < \infty,$$

the continuous shearlet system is defined as the family of functions $(\psi_\lambda)_\lambda$, with $\lambda = (a_\lambda, s_\lambda, \mathbf{b}_\lambda) \in \mathbb{R}^+ \times \mathfrak{S} \times \mathbb{R}^2$, where

$$\psi_\lambda(\mathbf{x}) = a_\lambda^{-3/4} \psi(D_{1/a_\lambda} S_{s_\lambda}(\mathbf{x} - \mathbf{b}_\lambda)).$$

We refer to $\psi$ as the mother shearlet. For the construction to work the set $\mathfrak{S}$ should be a (symmetric) subset of $\mathbb{R}$ that contains $[-1, 1]$. There are two standard approaches here. Initial constructions of shearlet family took $\mathfrak{S}$ to be the whole $\mathbb{R}$, but this comes with a serious disadvantage. Namely, it is easy to notice that shearlets of this type would exhibit a certain directional bias. In other words, if we were to detect a singularity which is arbitrarily close to the $y$ axis we would need to consider the shearing parameter as it tends to infinity, and if it were on the $y$-axis we would need to look at the limit. This is clearly a situation we would not mind avoiding, but more importantly, it would pose great problems in applications.

To overcome these shortcomings of the classical shearlet construction, we instead typically consider the *cone-adapted shearlets*. Here we address this issue by splitting the frequency domain into 4 high frequency cones, and a low frequency box, as in figure 1. Then we can define the shearlets on the horizontal cones and restrict $\mathfrak{S}$ to a finite set, say $\mathfrak{S} = [-\Xi, \Xi]$ with $0 < \Xi < \infty$. The corresponding shearlets on the vertical cones are obtained by simply swapping the roles of $x$ and $y$ variables.

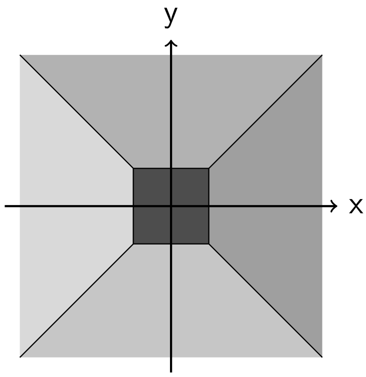


Figure 1: A partitioning of the (frequency) plane into a low-frequency box and four high frequency cones

Further details regarding construction of shearlets and related topics can be found in [6]. Let us now define continuous shearlet molecules, were we use the cone-adapted approach to shearlets.

**Definition 3.3.** Take $D^0_a = \operatorname{diag}(a, a^{1/2}), D^1_a = \operatorname{diag}(a^{1/2}, a)$, and $S^0_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$, $S^1_s = \left(S^0_s\right)^\tau$, and define the index set

$$\Lambda = \left\{(\epsilon, a, s, \mathbf{b}) : \epsilon \in \{0, 1\},\ a \in [0, a_0],\ s \in \mathfrak{S},\ \mathbf{b} \in \mathbb{R}^2\right\},\ \text{where } \mathfrak{S} = [-\Xi, \Xi].$$

For $\lambda = (\epsilon, a_\lambda, s_\lambda, \mathbf{b}_\lambda) \in \Lambda$, and suitable functions $\phi, \psi^\epsilon$, define the shearlet system $\Sigma = \{\sigma_\lambda : \lambda \in \Lambda\}$ by

$$\begin{aligned} \sigma^\epsilon_\lambda(\mathbf{x}) &= a_\lambda{}^{-3/4} \psi^{(\epsilon,\lambda)} \left(D^\epsilon_{1/a_\lambda} S^\epsilon_s (\mathbf{x} - \mathbf{b})\right), \\ \sigma_\lambda(\mathbf{x}) &= \phi(\mathbf{x} - \mathbf{b}) \quad \text{for } \lambda = (\epsilon, 0, 0, \mathbf{0}). \end{aligned}$$

We call $\Sigma$ a system of *Continuous Shearlet Molecules* of order $(R, M, N_1, N_2)$ if functions $\phi$ and $\psi^{(\epsilon,\lambda)}$ satisfy

$$|\partial^\beta \hat{\psi}^{(\epsilon,\lambda)}(\boldsymbol{\xi})| \lesssim \min(1, a_\lambda + |\xi_1| + a_\lambda{}^{1/2}|\xi_2|)^M \langle|\boldsymbol{\xi}|\rangle^{-N_1} \langle \xi_{2-\epsilon} \rangle^{-N_2} \tag{5}$$

for all $\beta \in \mathbb{N}^2_0$ such that $|\beta| \leqslant R$.

Following this rather lengthy definition we will now show that shearlet molecules are a special case of continuous parabolic molecules.

**Proposition 3.4.** *Assume that the system $\Sigma$ constitutes a system of shearlet molecules of order $(R, M, N_1, N_2)$. Then $\Sigma$ constitutes a system of continuous parabolic molecules of the same order, with parametrisation $\left(\Lambda^\Sigma_i, \Psi^\Sigma_i\right)_{i \in I}$ where*

$$\begin{aligned} \Lambda^\Sigma_0 = \Lambda^\Sigma_1 &= [0, a_0] \times \mathfrak{S} \times \mathbb{R}^2, \\ \Phi^\Sigma_i(a, s, \boldsymbol{b}) &= \left(a, i\frac{\pi}{2} + \arctan(-s), \boldsymbol{b}\right), \end{aligned}$$

*for* $i = 0, 1$.

*Proof.* Without loss of generality we will restrict our discussion to the case when $i = 0$. Let us recall that CPM systems are to be written in the form

$$m_\lambda(\mathbf{x}) = a_\lambda^{-3/4} \varphi^{(\lambda)}(D_{1/a_\lambda} R_{\theta_\lambda}(\mathbf{x} - \mathbf{b}_\lambda)).$$

We can take $\mathbf{b} = \mathbf{0}$, without loss of generality. Therefore, we write (omitting the index $\lambda$)

$$\varphi(\mathbf{x}) = \psi(D_{1/a_\lambda} S_{s_\lambda} R_{\theta_\lambda}^{-1} D_{a_\lambda} \mathbf{x}).$$

The Fourier transform is given by

$$\hat{\varphi}(\boldsymbol{\xi}) = \hat{\psi}(D_{a_\lambda} S_{s_\lambda}^{-\tau} R_{-\theta_\lambda} D_{1/a_\lambda} \boldsymbol{\xi}).$$

Denote $A = D_{a_\lambda} S_{s_\lambda}^{-\tau} R_{-\theta_\lambda} D_{1/a_\lambda}$. Since $\theta_\lambda = \arctan(-s)$, we have

$$A = \begin{pmatrix} \tau_1(-s) & a_\lambda^{-1/2} \sin(\arctan(-s)) \\ 0 & \tau_2(-s) \end{pmatrix}$$

where $\tau_1(t) = \cos(\arctan(t))$ and $\tau_2(t) = t\sin(\arctan(t)) + \cos(\arctan(t))$. Since $\mathfrak{S}$ is bounded we have

$$c_1 \leqslant \tau_1(t) \leqslant C_1, \quad \text{and} \quad c_2 \leqslant \tau_2(t) \leqslant C_2, \tag{6}$$

where the constants in question depend only on $\mathfrak{S}$. The estimates (6) follow easily once we use trigonometric identities to rewrite $\tau_1$ and $\tau_2$ as $\tau_1(t) = (t^2+1)^{-1/2}$, that is, $\tau_2(t) = (t^2+1)^{1/2}$.

In order to obtain bounds on the derivatives of $\hat{\varphi}$, we will now use the assumptions, that is, inequality (5), regarding the decay of shearlet molecules. We have

$$|\partial^\beta \hat{\varphi}(\boldsymbol{\xi})| \lesssim \sup_{|\gamma| \leqslant R} |\partial^\gamma \hat{\psi}(A\boldsymbol{\xi})| \lesssim \min(1, a_\lambda + |(A\boldsymbol{\xi})_1| + a_\lambda^{1/2}|(A\boldsymbol{\xi})_2|)^M \langle |A\boldsymbol{\xi}| \rangle^{-N_1} \langle (A\boldsymbol{\xi})_2 \rangle^{-N_2}.$$

What is left is to estimate the terms in the previous equation to ensure that $\hat{\varphi}$ satisfies decay conditions (1). We have

$$|\boldsymbol{\xi}| \leqslant \|A^{-1}\| |A\boldsymbol{\xi}| \Rightarrow \langle |A\boldsymbol{\xi}| \rangle^{-N_1} \leqslant (\min(1, \|A^{-1}\|^{-1})^{-N_1} \langle |\boldsymbol{\xi}| \rangle^{-N_1} \lesssim \langle |\boldsymbol{\xi}| \rangle^{-N_1},$$

where the matrix norm is bounded due to (6). Through a similar argument we can find bounds for other terms. In conclusion, we have

$$|\partial^\beta \hat{\varphi}(\boldsymbol{\xi})| \lesssim \min(1, a_\lambda + |\xi_1| + a_\lambda{}^{1/2}|\xi_2|)^M \langle \|\boldsymbol{\xi}\| \rangle^{-N_1} \langle \xi_2 \rangle^{-N_2},$$

which is what we wanted to show. □

The next step is to establish the $k$-admissibility of the shearlet parametrisation.

**Proposition 3.5.** *The set* $\Lambda^\Sigma = [0, a_0] \times \mathfrak{S} \times \mathbb{R}^2$, *where* $\mathfrak{S} = [-\Xi, \Xi]$ *is* $k$*-admissible for all* $k > 2$.

*Proof.* We only need to show

$$\sup_{\nu\in\Lambda_0}\int_{\Lambda^\Sigma} w(\lambda,\nu)^{-k}d\mu(\lambda) < \infty,$$

since the other statement required by Definition 2.5 follows along exactly the same lines as the proof of Lemma 3.2. We have

$$\begin{aligned}\int_{\Lambda^\Sigma} w(\lambda,\nu)^{-k}d\mu(\lambda) &= \int_{\Lambda^\Sigma}\left[\frac{a_M}{a_m}(1+a_M^{-1}d(\lambda,\nu))\right]^{-k}d\mu(\lambda)\\ &= \int_{[0,a_0]}\frac{a_m^k}{a_M^k}\left(\int_{\mathfrak{S}}\int_{\mathbb{R}^2}(1+a_M^{-1}d(\lambda,\nu))^{-k}dsd\mathbf{b}\right)\frac{da}{a^3}\end{aligned}$$

Therefore, if we could prove an analogue of (4), the rest of the proof would be the same as the proof of Lemma 3.2. We have

$$\begin{aligned}&\int_{\mathfrak{S}}\int_{\mathbb{R}^2}(1+qd(\lambda,\nu))^{-k}dsd\mathbf{b}_\lambda\\ &\quad= \int_{\mathfrak{S}}\int_{\mathbb{R}^2}\left(1+q(|\mathbf{b}_\lambda-\mathbf{b}_\nu|^2+|\arctan(-s)-\theta_\nu|^2+|\langle\mathbf{e}_\lambda,\mathbf{b}_\lambda-\mathbf{b}_\nu\rangle|\right)^{-k}dsd\mathbf{b}_\lambda\\ &\quad\leqslant \int_{\arctan(-\mathfrak{S})}\int_{\mathbb{R}^2}\left(1+q(|\mathbf{b}_\lambda-\mathbf{b}_\nu|^2+|\theta_\lambda-\theta_\nu|^2+|\langle\mathbf{e}_\lambda,\mathbf{b}_\lambda-\mathbf{b}_\nu\rangle|\right)^{-k}(\theta_\lambda^2+1)d\theta_\lambda d\mathbf{b}_\lambda\\ &\quad\leqslant C_{\mathfrak{S}}\int_{\mathbb{R}}\int_{\mathbb{R}^2}\left(1+q(|\mathbf{b}_\lambda-\mathbf{b}_\nu|^2+|\theta_\lambda-\theta_\nu|^2+|\langle\mathbf{e}_\lambda,\mathbf{b}_\lambda-\mathbf{b}_\nu\rangle|\right)^{-k}d\theta_\lambda d\mathbf{b}_\lambda\\ &\quad\lesssim q^{-2}\end{aligned}$$

where we used the change of variables ($\theta = \arctan(-s)$), the boundednes of $\mathfrak{S}$ and Lemma A.3. Hence, the claim follows. ☐

There are two further important sub-types of shearlets. First one is that of band limited shearlets. These are shearlets such that the Fourier transform of the mother shearlet has a compact support. This paper includes a very specific construction of a band limited, cone-adapted shearlet family that is also a family of CPMs, allows a certain reconstruction formula, and whose dual is also a family of CPMs. Details are in Appendix B. The importance of having such a family will become clear in Chapter 5.

Another important class of shearlet systems is the class of compactly supported shearlets. They will be of great importance in later parts of Chapter 5. Here we shall consider compactly supported shearlets with separable generators. In other words, let

$$\psi^0(\mathbf{x}) = \psi_1(x_1)\psi_2(x_2),\ \text{and}\ \psi^1(\mathbf{x}) = \psi^0(x_2,x_1).$$

Given a dilation parameter $a$, a shearing parameter $s$ and a location $\mathbf{b}$, we define

$$\psi^i_{as\mathbf{b}} = a^{-3/4}\psi^\epsilon\left(D_{1/a}S_s(\mathbf{x}-\mathbf{b})\right),\quad i\in\{0,1\}. \tag{7}$$

In order to ensure that this defines a system of parabolic molecules we need to endow the generators $\psi_1$ and $\psi_2$ with sufficient smoothness and moments. We say that a real function $\rho$ has $K$ (anistropic) vanishing moments if

$$\int_{\mathbb{R}}\frac{|\hat{\rho}(\xi)|^2}{|\xi|^{2K}}d\xi < \infty.$$

Notice that this is equivalent to saying that $\rho(x) = \frac{\partial^K}{\partial x^K}\theta$, where $\hat{\theta} \in L^2(\mathbb{R})$. Assuming that we have sufficient smoothness and moments, we can now show that the system (7) is a system of CPMs, and furthermore, that a projection of the shearlets onto a frequency cone $\mathcal{C}_{u,v}$ has the additional property of admitting a representation formula for $f \in L^2(\mathcal{C}_{u,v})\check{}$. The details can be found in Appendix A.

**Proposition 3.6.** *Consider the shearlet system (7), such that* $\psi_1 \in C^{N_1}(\mathbb{R})$ *has compact support and* $M + R$ *anisotropic moments, and that* $\psi_2 \in C^{N_1+N_2}(\mathbb{R})$ *also has compact support, where* $M, R, N_1$ *and* $N_2$ *satisfy*

$$2(M + R) - 1/2 > N_1 + N_2 > M + R > 1/2.$$

*Then (7) constitutes a system of continuous parabolic molecules of order* $(R, M + N_1, N_1, N_2)$. *Furthermore, the system*

$$\{P_{\mathcal{C}_{u,v}}\psi_{as\boldsymbol{b}} : a \in [0, 1] s \in [-\Xi, \Xi], \boldsymbol{b} \in \mathbb{R}^2\} \cup \{T_{\boldsymbol{b}} P_{\mathcal{C}_{u,v}} W : \boldsymbol{b} \in \mathbb{R}^2\} \tag{8}$$

*is a tight frame for* $L^2(\mathcal{C}_{u,v})\check{}$, *provided* $u > 0$, $\Xi > v$.

# 4 Almost Orthogonality

We will now state and prove the almost orthogonality of continuous parabolic molecules, the essential tool used in this paper.

**Theorem 4.1.** *Let* $\Gamma = \{m_\lambda : \lambda \in \Lambda^\Gamma\}$ *and* $\Sigma = \{n_\nu : \nu \in \Lambda^\Sigma\}$ *be two families of continuous parabolic molecules, both of order* $(R, M, N_1, N_2)$*. Then*

$$|\langle m_\lambda, n_\nu \rangle| \leqslant w(\lambda, \nu)^{-N}$$

*holds for every* $N \in \mathbb{N}$ *such that*

$$R \geqslant 2N, \quad M > 3N - \frac{5}{4}, \quad N_1 \geqslant N + \frac{3}{4}, \quad N_2 \geqslant 2N.$$

This result certainly should not come as a surprise, as its discrete setting analogue can be found in [3], though it has not been yet shown in the continuous setting. The proof is quite similar to its discrete analogues, apart from some technical differences.

*Proof.* Since $\Gamma$ and $\Sigma$ are CPMs, we can write

$$m_\lambda(\mathbf{x}) = a_\lambda^{-3/4} \varphi^{(\lambda)} \left( D_{1/a_\lambda} R_{\theta_\lambda} (\mathbf{x} - \mathbf{b}_\lambda) \right),$$
$$n_\nu(\mathbf{x}) = a_\nu^{-3/4} \psi^{(\nu)} \left( D_{1/a_\nu} R_{\theta_\nu} (\mathbf{x} - \mathbf{b}_\nu) \right).$$

Parseval's equality and Lemma A.4 give

$$\begin{aligned} \langle m_\lambda, n_\nu \rangle &= \langle \hat{m}_\lambda, \hat{n}_\nu \rangle = \int_{\mathbb{R}^2} \hat{m}_\lambda(\xi) \overline{\hat{n}_\nu(\xi)} d\xi \\ &= (a_\lambda a_\nu)^{3/4} \int_{\mathbb{R}^2} \hat{\varphi}^{(\lambda)} \left( D_{a_\lambda} R_{\theta_\lambda} \xi \right) \overline{\hat{\psi}^{(\nu)} \left( D_{a_\nu} R_{\theta_\nu} \xi \right)} e^{-2\pi\imath (\mathbf{b}_\lambda - \mathbf{b}_\nu) \cdot \xi} d\xi. \end{aligned} \tag{9}$$

Integration by parts gives

$$\begin{aligned} &\int_{\mathbb{R}^2} \hat{\varphi}^{(\lambda)} \left( D_{a_\lambda} R_{\theta_\lambda} \xi \right) \overline{\hat{\psi}^{(\nu)} \left( D_{a_\nu} R_{\theta_\nu} \xi \right)} e^{-2\pi\imath (\mathbf{b} - \mathbf{b}_\nu) \cdot \xi} d\xi = \\ &\quad = \int_{\mathbb{R}^2} \mathcal{L}_{\lambda,\nu}^{k} \left( \hat{\varphi}^{(\lambda)} \left( D_{a_\lambda} R_{\theta_\lambda} \xi \right) \overline{\hat{\psi}^{(\nu)} \left( D_{a_\nu} R_{\theta_\nu} \xi \right)} \right) \mathcal{L}_{\lambda,\nu}^{-k} \left( e^{-2\pi\imath (\mathbf{b}_\lambda - \mathbf{b}_\nu) \cdot \xi} \right) d\xi, \end{aligned}$$

where the differential operator $\mathcal{L}_{\lambda,\nu}$ is defined via

$$\mathcal{L}_{\lambda,\nu} = \mathcal{I} - a_M^{-1} \Delta - \frac{a_M^{-2}}{1 + a_M^{-1} |\theta_\lambda - \theta_\nu|^2} \frac{\partial^2}{\partial e_\lambda^2}, \tag{10}$$

where $a_M = \max(a_\lambda, a_\nu)$.

Let us introduce some short hand notation. Denote $\delta\mathbf{b} = \mathbf{b}_\lambda - \mathbf{b}_\nu$ and $\delta\theta = \theta_\lambda - \theta_\nu$. Lemma A.5 states that the exponentials are eigenfunctions of $\mathcal{L}_{\lambda,\nu}$. Thus, we have

$$\mathcal{L}_{\lambda,\nu}^{-k} \left( e^{-2\pi\imath \xi \cdot \delta\mathbf{b}} \right) = \left[ 1 + 4\pi^2 a_M^{-1} |\delta\mathbf{b}|^2 + 4\pi^2 \frac{a_M^{-2}}{1 + a_M^{-1} |\delta\theta|^2} \langle \mathbf{e}_\lambda, \delta\mathbf{b} \rangle^2 \right]^{-k} e^{-2\pi\imath \xi \cdot \delta\mathbf{b}}. \tag{11}$$

On the other hand, Lemma A.6 for $k \leqslant \frac{R}{2}$ provides a bound for $\mathcal{L}^k_{\lambda,\nu}$

$$\left| \mathcal{L}^k_{\lambda,\nu}\left(\hat{\varphi}^{(\lambda)}\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{\psi}^{(\nu)}\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}\right)\right| \lesssim S_{\lambda,M-N_2,N_1,N_2}(\xi)S_{\nu,M-N_2,N_1,N_2}(\xi). \tag{12}$$

Plugging (11) and (12) into (9) yields

$$|\langle m_\lambda, n_\nu\rangle \lesssim \mathcal{S}\left[1 + a_M^{-1}|\delta\mathbf{b}|^2 + \frac{a_M^{-2}}{1+a_M^{-1}|\delta\theta|^2}\langle \mathbf{e}_\lambda, \delta\mathbf{b}\rangle^2\right]^{-N}, \tag{13}$$

where

$$\mathcal{S} = (a_\lambda a_\nu)^{3/4}\int_{\mathbb{R}^2} S_{\lambda,M-N_2,N_1,N_2}(\xi)S_{\nu,M-N_2,N_1,N_2}(\xi)d\xi.$$

Lemma A.8 gives a bound on $\mathcal{S}$

$$\mathcal{S} \lesssim \left(\frac{a_M}{a_m}\right)^{-N}\left(1 + a_M^{-1/2}|\delta\theta|\right)^{-N}.$$

Hence, it follows

$$\begin{aligned}|\langle m_\lambda, n_\nu\rangle| &\lesssim \left(\frac{a_M}{a_m}\right)^{-N}\left(1 + a_M^{-1/2}|\delta\theta|\right)^{-N}\left[1 + a_M^{-1}|\delta\mathbf{b}|^2 + \frac{a_M^{-2}}{1+a_M^{-1}|\delta\theta|^2}\langle \mathbf{e}_\lambda, \delta\mathbf{b}\rangle^2\right]^{-N} \\ &\lesssim \left(\frac{a_M}{a_m}\right)^{-N}\left(1 + a_M^{-1}|\delta\theta|^2 + a_M^{-1}|\delta\mathbf{b}|^2 + \frac{1}{a_M^2 + a_M|\delta\theta|^2}\langle \mathbf{e}_\lambda, \delta\mathbf{b}\rangle^2\right)^{-N}\end{aligned}$$

Lastly, we have

$$\begin{aligned}1 + a_M^{-1}|\delta\theta|^2 + \frac{1}{a_M^2 + a_M|\delta\theta|^2}\langle \mathbf{e}_\lambda, \delta\mathbf{b}\rangle^2 &= \left(\sqrt{1 + a_M^{-1}|\delta\theta|^2}\right)^2 + \left(\frac{|\langle \mathbf{e}, \delta\mathbf{b}\rangle|}{\sqrt{a_M^2 + a_M|\delta\theta|^2}}\right)^2 \\ &\geqslant 2\sqrt{1 + a_M^{-1}|\delta\theta|^2}\frac{|\langle \mathbf{e}, \delta\mathbf{b}\rangle|}{\sqrt{a_M^2 + a_M|\delta\theta|^2}} = 2a_M^{-1}|\langle \mathbf{e}, \delta\mathbf{b}]\rangle|\end{aligned}$$

Therefore,

$$\begin{aligned}|\langle m_\lambda, n_\nu\rangle| &\lesssim \left(\frac{a_M}{a_m}\right)^{-N}\left(1 + a_M^{-1}|\delta\mathbf{b}|^2 + a_M^{-1}|\langle \mathbf{e}_\lambda, \delta\mathbf{b}\rangle|\right)^{-N} \\ &\lesssim \left(\frac{a_M}{a_m}\right)^{-N}\left(1 + a_M^{-1}\left(|\delta\mathbf{b}|^2 + |\delta\theta|^2 + |\langle \mathbf{e}_\lambda, \delta\mathbf{b}\rangle|\right)\right)^{-N} \\ &\lesssim w(\lambda,\nu)^{-N},\end{aligned}$$

which concludes the proof.

□

# 5 Microlocal Analysis

Now that we have all the necessary tools in place we can put them to use. As the title of the current chapter would suggest, this will be done in the framework of microlocal analysis.

The goal is to show that all members of the class decomposition systems based on parabolic scaling answer the same questions. In other words, we want to show that there are results in microlocal analysis that hold for all sufficiently good families of parabolic molecules. Since we are dealing with decomposition systems, an important object of interest are the frame coefficients. Therefore, in view of our objective, we would need tools that would enable bridging the gap between statements regarding frame coefficients of one family to analogous statements including frame coefficients of some other family of continuous parabolic molecules. To that end, we will need reconstruction formulas, that is, given a sequence of frame coefficients we require a formula that puts the pieces back together and reconstructs the original signal. We can distinguish between two cases. The first case is that of system which admit a reconstruction formula which is valid for functions without any specific restrictions on the support of their Fourier transform. The other case is that of systems which admit a reconstruction formula which is valid only for functions such that the support of their Fourier transform is inside some cone in the frequency plane.

## 5.1 Parabolic Molecules and Frames

Let us get back to the task at hand. We first need to set the framework we will be working in. We shall begin by defining microlocal Sobolev regularity.

**Definition 5.1.** We say that a distribution $f$ is microlocally in the $L^2$ Sobolev space $H^k$ at $(\theta_0, \mathbf{x}_0)$, written $f \in H^k(\theta_0, \mathbf{x}_0)$, if for some smooth bump function $\varphi \in C^\infty(\mathbb{R}^2)$, with $\varphi(\mathbf{x}_0) \neq 0$, localised to a ball near $\mathbf{x}_0$, and for some smooth bump function $\beta \in C^\infty_{per}[0, 2\pi)$, obeying $\beta(\theta_0) = 1$ and localised to a ball near $\theta_0$, the space/direction localised function $f_{\varphi,\beta}$, defined in polar Fourier coordinates by $\beta(w)\widehat{\varphi f}(r\cos(w), r\sin(w))$ belongs to the weighted $L^2$ space $L^2((1 + |\xi|^2)^{k/2} d\xi)$.

Candés and Donoho showed in [1] that this notion of microlocal regularity can be determined by an $L^2$ condition, as stated in the following theorem.

**Theorem 5.1.** *Let $S^k_2(\theta, \boldsymbol{x})$ denote the (normal-approach, parabolic scaling) square function*

$$S^k_2(\theta, \boldsymbol{x}) = \left( \int_0^{a_0} |\langle f, \gamma_{a\theta\boldsymbol{x}} \rangle|^2 a^{-2k} \frac{da}{a^3} \right)^{1/2}, \tag{14}$$

*where $\{\gamma_{a\theta\boldsymbol{x}}, a \in [0, a_0], \theta \in [0, 2\pi), \boldsymbol{x} \in \mathbb{R}^2\}$ are second generation curvelets.*

*The distribution $f$ is in $H^k(\theta_0, \boldsymbol{x}_0)$ if and only if for some neighbourhood $\mathcal{N}$ of $(\theta_0, \boldsymbol{x}_0)$ we have*

$$\int_{\mathcal{N}} (S^k_2(\theta, \boldsymbol{x}))^2 d\theta d\boldsymbol{x} < \infty.$$

In Theorem 5.1 we use curvelets to resolve microlocal Sobolev regularity, but there is no reason that should make us think that such a statement should hold only for curvelets. Rather, it would seem sensible to assume that an analogous result should hold for many other directional representation systems that are based on parabolic scaling. Indeed, it is this directional focus based on parabolic scaling that makes all the difference.

A slightly different reading of Theorem 5.1 would be to read it as a saying that $f$ is in $H^k(\theta_0, \mathbf{x}_0)$ if and only if $a^{-k}\langle f, \gamma_{a\theta\mathbf{x}}\rangle$ is in $L^2([0, a_0] \times \mathcal{N}, \mu)$. This is the interpretation we will use. Our goal now is to extend this result to other systems of CPMs. To begin, take $\Gamma = \{m_\lambda : \lambda \in \Lambda_\Gamma\}$ and $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$ to be two families of CPMs, with parametrisations $(\Lambda_\Gamma, \Phi_\Gamma)$ and $(\Lambda_\Sigma, \Psi_\Sigma)$, and denote their Gramian by $G(\lambda, \nu) = \langle m_\lambda, n_\nu\rangle$.

As we have mentioned, reconstruction formulas of representation systems will play a crucial role. Such formulas for curvelets can be found in [1], whereas for shearlets they can for example be found in [8], and in Appendix B of this paper. Reconstruction formulas (the high-frequency case) are generally of the form

$$f = \int_{\Lambda_\Sigma} \langle f, \tilde{n}_\nu\rangle \tilde{n}_\nu \, d\mu(\nu), \tag{15}$$

which is valid in (at least) the weak sense. In (15) we use $\tilde{n}_\nu$ to denote the elements of the dual system (which is assumed to also be a system of CPMs). The idea now is to use (15) to establish a relationship between frame coefficients associated to $\Gamma$ and $\Sigma$. Taking the inner product of (15) with $m_\lambda$ we immediately have

$$a_\lambda^{-k}\langle f, m_\lambda\rangle = a_\lambda^{-k} \int_{\Lambda_\Sigma} \langle f, \tilde{n}_\nu\rangle\langle n_\nu, m_\lambda\rangle d\mu(\nu). \tag{16}$$

Therefore, we would need bounds on the weighted $L^2$ norm of the coefficients in (16). We can actually show a bit more, namely, that we can consider not only the $L^2$ norm but rather any $L^p$ norm, for $p \in [1, \infty]$.

Notice first that the integral in (16) can be split up in two parts; an integral over $[0, a_0] \times \mathcal{N}$ and and integral over the complement. The first of these integrals is directly related to the square function $S_2^k$ through a bounded integral operator.

**Lemma 5.2.** *Let $\Gamma = \{m_\lambda : \lambda \in \Lambda_\Gamma\}$ and $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$ be two families of continuous parabolic molecules and take $N \in \mathbb{N}$ as given in Theorem 4.1. The operator $T : L^p([0, a_0] \times \mathcal{N}, \mu) \to L^p([0, a_0] \times \mathcal{M}, \mu)$, where $\mathcal{N}$ and $\mathcal{M}$ are open and bounded subsets of $[0, 2\pi) \times \mathbb{R}^2]$ and $p \in [1, \infty]$, defined via*

$$(Tu)(\nu) = \int_{[0,a_0]\times\mathcal{N}} \left(\frac{a_\lambda}{a_\nu}\right)^k G(\lambda, \nu) u(\lambda) d\mu(\lambda),$$

*is bounded (in $L^p$) provided the parametrisations of $\Gamma$ and $\Sigma$ are $N - k$ admissible.*

*Proof.* In order to show the boundedness of $T$ we will use Schur's test which says that $T$ is bounded, and the bound is given by

$$\|T\| \leqslant \left[\sup_\nu \int \left(\frac{a_\lambda}{a_\nu}\right)^k |G(\lambda, \nu)| d\mu(\lambda)\right]^{1/p} \left[\sup_\lambda \int \left(\frac{a_\lambda}{a_\nu}\right)^k |G(\lambda, \nu)| d\mu(\nu)\right]^{(p-1)/p}$$

provided that the right hand side of the expression is finite. To show that these integrals are indeed bounded we employ Theorem 4.1, which gives bounds on the integral kernel. In other words, we have

$$\int_{[0,a_0]\times\mathcal{N}} \left(\frac{a_\lambda}{a_\nu}\right)^k |G(\lambda,\nu)| d\mu(\lambda) \lesssim \int_{[0,a_0]\times\mathcal{N}} \left(\frac{a_\lambda}{a_\nu}\right)^k \left(\frac{a_M}{a_m}\right)^{-N} \left(1+a_M^{-1}d(\lambda,\nu)\right)^{-N} d\mu(\lambda). \tag{17}$$

Since $a_m \leqslant a_\lambda, a_\nu \leqslant a_M$ and $\left(1+a_M^{-1}d(\lambda,\nu)\right)^{-N} \leqslant \left(1+a_M^{-1}d(\lambda,\nu)\right)^{-(N-k)}$ we have

$$(17) \lesssim \int_{[0,a_0]\times\mathcal{N}} \left[\frac{a_M}{a_m}\left(1+a_M^{-1}d(\lambda,\nu)\right)\right]^{-(N-k)} d\mu(\lambda) \leqslant \int_{\Lambda_0} w(\lambda,\nu)^{-(N-k)} d\mu(\lambda) < \infty \tag{18}$$

The boundedness of the last expression follows from the admissibility of the parametrisation. The other integral is treated analogously. Hence, T is bounded. □

We are now ready to prove our first universality-type result. It will allow us to infer that assuming the frame coefficients of one CPM family are in a certain weighted $L^p$ space, then the frame coefficients of any other suitable CPM family are also in a weighted $L^p$ space. In order for the proof to work we need a further assumption on the parametrisation mapping $\Psi_\Sigma$, namely, we require $(\Psi_\Sigma)^{-1}$ to have a uniformly bounded Jacobian.

**Theorem 5.3.** *Let* $k \in \mathbb{N}$ *and take* $\Gamma = \{m_\lambda : \lambda \in \Lambda_\Gamma\}$ *and* $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$ *to be two families of continuous parabolic molecules, with parametrisations* $(\Phi_\Gamma, \Lambda_\Gamma)$ *and* $(\Psi_\Sigma, \Lambda_\Sigma)$, *such that* $\Sigma$ *admits a reproduction formula of the form (15), that the Jacobian of* $\Psi_\Sigma^{-1}$ *is uniformly bounded and that the conditions of Lemma 5.2 are satisfied. Take* $p \in [1,\infty]$. *Then if for some open and bounded neighbourhood* $\mathcal{N}$ *of* $(\theta_0, \boldsymbol{x}_0)$

$$a^{-k}_{\Psi_\Sigma^{-1}(\cdot)}\langle f, n_{\Psi_\Sigma^{-1}(\cdot)}\rangle \in L^p\left([0,a_0]\times\mathcal{M},\mu\right)$$

*holds then*

$$a^{-k}_{\Phi_\Gamma^{-1}(\cdot)}\langle f, m_{\Phi_\Gamma^{-1}(\cdot)}\rangle \in L^p\left([0,a_0]\times\mathcal{N},\mu\right)$$

*holds for some open and bounded neighbourhood* $\mathcal{M}$ *of* $(\theta_0, \boldsymbol{x}_0)$.

To reduce the notation we will write $a_\nu$ instead of $a_{\Phi_\Sigma{}^{-1}(\nu)}$. Analogous abbreviations will be used in other similar cases.

*Proof.* Take $\mathcal{M}$ to be an open and bounded neighbourhood of $(\theta_0, \mathbf{x}_0)$ such that $\mathtt{dist}(\mathcal{M},\mathcal{N}^c) > 0$ We have

$$a_\lambda^{-k}\langle f, m_\lambda\rangle = a_\lambda^{-k}\int_{\Lambda_\Sigma}\langle f, \tilde{n}_\nu\rangle\langle n_\nu, m_\lambda\rangle d\mu(\nu) = A_1 + A_2$$

where

$$A_1 = a_\lambda^{-k}\int_{\Psi_\Sigma^{-1}([0,a_0]\times\mathcal{N})}\langle f, \tilde{n}_\nu\rangle\langle n_\nu, m_\lambda\rangle d\mu(\nu)$$

$$A_2 = a_\lambda^{-k}\int_{\Psi_\Sigma^{-1}([0,a_0]\times\mathcal{N}^c)}\langle f, \tilde{n}_\nu\rangle\langle n_\nu, m_\lambda\rangle d\mu(\nu)$$

For $A_2$ we have

$$\begin{aligned} A_2 &= a_\lambda^{-k} \int_{\Psi_\Sigma^{-1}([0,a_0]\times\mathcal{N}^c)} \langle f, \tilde{n}_\nu \rangle \langle n_\nu, m_\lambda \rangle d\mu(\nu) \\ &\lesssim a_\lambda^{-k} \int_{[0,a_0]\times\mathcal{N}^c} \langle f, \tilde{n}_\nu \rangle \langle n_\nu, m_\lambda \rangle d\mu(\nu) \\ &\lesssim a_\lambda^{-k} \|f\|_{L^2(\mathbb{R}^2)} \left( \int_{[0,a_0]\times\mathcal{N}^c} w(\nu,\lambda)^{-2N} d\mu(\nu) \right)^{1/2}, \end{aligned}$$

where we used the boundedness of the Jacobian in the second line, and Theorem 4.1 in the third line. We can now write

$$w(\nu,\lambda) = \frac{a_M}{a_m}\left(1 + a_M^{-1} d(\nu,\lambda)\right) \geqslant a_m^{-1}\left(|\theta_\lambda - \theta_\nu|^2 + |\mathbf{b}_\lambda - \mathbf{b}_\nu|^2\right)$$

to get

$$A_2 \lesssim a_\nu^{-k} \left( \int_0^{a_0} a_m^{2N} \frac{da}{a^3} \right)^{1/2} \left( \int_{\mathcal{N}^c} \left(|\theta_\lambda - \theta_\nu|^2 + |\mathbf{b}_\lambda - \mathbf{b}_\nu|^2\right)^{-2N} \right)^{1/2}.$$

The second integral is clearly uniformly bounded as long as $\mathrm{dist}(\mathcal{M}, \mathcal{N}^c) \geqslant \epsilon > 0$. It follows

$$A_2 \lesssim a_\lambda^{-k} \left( \int_0^{a_0} a_m^{2N} \frac{da}{a^3} \right)^{1/2} \lesssim a_\lambda^{N-1-k}.$$

In other words, $A_2 = A_2(a_\lambda)$ will have a finite $L^p([0,a_0] \times \mathcal{M}, \mu)$ norm provided $N \geqslant k$ for $p = \infty$, and $N > \frac{2}{p} + k + 1$, for $p \in [1,\infty)$.

Turning our attention to $A_1$, notice first that due to the boundedness of the Jacobian we have

$$A_1 \lesssim (Tu)(\lambda),$$

with $T$ as in Lemma 5.2. Hence, by Lemma 5.2, it follows that $A_1$ also has a finite $L^p([0,a_0] \times \mathcal{M}, \mu)$ norm for $p \in [1,\infty]$ and the statement follows.

□

*Remark.* Notice that the argument for the boundedness of $A_2$ can be also seen as searching for a bound for the operator $\tilde{T}$, defined by

$$\left(\tilde{T}u\right)(\lambda) = \int_{[0,a_0]\times\mathcal{N}^c} \left(\frac{a_\nu}{a_\lambda}\right)^k G(\nu,\lambda) u(\nu) d\mu(\nu).$$

In that case, we would need a further assumption, namely, we need $a_\nu^k u(\nu) \in L^2$. Alternatively, we may assume $a_\nu^k u(\nu) \in L^q$, for some $q$. This would give $N > \frac{2}{p} + \frac{2}{q'} + k$, where $q'$ is such that $1/q + 1/q' = 1$.

Depending on the choice of $p$, we can now apply the theorem to different situations. For the first application we will address the question of the universality of microlocal Sobolev regularity with respect to CPMs. In other words, let take $p = 2$. The following theorem says that we can infer whether $f$ is in $H^k(\theta_0, \mathbf{x}_0)$ by looking at the $L^2$ condition on $S_2^k$, of the type (14), with respect to not just curvelets but also other families of continuous parabolic molecules.

**Theorem 5.4.** *Let $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$ be a family of CPMs of order $(R, M, N_1, N_2)$, and satisfying the assumptions of Theorem 5.3. Then $f$ is in $H^k(\theta_0, \boldsymbol{x}_0)$ if and only if*

$$a^{-k}_{\Psi^{-1}_\Sigma(\cdot)} \langle f, n_{\Psi^{-1}_\Sigma(\cdot)} \rangle \in L^2\left([0, a_0] \times \mathcal{N}, \mu\right), \tag{19}$$

*for some open neighbourhood $\mathcal{N}$ of $(\theta_0, \boldsymbol{x}_0)$.*

*Proof.* Assume (19) holds. Take $\Gamma = \{m_\lambda : \lambda \in \Lambda_\Gamma\}$ to be a family of second generation curvelets. It follows by Proposition 3.3 that $\Sigma$ is a family of CPMs of order $(\infty, \infty, R/2, R/2)$, where we can take any $R$. Therefore, we can apply Theorem 5.3, which gives

$$a^{-k}_{\Phi^{-1}_\Gamma(\cdot)} \langle f, m_{\Phi^{-1}_\Gamma(\cdot)} \rangle \in L^p\left([0, a_0] \times \mathcal{M}, \mu\right)$$

for some neighbourhood $\mathcal{M}$ of $(\theta_0, \mathbf{x}_0)$. Hence, the claim follows by Theorem 5.1. The converse follows trivially since if $f$ is in $H^k(\theta_0, \mathbf{x}_0)$ then by Theorem 5.1

$$a^{-k} \langle f, \gamma_{a\theta\mathbf{x}} \rangle \in L^2\left([0, a_0] \times \mathcal{N}, \mu\right).$$

The parametrisation mapping for curvelets is the identity. Hence, its Jacobian is trivially bounded and we can apply Theorem 5.3, yielding the statement. □

We can be more specific and take $\Sigma$ to be shearlets. Having Theorem 5.3 in mind, we need to ensure that the shearlet family we will use admits a representation formula.

**Corollary 5.5.** *Consider the family $\Sigma$ of cone-adapted, band-limited shearlets ([9, 10] or Appendix B). Then $f \in H^k(\theta_0, \boldsymbol{x}_0)$ if and only if*

$$\int_{\mathcal{N}_s} \left( \int_0^{a_0} |\langle f, \sigma_{as\boldsymbol{x}} \rangle|^2 a^{-2k} \frac{da}{a^3} \right) ds\, d\boldsymbol{x} < \infty$$

*where $\mathcal{N}_s$ is some neighbourhood of $(s_0, \boldsymbol{x}_0)$ and $s_0$ is the corresponding shearing parameter.*

*Proof.* Without loss of generality we can assume that $\theta \in \left(\frac{-\pi}{4}, \frac{\pi}{4}\right) \cup \left(\frac{3\pi}{4}, \frac{5\pi}{4}\right)$. The other case can be treated analogously. This is equivalent to $s = -\tan(\theta_0) \in (-1, 1)$, which means that we can find $\varepsilon > 0$ small enough such that

$$(s_0 - \varepsilon, s_0 + \varepsilon) \subseteq (-1, 1)$$

and we can use only horizontal shearlets for the analysis (this helps to simplify the expression for the parametrisation). Therefore, by B.2, the shearlet system is a system of parabolic molecules of arbitrary order, and is admissible for all $k > 2$. Furthermore, the Jacobian of $\Psi^{-1}_\Sigma$ is equal to $\frac{-2}{\cos(2\theta)+1}$, which is uniformly bounded for $\theta$ in $\left(\frac{-\pi}{4}, \frac{\pi}{4}\right) \cup \left(\frac{3\pi}{4}, \frac{5\pi}{4}\right)$. Thus, by Theorem 5.4 it suffices to show that

$$a^{-k}_{\Psi^{-1}_\Sigma(\cdot)} \langle f, \sigma_{\Psi^{-1}_\Sigma(\cdot)} \rangle \in L^2\left([0, a_0] \times \mathcal{N}, \mu\right).$$

Without loss of generality we can assume that $\mathcal{N}_s = (s_0 - \varepsilon, s_0 + \varepsilon) \times B_\varepsilon(\mathbf{x}_0)$. Let us define

$$\mathcal{N} = \arctan(s_0 - \varepsilon, s_0 + \varepsilon) \times B_\varepsilon(\mathbf{x}_0)$$

and notice that $\mathcal{N}$ is an open neighbourhood of $(\theta_0, \mathbf{x}_0)$. It follows

$$\begin{aligned}\left\| a^{-k}_{\Psi_\Sigma^{-1}(\cdot)} \langle f, \sigma_{\Psi_\Sigma^{-1}(\cdot)} \rangle \right\|^2 &= \int_{\mathcal{N}} \int_0^{a_0} \left| \langle f, \sigma_{\Psi_\Sigma^{-1}(a\theta\mathbf{x})} \right| a^{-2k} \frac{da}{a^3} d\theta d\mathbf{x} \\ &= \int_{\mathcal{N}_s} \int_0^{a_0} |\langle f, \sigma_{as\mathbf{x}} \rangle| \, a^{-2k} \, |\det J\Psi_\Sigma(as\mathbf{x})| \, \frac{da}{a^3} ds d\mathbf{x} \\ &\lesssim \int_{\mathcal{N}_s} \int_0^{a_0} |\langle f, \sigma_{as\mathbf{x}} \rangle| \, a^{-2k} \frac{da}{a^3} ds d\mathbf{x} < \infty.\end{aligned}$$

Here we used the fact that $|J\Psi_\Sigma(a, s, \mathbf{x})| = \left| \frac{1}{1+s^2} \right|$ is uniformly bounded. Therefore, by Theorem 5.4 the claim follows. □

It is important to note at this point that the argument used in the proofs of Corollary 5.5 and Theorem 5.4 would not go through for the non cone-adapted shearlets. The reason behind this lies in the fact the Jacobian of $\Psi_\Sigma^{-1}$ is not uniformly bounded on $[0, 2\pi]$. This agrees with the intuition, since regular shearlets exhibit a directional bias in the sense that the singularities on the $y$-axis can only be resolved as the shearing parameter tends to infinity. Cone-adapted shearlets on the other hand, always take the shearing parameters from a bounded set, which, in terms of Theorem 5.3, has the effect that the Jacobian will be bounded.

For the second application of Theorem 5.3 we will look at the resolution of the wavefront set. The notion of wavefront sets is related to the notion singular support, but whereas singular support only tells us where is a given function singular, wavefront set tries to give insight into how is the function singular by giving information about the direction of its singularities. There are various applications of wavefront sets, perhaps the most well known one comes in the study of the propagation of singularities of solutions of partial differential equations.

**Definition 5.2.** The wavefront set of a distribution $f$, denoted $WF(f)$ is the complement of the set of all points $(\theta_0, \mathbf{x}_0)$ such that there exists a smooth window function $\phi \in C_0^\infty$, $\phi(\mathbf{x}_0) \neq 0$ and an open cone $\mathcal{C}$ such that $\theta_0 \in \mathcal{C}$, with the property that for all $N \in \mathbb{N}$

$$|\widehat{\phi f}(\boldsymbol{\xi})| \leqslant C_N(1 + |\boldsymbol{\xi}|)^{-N}, \text{ for all } \boldsymbol{\xi} \in \mathcal{C}. \tag{20}$$

Condition (20) is typically called the rapid decay. It follows from the definition that if $(\theta_0, \mathbf{x}_0)$ is in $WF(f)$ that $\mathbf{x}_0$ is then in $\operatorname{sing\,supp}(f)$. Clearly, wavefront sets are only worth considering when the space is at least two dimensional, but in that case wavelets would not be of much use. This is because wavelets have no information regarding orientation, so it is immediately clear that they cannot resolve the wavefront set of a given function. On the other hand, directional representation systems are well suited for answering this questions and it is has been established that curvelets and shearlets can both resolve the wavefront set [6, 1].

Let us now generalise this fact in our framework of parabolic molecules. In terms of Theorem 5.3 this equates to taking $p$ to be $\infty$.

**Theorem 5.6.** *Let $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$ be a family of continuous parabolic molecules of order $(R, M, N_1, N_2)$, and $(\Psi_\Sigma, \Lambda_\Sigma)$ its parametrisation, satisfying the conditions of Theorem 5.3. The wavefront set of* f *is the complement of*

$$\mathcal{R}_\Sigma = \Bigg\{ (\theta_0, \boldsymbol{x}_0) : \textit{ for all } k \in \mathbb{N} \textit{ we have } |\langle f, n_{\Psi_\Sigma^{-1}(a,\theta,\boldsymbol{x})}\rangle| = \mathcal{O}(a^k) \textit{ as } a \to 0,$$

$$\textit{for some neighbourhood } \mathcal{N} \textit{ of } (\theta_0, \boldsymbol{x}_0) \Bigg\}. \quad (21)$$

*Proof.* The condition in the definition of $\mathcal{R}_\Sigma$ says that $(\theta_0, \mathbf{x}_0)$ is in $\mathcal{R}_\Sigma$ if

$$a^{-k}\langle f, n_{\Psi_\Sigma^{-1}(a,\theta,\mathbf{x})}\rangle \in L^\infty\left([0, a_\epsilon] \times \mathcal{N}, \mu\right), \text{ for all } k \in \mathbb{N},$$

for some open neighbourhood $\mathcal{N}$ of $(\theta_0, \mathbf{x}_0)$ and $0 < a_\epsilon < a_0$. On the other hand, by [1], second generation curvelets have the property that $WF(f)$ is the complement of $\mathcal{R}_\Gamma$, where by $\Gamma$ we denote the family of second generation curvelets, and $\mathcal{R}_\Gamma$ is defined analogously to (21). We can rewrite this as

$$a^{-k}\langle f, \gamma_{a\theta\mathbf{x}}\rangle \in L^\infty\left([0, a_\epsilon] \times \mathcal{N}, \mu\right).$$

Hence, the claim follows by Theorem 5.3, since replacing $a_0$ with $a_\epsilon$ has no bearing on the statement of the theorem. □

## 5.2 Representation Systems for Cone-Supported Functions

Results of the previous section relied heavily upon the requirement that we have representation families with good reconstruction formulas. To be more precise, we needed families of parabolic molecules that can reconstruct functions whose Fourier transform has a support which covers (possibly) the entire frequency plane. However, there are families of functions that admit a representation formula which is valid only for functions with frequency support inside a certain cone. Thus, the approach described in the previous subsection cannot be immediately applied because we have no means of controlling the decay of the frame coefficients outside of the cone in which the representation formula is valid. Our goal is to show that we can work around this problem and that analysis of the same type is still applicable. In order to get a grip on those bounds we will need stronger assumptions, as the conditions (1) are not sufficient. Also, the proofs will be a bit more technical.

Let us begin by describing the situation at hand. As was the case in previous chapters, we will only be concerned with the high frequency case. Define two cones in the frequency domain by

$$\mathcal{C}_{u,v} = \left\{ \xi \in \mathbb{R}^2 : |\xi_1| \geqslant u, \left|\frac{\xi_2}{\xi_1}\right| \leqslant v \right\}$$

and

$$\mathcal{C}^c_{u,v} = \left\{ \xi \in \mathbb{R}^2 : |\xi_2| \geqslant u, \left|\frac{\xi_1}{\xi_2}\right| \leqslant v \right\},$$

where $u, v > 0$. For a set $\mathcal{D}$ we define the following $L^2$ space

$$L^2(\mathcal{D})^{\vee} = \{f \in L^2(\mathbb{R}^2) L \operatorname{supp} \hat{f} \subseteq \mathcal{D}.$$

The standard situation is when $u = v = 1$. In that case we use denote $\mathcal{C} := \mathcal{C}_{u,v}$ and $\mathcal{C}^c := \mathcal{C}^c_{u,v}$. Furthermore, denote by $P_{\mathcal{C}}$ and $P_{\mathcal{C}^c}$ the respective frequency-domain projections onto cones $\mathcal{C}$ and $\mathcal{C}^c$. We consider a point $(\theta_0, \mathbf{x}_0)$ with open and bounded neighbourhoods $\mathcal{N}$ and $\mathcal{M}$, whose closures are contained inside the cone $\mathcal{C}$.

Let us assume that we have two families of parabolic molecules. The first family, $\Gamma = \{m_\lambda : \lambda \in \Lambda_\Gamma\}$, admits a representation formula

$$g = \int \langle g, m_\lambda \rangle \tilde{m}_\lambda \, d\mu(\lambda), \tag{22}$$

which holds at least in the weak sense, and is valid for $g \in L^2(\mathcal{C})^{\vee}$. A function $f \in L^2(\mathbb{R}^2)$ can hence be decomposed as

$$f = P_{\mathcal{C}} f + P_{\mathcal{C}^c} f.$$

The term $P_{\mathcal{C}} f$ can be treated in a manner entirely analogous to that of the previous section, when we had a representation formula which was valid on the entire $\mathbb{R}^2$. Hence, the problem lies in bounding the frame coefficients on $\mathcal{C}^c$, that is, outside of the cone in which we can represent $f$ using frame coefficients given through the members of $\Gamma$. To do that we will need stronger assumptions on the second family of parabolic molecules, $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$. Let us recall the Definition 2.2, which gives

$$n_\nu(\boldsymbol{\xi}) = a_\nu^{-3/4} \varphi^{(\nu)} \left( D_{1/a_\nu} R_{\theta_\nu} (\mathbf{x} - \mathbf{b}_\nu) \right).$$

Our first assumption is that functions $\varphi^\nu$ have $M$ vanishing moments in $x_1$ direction, that is, $\hat{\varphi}^{(\nu)}(\boldsymbol{\xi}) = \xi_1^M \hat{\rho}^{(\nu)}(\boldsymbol{\xi})$, with $\theta^\nu \in L^2$. This is a very common assumption. We also require a Sobolev condition on $\varphi^\nu$, namely, $\frac{\partial^L}{\partial^L x_2} \varphi \in L^2$. The last assumption is that all the $L^2$ norms are uniformly bounded. Curvelets, shearlets and a large number of other systems of note satisfy the boundedness assumption trivially, since they either have only finitely many generators (e.g. shearlets), or they have generators which are simple variations of one another (e.g. curvelets).

Let us briefly summarise the assumptions,

$$\begin{aligned} \hat{\varphi}^{(\nu)}(\boldsymbol{\xi}) = \xi_1^M \hat{\rho}^{(\nu)}(\boldsymbol{\xi}), &\text{ with } \sup_\nu \|\hat{\rho}^{(\nu)}\| < \infty \\ \frac{\partial^L}{\partial^L x_2} \varphi^{(\nu)} \in L^2, &\text{ with } \sup_\nu \left\| \frac{\partial^L}{\partial^L x_2} \varphi^{(\nu)} \right\| < \infty \end{aligned} \tag{23}$$

We are now ready to state and prove our result.

**Theorem 5.7.** *Let $\Gamma = \{m_\lambda : \lambda \in \Lambda_\Gamma\}$ and $\Sigma = \{n_\nu : \nu \in \Lambda_\Sigma\}$ be two families of continuous parabolic molecules satisfying assumptions (23), with parametrisations $(\Phi_\Gamma, \Lambda_\Gamma)$ and $(\Psi_\Sigma, \Lambda_\Sigma)$, such that $\Gamma$ admits a reproduction formula of the form (22) for functions*

$g$ *with* $\operatorname{supp}\hat{g} \subset \mathcal{C}$*. Take a function* $f \in L^2(\mathbb{R}^2)$*. Then if for some finite neighbourhood* $\mathcal{N}$ *of* $(\theta_0, \boldsymbol{x}_0)$ *we have*

$$a^{-k}_{\Phi_\Gamma^{-1}(\cdot)}\langle f, m_{\Phi_\Gamma^{-1}(\cdot)}\rangle \in L^2\left([0, a_0] \times \mathcal{N}, \mu\right)$$

*then*

$$a^{-k}_{\Psi_\Sigma^{-1}(\cdot)}\langle f, n_{\Psi_\Sigma^{-1}(\cdot)}\rangle \in L^2\left([0, a_0] \times \mathcal{M}, \mu\right)$$

*holds for some neighbourhood* $\mathcal{M}$ *of* $(\theta_0, \boldsymbol{x}_0)$*.*

*Proof.* We write

$$\left\|a_\nu^{-k}\langle f, n_\nu\rangle\right\| \leqslant \left\|a_\nu^{-k}\langle P_{\mathcal{C}} f, n_\nu\rangle\right\| + \left\|a_\nu^{-k}\langle P_{\mathcal{C}^c} f, n_\nu\rangle\right\|.$$

Using the assumptions we can apply the reconstruction formula on the first summand, $\left\|a_\nu^{-k}\langle P_{\mathcal{C}} f, n_\nu\rangle\right\|$, and furthermore, apply the same steps as in the proof of Theorem 5.3 to obtain the required bound. Therefore, what is left is to find bounds on the coefficients pertaining to $\mathcal{C}^c$.

Let us write $g = P_{\mathcal{C}^c} f$. We have

$$|\langle g, n_\nu\rangle| = a^{3/4}\int |\hat{g}(\xi)||\hat{\varphi}^{(\nu)}\left(D_a R_\theta \xi\right)|\, d\xi = a^{3/4}\int_{|\xi_2|<a^{-\alpha}} + a^{3/4}\int_{|\xi_2|>a^{-\alpha}} = I_1 + I_2, \tag{24}$$

where $1/2 < \alpha < 1$.

Since $\varphi^\nu$ has $M$ vanishing moments in the $x_1$ direction we have

$$I_1 \leqslant a^{3/4}\int_{|\xi_2|<a^{-\alpha}} a^M|\hat{g}(\xi)||\cos(\theta_\nu)\xi_1 - \sin(\theta_\nu)\xi_2|^M|\hat{\rho}^{(\nu)}\left(D_{a_\nu}R_{\theta_\nu}\xi\right)|\, d\xi,$$

For $\xi \in \mathcal{C}^c$ with $|\xi_2| < a^{-\alpha}$ we have

$$|\cos(\theta)\xi_1 - \sin(\theta)\xi_2| \leqslant \|\xi\| \lesssim a^{-\alpha}.$$

Hence, it follows

$$\begin{aligned} I_1 &\leqslant a^{M-M\alpha}\int a^{3/4}|\hat{g}(\xi)||\hat{\rho}^{(\nu)}\left(D_{a_\nu}R_{\theta_\nu}\xi\right)|\, d\xi = a^{M(1-\alpha)}\langle|\hat{g}|, |\hat{\rho}^{(\nu)}\left(D_{a_\nu}R_{\theta_\nu}\cdot\right)|\rangle \\ &\leqslant a^{M(1-\alpha)}\|f\|\|\hat{\rho}^{(\nu)}\|. \end{aligned}$$

Regarding $I_2$ we have

$$\begin{aligned} I_2 &= a^{3/4}\int_{|\xi_2|<a^{-\alpha}} |\hat{g}(\xi)||\hat{\varphi}^{(\nu)}\left(D_{a_\nu}R_{\theta_\nu}\xi\right)|\, d\xi \\ &= a^{-3/4}\int_{|a^{-1/2}\cos(\theta_\nu)\tilde{\xi}_2 - a^{-1}\sin(\theta_\nu)\tilde{\xi}_1|<a^{-\alpha}} |\hat{g}(R_{-\theta_\nu}D_{1/a_\nu}\tilde{\xi})||\hat{\varphi}(\xi)| d\xi \\ &= a^{-3/4}\int_{|a^{-1/2}\cos(\theta_\nu)\tilde{\xi}_2 - a^{-1}\sin(\theta_\nu)\tilde{\xi}_1|<a^{-\alpha}} |\hat{g}(R_{-\theta_\nu}D_{1/a_\nu}\tilde{\xi})||\tilde{\xi}_2|^{-L}\left|\left(\widehat{\frac{\partial^L}{\partial x_2^L}\varphi}\right)\right|(\tilde{\xi})\, d\tilde{\xi} \end{aligned}$$

Using the fact that $\theta_\nu \in [\theta_0 - \varepsilon, \theta_0 + \varepsilon] \subsetneq \mathcal{C}$ gives $|R_{\theta_\nu}\xi|_2 > C(\theta_0)|\xi_2|$. Thus, for $|\xi_2| > a^{-\alpha}$ we have

$$|\tilde{\xi}_2| = a^{1/2}|R_{\theta_\nu}\xi|_2 > C(\theta_0)a^{1/2-\alpha}.$$

This in turn yields

$$I_2 \leqslant a^{L(\alpha-1/2)} \|f\| \left\| \frac{\partial^L}{\partial x_2^L} \varphi \right\|.$$

Plugging it all in (24) we have

$$|\langle g, n_\nu \rangle| \lesssim a^{L(\alpha-1/2)+M(1-\alpha)}.$$

which yields

$$\left\| a_\nu^{-k} \langle P_{\mathcal{C}^c} f, n_\nu \rangle \right\| \lesssim 1.$$

Using the assumptions (23), the assumptions of the theorem, and the result on $\left\| a_\nu^{-k} \langle P_{\mathcal{C}} f, n_\nu \rangle \right\|$ the claim of the theorem follows. □

We have everything in place to state a result analogous to Corollary 5.5. The statement there depended on very special constructions of shearlets, where we had to pay special attention to the behaviour at the seam lines and ensuring that the dual frame was also a shearlet type system. Now we have the same result, but the assumptions are such that they admit more general shearlet constructions. That is, we have that shearlets resolve the microlocal Sobolev regularity of a function provided the shearlet generators are sufficiently smooth.

**Proposition 5.8.** *Let $\psi(\boldsymbol{x}) = \psi_1(x_1)\psi_2(x_2)$ be a generator of a shearlet system (7), such that $\psi_1, \psi_2$ are compactly supported and where $\psi_1$ has $M + R$ vanishing moments and Fourier decay of order $N_1$, and $\psi_2$ has Fourier decay of order $N_1 + N_2$. Assume the angle $\theta_0$ lies in the cone $\mathcal{C}_{u,\nu}$. Then $f$ is in $H^k(\theta_0, \boldsymbol{x}_0)$ if and only if*

$$\int_{\mathcal{N}_s} \left( \int_0^{a_0} |\langle f, \sigma_{as\boldsymbol{x}} \rangle|^2 a^{-2k} \frac{da}{a^3} \right) ds d\boldsymbol{x} < \infty \tag{25}$$

*where $\mathcal{N}_s$ is some open and bounded neighbourhood of $(s_0, \boldsymbol{x}_0)$ and $s_0$ is the corresponding shearing parameter.*

*Proof.* Let us first assume that (25) holds. Proposition 3.6 tells us that under the assumptions of this proposition the compact shearlets (7) constitute a family of CPMs of order $(R, M + N_1, N_1, N_2)$. It also tells us that (8) constitutes a frame for $\mathcal{L}^2(\mathcal{C}_{u,\nu}\check{)}$, and admits a representation formula

$$f = \frac{1}{C_\psi} \int_{\mathbb{R}^2} \langle f, T_{\mathbf{b}} W \rangle P_{\mathcal{C}_{u,\nu}} T_{\mathbf{b}} W d\mathbf{b} + \frac{1}{C_\psi} \int_{\mathbb{R}^2} \int_{s \in [-\Xi, \Xi]} \int_{a \in [0,1]} \langle f, \psi_{as\mathbf{b}} \rangle P_{\mathcal{C}_{u,\nu}} \psi_{as\mathbf{b}} da ds d\mathbf{b}.$$

Furthermore, the conditions (23) are satisfied by construction and since we have only one generator the supremum goes over a set which contains only one element. Applying the same arguments as in the proof of Corollary 5.5, we can apply Theorem 5.7 and the claim follows by Theorem 5.1. As in the previous proofs, the converse follows trivially. □

These conditions can be somewhat weakened, for example the condition on the separability of $\psi$ simplifies the computations, but can it be avoided.

# A Various Proofs

## A.1 Proof of Proposition 3.6

**Proposition A.1.** *Consider the shearlet system (7), such that* $\psi_1 \in C^{N_1}(\mathbb{R})$ *has compact support and* $M + R$ *anisotropic moments, and that* $\psi_2 \in C^{N_1+N_2}(\mathbb{R})$ *also has compact support, where* $M, R, N_1$ *and* $N_2$ *satisfy*

$$2(M+R) - 1/2 > N_1 + N_2 > M + R > 1/2.$$

*Then (7) constitutes a system of continuous parabolic molecules of order* $(R, M, N_1, N_2)$. *Furthermore, the system*

$$\{P_{\mathcal{C}_{u,v}}\psi_{as\boldsymbol{b}} : a \in [0,1] s \in [-\Xi, \Xi], \boldsymbol{b} \in \mathbb{R}^2\} \cup \{T_{\boldsymbol{b}} P_{\mathcal{C}_{u,v}} W : \boldsymbol{b} \in \mathbb{R}^2\}$$

*is a tight frame for* $L^2(\mathcal{C}_{u,v})^{\vee}$, *provided* $u > 0$, $\Xi > v$.

*Proof.* Without loss of generality, let us take $i = 0$ and drop the indices. Let us show that the generator $\psi$ satisfies the condition (5), that is, the bound

$$|\partial^\beta \hat{\psi}(\boldsymbol{\xi})| \lesssim \min\left(1, a + |\xi_1| + a^{1/2}|\xi_2|\right)^M \langle |\boldsymbol{\xi}| \rangle^{-N_1} \langle \xi_2 \rangle^{-N_2}.$$

We will split the proof into two cases. First, let $|\xi_1| \geqslant 1$. It suffices to show

$$|\xi_1^{N_1} \xi_2^{N_1+N_2} \partial^\beta \hat{\psi}(\boldsymbol{\xi})| \lesssim 1$$

since this implies

$$\langle \xi_1 \rangle^{N_1} \langle \xi_2 \rangle^{N_1+N_2} |\partial^\beta \hat{\psi}(\boldsymbol{\xi})| \lesssim 1.$$

Then the claim would follow from the inequality

$$\langle \xi_1 \rangle^{N_1} \langle \xi_2 \rangle^{N_1+N_2} \geqslant \langle |\boldsymbol{\xi}| \rangle^{N_1} \langle \xi_2 \rangle^{N_2}.$$

Notice that we have

$$\xi_1^{N_1} \xi_2^{N_1+N_2} \partial^\beta \hat{\psi}(\boldsymbol{\xi}) = C\left(\partial^{(N_1, N_1+N_2)} \mathbf{x}^\beta \psi(\mathbf{x})\right)^{\wedge}(\boldsymbol{\xi}),$$

where the constant $C$ depends on $N_1, N_2$ and $\beta$. Thus, since $\psi$ has compact support and is sufficiently smooth, it follows that $\partial^{(N_1,N_1+N_2)} \mathbf{x}^\beta \psi(\mathbf{x})$ is in $L^1(\mathbb{R})$. Therefore, $|\xi_1^{N_1} \xi_2^{N_1+N_2} \partial^\beta \hat{\psi}(\boldsymbol{\xi})| \lesssim 1$.

Now, let $|\xi_1| \leqslant 1$. Using the separability of $\psi$ we have

$$\mathbf{x}^\beta \psi(\mathbf{x}) = x_1^{\beta_1} \psi_1(x_1) x_2^{\beta_2} \psi_2(x_2),$$

and it follows

$$\left(\partial^{(N_1,N_1+N_2)} \mathbf{x}^\beta \psi(\mathbf{x})\right)^{\wedge}(\boldsymbol{\xi}) = \left(\partial^{N_1} x_1^{\beta_1} \psi_1(x_1)\right)^{\wedge}(\xi_1) \left(\partial^{N_1+N_2} x_2^{\beta_2} \psi_2(x_2)\right)^{\wedge}(\xi_2).$$

Using the same arguments as in the first case we can deal with the second term and show

$$\left| \left(\partial^{N_1+N_2} x_2^{\beta_2} \psi_2(x_2)\right)^{\wedge}(\xi_2) \right| \lesssim 1.$$

To deal with the first term we shall use the vanishing moments. The assumption that $\psi_1$ has $M+R$ vanishing moments implies that $\psi(x_1) = \frac{\partial^{M+R}}{\partial x_1^{M+R}}\rho(x_1)$, where $\hat{\rho} \in \mathcal{L}^2(\mathbb{R})$. Taking the Fourier transform we have $\hat{\psi}_1(\xi_1) = (\imath\xi_1)^{M+R}\hat{\rho}(\xi_1)$. Therefore,

$$\frac{\partial^n}{\partial \xi_1^n}(\partial^{\beta_1}\hat{\psi}_1)(0) = 0 \text{ for all } n = 0, \dots, M+R-\beta_1-1, \text{ where } \beta_1 \leqslant R.$$

Furthermore, $\partial^{\beta_1}\hat{\psi}_1$ is an analytic function, since it is the Fourier transform of a compactly supported, continuous function whose derivatives (of order up to $M$) vanish at 0. It follows that the function $\partial^{\beta_1}\hat{\psi}(\xi_1)$ is uniformly bounded for $|\xi_1| \leqslant 1$, while for small $\xi_1$ we have $|\partial^{\beta_1}\hat{\psi}_1(\xi_1)| \lesssim |\xi_1|^M$. Combining those two statements gives

$$|\partial^{\beta_1}\hat{\psi}_1(\xi_1)| \lesssim \min(1, |\xi_1|)^M \Rightarrow \left|\widehat{\left(\partial^{N_1} x_1^{\beta_1}\psi_1(x_1)\right)}(\xi_1)\right| \lesssim \min(1, |\xi_1|)^{M+N_1}.$$

Plugging all the estimates in we have

$$\widehat{\left(\partial^{(N_1, N_1+N_2)}\mathbf{x}^\beta\psi(\mathbf{x})\right)}(\xi) \lesssim \min(1, |\xi_1|)^{M+N_1}.$$

which implies

$$|\xi_1^{N_1}\xi_2^{N_1+N_2}\partial^\beta\hat{\psi}(\xi)| \lesssim \min(1, |\xi_1|)^{M+N_1},$$

Therefore, we have a system of shearlet molecules of order $(R, M+N_1, N_1, N_2)$, and by Proposition 3.4 it is a system of CPMs of the same order.

The second part of the claim follows from applying Theorem 4.9 of [8], which says that assuming $\psi$ has $M$ vanishing moments and Fourier decay of sufficient order in both of the variables, then it admits a representation formula. The assumptions of this proposition are such that these conditions are immediately satisfied. Namely, $\psi$ satisfies the Fourier decay conditions since $\psi_1$ is in $C^{N_1}(\mathbb{R})$ and $\psi_2$ is in $C^{N_1+N_2}(\mathbb{R})$, and it has $M+R$ vanishing moments by assumption. Therefore, (8) is a tight frame for $L^2(\mathbb{R}^2)$. □

## A.2 Additional Proofs For Section 3

**Proposition A.2.** *A system of curvelet molecules of regularity* $3R/2$ *constitutes a system of CPMs of order* $(R, R, R/2, R/2)$*, with canonical parametrisation.*

*Proof.* What we only have to do is to establish that the decay estimates (1) holds. Regarding the vanishing moments we can use [2] (similar claims can also be found in [11]), where it is stated that definition of curvelet molecules implies that (3) also holds for derivatives of $\varphi^{(\lambda)}$. On the other hand, since

$$|\partial^\beta\varphi^{(\lambda)}(\mathbf{x})| \leqslant C_N\langle|\mathbf{x}|\rangle^{-N}$$

holds for all $N$ and all $|\beta| \leqslant R$ we have

$$\mathbf{x}^\alpha\partial^\beta\varphi(\mathbf{x}) \in L^1(\mathbb{R}^2),$$

where $\alpha \in \mathbb{N}_0^2$ is an arbitrary multi-index, and $|\beta| \leqslant R$. Thus, for $|\beta| \leqslant R$ we have

$$\xi_1^{R/2}\xi_2^R \partial^\beta \hat{\varphi}(\boldsymbol{\xi}) = (\partial^{(R/2,R)}(\mathbf{x}^\beta \varphi(\mathbf{x})))\hat{}(\boldsymbol{\xi}) \in L^\infty(\mathbb{R}^2),$$

which gives

$$|\xi_1^{R/2}\xi_2^R \partial^\beta \hat{\varphi}(\boldsymbol{\xi})| \lesssim 1.$$

It follows

$$\langle|\boldsymbol{\xi}|\rangle^{R/2}\langle\xi_2\rangle^{R/2}|\partial^\beta \hat{\varphi}(\boldsymbol{\xi})| \lesssim 1.$$

□

**Lemma A.3.** *Let* d *be as in Definition 2.3. Then*

$$\int_{[0,2\pi)\times\mathbb{R}^2} \left[1 + q^{-1}d(\lambda,\nu)\right]^{-k} d\theta d\boldsymbol{b} \lesssim q^2,$$

*holds for all* $q > 0$

*Proof.* Denote $\delta\theta = \theta_\lambda - \theta_\nu$ and $\delta\mathbf{b} = \mathbf{b}_\lambda - \mathbf{b}_\nu$. For an arbitrary $q \in \mathbb{R}^+$ we have

$$\begin{aligned}
\int_{[0,2\pi)\times\mathbb{R}^2} \left[1 + q^{-1}d(\lambda,\nu)\right]^{-k} d\mathbf{b}d\theta &= \int_{[0,2\pi)\times\mathbb{R}^2} \left[1 + q^{-1}\left(|\delta\theta|^2 + |\delta\mathbf{b}|^2 + |\langle e_\lambda, \delta\mathbf{b}\rangle|\right)\right]^{-k} d\mathbf{b}d\theta \\
&= \int_{[0,2\pi)\times\mathbb{R}^2} \left[1 + q^{-1}\left(|\delta\theta|^2 + |\mathbf{b}|^2 + |\langle e_\alpha, \mathbf{b}\rangle|\right)\right]^{-k} \mathbf{b}d\theta \\
&= \int_{[\theta_\nu, 2\pi+\theta_\nu)\times\mathbb{R}^2} \left[1 + q^{-1}\left(\theta^2 + |\mathbf{b}|^2 + |\langle R^\tau_{\theta_\lambda+\theta_\nu}\, e_\lambda, \mathbf{b}\rangle|\right)\right]^{-k} d\mathbf{b}d\theta \\
&= \int_{[\theta_\nu, 2\pi+\theta_\nu)\times\mathbb{R}^2} \left[1 + q^{-1}\left(\theta^2 + |\mathbf{b}|^2 + |b_1|\right)\right]^{-k} d\mathbf{b}d\theta \\
&\leqslant \int_{\mathbb{R}\times\mathbb{R}^2} \left[1 + q^{-1}\left(\theta^2 + b_2^2 + |b_1|\right)\right]^{-k} d\mathbf{b}d\theta \\
&\leqslant \int_{\mathbb{R}\times\mathbb{R}^2} \left[1 + q^{-1}\left(q\theta^2 + qb_2^2 + q|b_1|\right)\right]^{-k} q^2 d\mathbf{b}d\theta \\
&\leqslant q^2 \int_{\mathbb{R}\times\mathbb{R}^2} \left[1 + \left(\theta^2 + b_2^2 + |b_1|\right)\right]^{-k} d\mathbf{b}d\theta \\
&\lesssim q^2.
\end{aligned}$$

□

## A.3 Lemmas used in the proof of Theorem 4.1

**Lemma A.4.** *Let* $\Psi(\boldsymbol{x}) = a^{-3/4}\psi\left(D_{1/a}R_\theta(\boldsymbol{x} - \boldsymbol{b})\right)$. *Then*

$$\hat{\Psi}(\boldsymbol{\xi}) = a^{3/4}e^{-2\pi\imath \boldsymbol{b}\cdot\boldsymbol{\xi}}\hat{\psi}\left(D_a R_\theta \boldsymbol{\xi}\right)$$

*Proof.* We have

$$\begin{aligned}
\hat{\Psi}(\boldsymbol{\xi}) &= a^{-3/4}\int_{\mathbb{R}^d}\psi\left(D_{1/a}R_\theta(\mathbf{u}-\mathbf{b})\right)e^{-2\pi\imath\mathbf{u}\cdot\boldsymbol{\xi}}d\mathbf{u} = \left\{\mathbf{x}=\mathbf{u}-\mathbf{b}\right\}\\
&= a^{-3/4}\int_{\mathbb{R}^d}\psi\left(D_{1/a}R_\theta\mathbf{u}\right)e^{-2\pi\imath(\mathbf{x}+\mathbf{b})\cdot\boldsymbol{\xi}}d\mathbf{x} = \left\{\mathbf{y}=D_{1/a}R_\theta\mathbf{u},\ d\mathbf{y}=|\det(D_{1/a}R_\theta)|d\mathbf{u}\right\}\\
&= a^{3/2-3/4}e^{-2\pi\imath\mathbf{b}\cdot\boldsymbol{\xi}}\int_{\mathbb{R}^d}\psi(\mathbf{y})e^{-2\pi\imath\left(\left(D_{1/a}R_\theta\right)^{-1}\mathbf{y}\right)\cdot\boldsymbol{\xi}}d\mathbf{y}\\
&= a^{3/4}e^{-2\pi\imath\mathbf{b}\cdot\boldsymbol{\xi}}\int_{\mathbb{R}^d}\psi(\mathbf{y})e^{-2\pi\imath\mathbf{y}\cdot\left(\left(D_{1/a}R_\theta\right)^{-\tau}\boldsymbol{\xi}\right)}d\mathbf{y}\\
&= a^{3/4}e^{-2\pi\imath\mathbf{b}\cdot\boldsymbol{\xi}}\hat{\psi}\left(\left(D_{1/a}R_\theta\right)^{-\tau}\boldsymbol{\xi}\right).
\end{aligned}$$

The last step is to use $\left(D_{1/a}R_\theta\right)^{-\tau} = D_aR_\theta$, which clearly holds since $D_a$ is a diagonal matrix and $R_\theta$ is an orthogonal matrix. Thus

$$\hat{\Psi}(\boldsymbol{\xi}) = a^{3/4}e^{-2\pi\imath\mathbf{b}\cdot\boldsymbol{\xi}}\hat{\psi}\left(D_aR_\theta\boldsymbol{\xi}\right).$$

□

**Lemma A.5.** *Consider the differential operator $\mathcal{L}_{\lambda,\nu}$ as defined in equation (10) The following holds*

$$\mathcal{L}_{\lambda,\nu}\left(e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\boldsymbol{x}}\right) = \alpha e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\boldsymbol{x}},$$

*where*

$$\alpha = 1+4\pi^2a_M^{-1}|\delta\mathbf{x}|^2+4\pi^2\frac{a_M^{-2}}{1+a_M^{-1}|\delta\theta|^2}\langle\boldsymbol{e}_\lambda,\delta\boldsymbol{x}\rangle^2.$$

*Consequently, we have*

$$\mathcal{L}_{\lambda,\nu}^{-k}\left(e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\boldsymbol{x}}\right) = \alpha^{-k}\left(e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\boldsymbol{x}}\right).$$

*Proof.* To simplify the notation, we will omit the indices $\lambda$ and $\mu$ throughout the proof. Recall the definition of the differential operator,

$$\mathcal{L}_{\lambda,\nu} = \mathcal{I} - a_M^{-1}\Delta - \frac{a_M^{-2}}{1+a_M^{-1}|\theta_\lambda-\theta_\nu|^2}\frac{\partial^2}{\partial e_\lambda^2}.$$

Hence, $\mathcal{L}_{\lambda,\nu}$ is a sum of three operators and we can deal with them one at a time. The identity operator is trivial. We have

$$\mathcal{I}\left(e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}\right) = e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}.$$

Laplace operator yields

$$\begin{aligned}
\Delta\left(e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}\right) &= \left(\frac{\partial^2}{\partial\xi_1^2}+\frac{\partial^2}{\partial\xi_2^2}\right)\left(e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}\right)\\
&= \frac{\partial}{\partial\xi_1}\left((-2\pi\imath\delta\mathbf{x}_1)e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}\right)+\frac{\partial}{\partial\xi_2}\left((-2\pi\imath\delta\mathbf{x}_2)e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}\right)\\
&= \left((-2\pi\imath\delta\mathbf{x}_1)^2+(-2\pi\imath\delta\mathbf{x}_2)^2\right)e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}} = -4\pi^2|\delta\mathbf{x}|^2e^{-2\pi\imath\boldsymbol{\xi}\cdot\delta\mathbf{x}}.
\end{aligned}$$

Lastly,

$$\frac{\partial^2}{\partial \mathbf{e}_\lambda} e^{-2\pi\imath\xi\cdot\delta\mathbf{x}} = \frac{\partial}{\partial \mathbf{e}_\lambda}\left(\nabla e^{-2\pi\imath\xi\cdot\delta\mathbf{x}}\cdot \mathbf{e}_\lambda\right) = \frac{\partial}{\partial \mathbf{e}_\lambda}\left(-2\pi\imath\delta\mathbf{x}\cdot\mathbf{e}_\lambda e^{-2\pi\imath\xi\cdot\delta\mathbf{x}}\right)$$
$$= \left(-2\pi\delta\mathbf{x}\cdot\mathbf{e}_\lambda\right)^2 e^{-2\pi\imath\xi\cdot\delta\mathbf{x}} = -4\pi^2\langle \mathbf{e}_\lambda, \delta\mathbf{x}\rangle^2 e^{-2\pi\imath\xi\cdot\delta\mathbf{x}}.$$

Thus, combining all of the above, we have

$$\mathcal{L}_{\lambda,\nu}\left(e^{-2\pi\imath\xi\cdot\delta\mathbf{x}}\right) = \left(1 + \frac{4\pi^2|\delta\mathbf{x}|^2}{a_M} + 4\pi^2\frac{a_M^{-2}}{1+a_M^{-1}|\delta\theta|^2}\langle \mathbf{e},\delta\mathbf{x}\rangle^2\right)e^{-2\pi\imath\xi\cdot\delta\mathbf{x}},$$

as desired. □

**Lemma A.6.** *Define*

$$S_{\lambda,M,N_1,N_2}(r,\phi) = \min\left(1, a_\lambda(1+r)\right)^M\left(1 + a_\lambda^{-1/2}|\sin(\phi+\theta_\lambda)|\right)^{-N_2}(1+a_\lambda r)^{-N_1}.$$

*The following holds*

$$\left|\mathcal{L}^k_{\lambda,\nu}\left(\hat{\psi}_\lambda\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{\phi}_\nu\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}\right)\right| \lesssim S_{\lambda,M-N_2,N_1,N_2}(\xi)S_{\nu,M-N_2,N_1,N_2}(\xi),$$

*where $\mathcal{L}_{\lambda,\nu}$ is the differential operator defined in (10).*

*Proof.* Let us denote

$$A_{\lambda,M,N_1,N_2}(r,\phi) = \min\left(1, a_\lambda(1+r)\right)^M(1+a_\lambda r)^{-N_1}\left(1 + a_\lambda^{1/2}|\sin(\phi+\theta_\lambda)|\right)^{-N_2}.$$

Notice the difference between the definitions of $A$ and $S$.

By Lemma A.7 it is sufficient to show

$$\left|\mathcal{L}^k_{\lambda,\nu}\left(\hat{\psi}_\lambda\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{\phi}_\nu\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}\right)\right| \lesssim A_{\lambda,M,N_1,N_2}(r,\varphi)A_{\nu,M,N_1,N_2}(r,\varphi).$$

We will construct our argument by induction in $k$. By Lemma A.10 the expression $\mathcal{L}^k_{\lambda,\nu}\left(\hat{\psi}_\lambda\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{\phi}_\nu\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}\right)$ can be written as a finite linear combination of terms of the form $\hat{c}_\lambda\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{d}_\nu\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}$ where $c_\lambda$ and $d_\nu$ satisfy (1) for $(R-2, M, N_1, N_2)$. Therefore, applying $\mathcal{L}^2_{\lambda,\nu}$ to $\hat{\psi}_\lambda\overline{\hat{\phi}_\nu}$ means applying $\mathcal{L}_{\lambda,\nu}$ to each of the terms $\hat{c}_\lambda(\cdot)\overline{\hat{d}_\nu(\cdot)}$, where in the process we lose two degrees of smoothness. Hence, if $k \leqslant R/2$ we conclude that the expression $\mathcal{L}^k_{\lambda,\nu}\left(\hat{\psi}_\lambda\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{\phi}_\nu\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}\right)$ can be written as a (finite) linear combination of terms of the form $\hat{c}_\lambda(\cdot)\overline{\hat{d}_\nu(\cdot)}$, where $c_\lambda$ and $d_\nu$ satisfy bounds of the form

$$|\hat{c}_\lambda(\xi)| \lesssim \min\left(1, a_\lambda + |\xi_1| + a_\lambda^{1/2}|\xi_2|\right)^M\langle|\xi|\rangle^{-N_1}\langle\xi_2\rangle^{-N_2}.$$

Now, since by Lemma A.10 the coefficients of the linear combination are uniformly bounded, we have

$$\left|\mathcal{L}^{k}_{\lambda,\nu}\left(\hat{\psi}_\lambda\left(D_{a_\lambda}R_{\theta_\lambda}\xi\right)\overline{\hat{\phi}_\nu\left(D_{a_\nu}R_{\theta_\nu}\xi\right)}\right)\right| \lesssim \tag{26}$$

$$\min\left(1, a_\lambda + |(D_{a_\lambda}R_{\theta_\lambda}\xi)_1| + a_\lambda^{1/2}|(D_{a_\lambda}R_{\theta_\lambda}\xi)_2|\right)^M \langle|(D_{a_\lambda}R_{\theta_\lambda}\xi)|\rangle^{-N_1}\langle(D_{a_\lambda}R_{\theta_\lambda}\xi)_2\rangle^{-N_2}$$

$$\cdot \min\left(1, a_\nu + |(D_{a_\nu}R_{\theta_\nu}\xi)_1| + a_\nu^{1/2}|(D_{a_\nu}R_{\theta_\nu}\xi)_2|\right)^M \langle|(D_{a_\nu}R_{\theta_\nu}\xi)|\rangle^{-N_1}\langle(D_{a_\nu}R_{\theta_\nu}\xi)_2\rangle^{-N_2}.$$

Now $D_a R_\theta \xi = \left(a(\cos(\theta)\xi_1 - \sin(\theta)\xi_2), a^{1/2}(\sin(\theta)\xi_1 + \cos(\theta)\xi_2)\right)^\tau$. Writing $\xi$ in polar coordinates as $\xi = (r\cos(\varphi), r\sin(\varphi))$ it follows

$$D_a R_\theta \xi = \begin{pmatrix} ar\cos(\theta+\varphi) \\ a^{1/2} r \sin(\theta+\varphi) \end{pmatrix}$$

Plugging those expressions in (26) yields the desired statement.

□

**Lemma A.7.** *For every* $0 \leqslant L \leqslant N_2$ *we have*

$$\min\left(1, a_\lambda(1+r)\right)^M (1+a_\lambda r)^{-N_1}\left(1 + a_\lambda^{1/2} r|\sin(\phi+\theta_\lambda)|\right)^{-N_2} \lesssim S_{\lambda,M-L,N,L}(r,\phi).$$

*Proof.* We will omit indices $\lambda$ throughout the proof. Since $L \leqslant N_2$ it follows

$$\left(1 + a^{1/2} r|\sin(\phi+\theta_\lambda)|\right)^{-N_2} \leqslant \left(1 + a^{1/2} r|\sin(\phi+\theta_\lambda)|\right)^{-L}.$$

Therefore,

$$\frac{\min(1, a(1+r))^M}{(1+ar)^{N_1}\left(1+a^{1/2}r|\sin(\phi+\theta)|\right)^{N_2}} \leqslant \frac{\min(1, a(1+r))^{M-L}}{(1+ar)^{N_1}}\left(\frac{\min(1, a(1+r))}{1+a^{1/2}r|\sin(\phi+\theta)|}\right)^L.$$

What is left for us to show is the following inequality

$$\frac{\min(1, a(1+r))}{1+a^{1/2}r|\sin(\phi+\theta)|} \lesssim \frac{1}{1+a^{-1/2}|\sin(\phi+\theta)|}.$$

We have the following cases

1. $r \geqslant a^{-1}$. Then $\min(1, a(1+r)) = 1$ and $a^{1/2}r \geqslant a^{-1/2}$. Thus

$$\frac{\min(1, a(1+r))}{1+a^{1/2}r|\sin(\phi+\theta)|} \leqslant \frac{1}{1+a^{-1/2}|\sin(\phi+\theta)|}.$$

2. $a_0^{-1} \leqslant r \leqslant a^{-1}$. We distinguish between two further cases.

   2A. $\min(1, a(1+r)) = 1$.
   This implies $a \geqslant \frac{1}{1+r}$. Therefore, because $(1+r)^{-1} \geqslant \frac{\min(1,a_0^{-1})}{2} r^{-1}$, we have $1 \lesssim ar$ and consequently $a^{-1/2} \lesssim a^{1/2}r$ which gives

$$\frac{\min(1, a(1+r))}{1+a^{1/2}r|\sin(\phi+\theta)|} = \frac{1}{1+a^{1/2}r|\sin(\phi+\theta)|} \lesssim \frac{1}{1+a^{-1/2}|\sin(\phi+\theta)|}.$$

2B. $\min(1, a(1+r)) = a(1+r)$.

$$\frac{\min(1, a(1+r))}{1+a^{1/2}r|\sin(\phi+\theta)|} = \frac{a(1+r)}{1+a^{1/2}r|\sin(\phi+\theta)|} = \frac{1+r}{r}\frac{1}{a^{-1}/r+a^{-1/2}|\sin(\phi+\theta)|}$$
$$\lesssim \frac{1}{1+a^{-1/2}|\sin(\phi+\theta)|}$$

because $r^{-1}+1 \leqslant (a_0+1)$.

3. $r \leqslant a_0^{-1}$. Here we clearly have

$$\frac{\min(1, a(1+r))}{1+a^{1/2}r|\sin(\phi+\theta)|} \leqslant \min(1, a(1+r)) \leqslant a(1+a_0^{-1}) \lesssim a.$$

Therefore, we want to show $a \lesssim \frac{1}{1+a^{-1/2}r|\sin(\phi+\theta)|}$. To this end, let us define $\mathbf{u} = (1, \sin(\phi+\theta))^\tau, \mathbf{v} = (1, a^{-1/2})^\tau$. Using Cauchy-Schwarz and $AM - GM$ inequalities we have

$$1+a^{-1/2}|\sin(\phi+\theta)| = |\langle \mathbf{u}, \mathbf{v}\rangle| \leqslant \|\mathbf{u}\|\|\mathbf{v}\| \leqslant \sqrt{2(1+a^{-1})} \leqslant \frac{3+a^{-1}}{2} \lesssim a^{-1}$$

Therefore $1+a^{-1/2}|\sin(\phi+\theta)| \lesssim a^{-1}$ which is what we wanted to show.

□

**Lemma A.8.** *For $M > A - 5/4$, $N_2 \geqslant B$ and $N_1 \geqslant A + 3/4$, we have*

$$(a_\lambda a_\nu)^{3/4}\int_{\mathbb{R}^2} S_{\lambda,M,N_1,N_2}(\xi)S_{\mu,M,N_1,N_2}(\xi)d\xi \lesssim \left(\frac{a_M}{a_m}\right)^{-A}\left(1+a_M^{-1/2}|\theta_\lambda-\theta_\nu|\right)^{-B}. \tag{27}$$

*Proof.* Without loss of generality we can assume $a_\lambda \leqslant a_\nu$, and denote

$$I_\phi = \left[\int_{\mathbf{T}}\left(1+a_\lambda^{-1/2}|\sin(\phi+\theta_\lambda)|\right)^{-N_2}\left(1+a_\nu^{-1/2}|\sin(\phi+\theta_\nu)|\right)^{-N_2}d\phi\right]$$

Writing (27) in polar coordinates gives

$$\int_{\mathbb{R}^2} S_{\lambda,M,N_1,N_2}(\xi)S_{\mu,M,N_1,N_2}(\xi)d\xi =$$
$$= I_\phi\int_{\mathbb{R}^+}\min(1, a_\lambda(1+r))^M\min(1, a_\nu(1+r))^M(1+a_\lambda r)^{-N_1}(1+a_\nu r)^{-N_1}r dr.$$

With the help of Lemma A.9 we have

$$I_\phi \lesssim a_\lambda^{1/2}\left(1+a_\nu^{-1/2}|\delta\theta|\right)^{-N_2}.$$

Therefore

$$(27) \lesssim \mathcal{S}(a_\nu/a_\lambda)^{3/4}\left(1+a_\nu^{-1/2}|\delta\theta|\right)^{-N_2}$$

where

$$\mathcal{S} = a_\nu^2\int_{\mathbb{R}^+}\min(1, a_\lambda(1+r))^M\min(1, a_\nu(1+r))^M(1+a_\lambda r)^{-N_1}(1+a_\nu r)^{-N_1}r dr$$

Hence, it remains to show $\mathcal{S} \lesssim (a_\nu/a_\lambda)^{-A-3/4}$. In order to show this, let us write $\mathbb{R}^+$ as

$$(0, \max(0, a_\nu^{-1}-1)) \cup \left(\max(0, a_\nu^{-1}-1), \max\left(0, a_\lambda^{-1}-1\right)\right) \cup \left(\max\left(0, a_\lambda^{-1}-1\right), \infty\right)$$

and split the integral in the definition of $\mathcal{S}$ accordingly. Therefore, we now write $\mathcal{S} = a_\nu{}^2(I_1 + I_2 + I_3)$.

Without loss of generality we can assume $\max(0, a_\nu^{-1}-1) = a_\nu^{-1}-1$ and $\max(0, a_\lambda^{-1}-1) = a_\lambda^{-1}-1$. It follows

$$\begin{aligned}
&\int_0^{a_\nu^{-1}-1} \min(1, a_\lambda(1+r))^M \min(1, a_\nu(1+r))^M \underbrace{(1+a_\lambda r)^{-N_1}}_{\leqslant 1} \underbrace{(1+a_\nu r)^{-N_1}}_{\leqslant 1} r dr \\
&\quad \leqslant \int_0^{a_\nu^{-1}-1} \min(1, a_\lambda(1+r))^M \min(1, a_\nu(1+r))^M r dr \\
&\quad \leqslant (a_\nu a_\lambda)^M \int_0^{a_\nu^{-1}-1} (1+r)^{2M+1} dr \leqslant (a_\lambda a_\nu)^M \frac{(1+r)^{2M+2}}{2M+2}\Big|_0^{a_\nu^{-1}-1} \lesssim a_\lambda^M a_\nu^{-M-2}
\end{aligned}$$

For $I_2$ we have

$$\begin{aligned}
&\int_{a_\nu^{-1}-1}^{a_\lambda^{-1}-1} \min(1, a_\lambda(1+r))^M \min(1, a_\nu(1+r))^M (1+a_\lambda r)^{-N_1} \underbrace{(1+a_\nu r)^{-N_1}}_{\leqslant 1} r dr \\
&\quad \leqslant \int_{a_\nu^{-1}-1}^{a_\lambda^{-1}-1} (a_\lambda(1+r))^M (1+a_\lambda r)^{-N_1} (1+r) dr
\end{aligned}$$

Since $a_\lambda \lesssim 1$ we have $a_\lambda(1+r) \lesssim 1 + a_\lambda r$, thus

$$\begin{aligned}
I_2 &\lesssim a_\lambda^M a_\nu^{-N_1} \int_{a_\nu^{-1}-1}^{a_\lambda^{-1}-1} (1+r)^{M-N_1+1} dr \lesssim a_\lambda^M a_\nu^{-N_1} \frac{(1+r)^{M+2-N_1}}{M+2-N_1}\Big|_{a_\nu^{-1}-1}^{a_\lambda^{-1}-1} \\
&\lesssim a_\lambda^M a_\nu^{-N_1} \left(a_\lambda^{N_1-2-M} - a_\nu^{N_1-2-M}\right) \lesssim a_\lambda^{N_1-2} a_\nu^{-N_1}
\end{aligned}$$

Lastly, for $I_3$ we have

$$\begin{aligned}
I_3 &= \int_{a_\lambda^{-1}-1}^{\infty} (1+a_\lambda r)^{-N_1} (1+a_\nu r)^{-N_1} r dr \leqslant (a_\lambda a_\nu)^{-N_1} \int_{a_\lambda^{-1}-1}^{\infty} (1+r)^{1-2N_1} dr \\
&\leqslant (a_\lambda a_\nu)^{-N_1} \frac{(1+r)^{2-2N_1}}{2-2N_1}\Big|_{a_\lambda^{-1}-1}^{\infty} \lesssim a_\lambda^{N_1-2} a_\nu^{-N_1}
\end{aligned}$$

Combining the bounds for $I_1, I_2$ and $I_3$, and due to the assumptions in the statement of the lemma, we have

$$\mathcal{S} \lesssim a_\nu^2 \left(a_\lambda^M a_\nu^{M-2} + a_\lambda^{N_1-2} a_\nu^{-N_1}\right) = (a_\lambda/a_\nu)^M + (a_\lambda/a_\nu)^{N_1-2} \lesssim (a_\lambda/a_\nu)^{-A-3/4}$$

which is what we wanted to show. ☐

**Lemma A.9.** *For $a_\lambda \leqslant a_\nu$ and a positive integer $N$ the following holds*

$$\int_{\mathbf{T}} \left(1 + a_\lambda^{-1/2}|\sin(\phi+\theta_\lambda)|\right)^{-N_2} \left(1 + a_\nu^{-1/2}|\sin(\phi+\theta_\nu)|\right)^{-N_2} d\phi \lesssim a_\lambda^{1/2} \left(1 + a_\nu^{-1/2}|\delta\theta|\right)^{-N_2}.$$

*Proof.* Application of Lemma 5.2 from [3], with $\phi = \phi + \theta_\lambda$ and $\theta = -\delta\theta$, yields

$$I_\phi \lesssim \max(a_\lambda^{-1/2}, a_\nu^{-1/2})^{-1} \left(1 + \min(a_\lambda^{-1/2}, a_\nu^{-1/2})|\theta_\nu - \theta_\lambda|\right)^{-N_2}.$$

Since, $a_\lambda \leqslant a_\nu$ we have $a_\lambda^{-1/2} \geqslant a_\nu^{-1/2}$ and it follows

$$I_\phi \lesssim a_\lambda^{1/2} \left(1 + a_\nu^{-1/2}|\delta\theta|\right)^{-N_2}.$$

□

**Lemma A.10.** *Let $\psi_\lambda$ and $\phi_\nu$ be functions that satisfy the conditions (1) for $(R, M, N_1, N_2)$. Then the expression $\mathcal{L}_{\lambda,\nu}\left(\hat{\psi}_\lambda(D_{a_\lambda} R_{\theta_\lambda} \xi)\overline{\hat{\phi}_\nu(D_{a_\nu} R_{\theta_\nu} \xi)}\right)$ can be written as a finite linear combination of terms of the form $\hat{c}_\lambda(D_{a_\lambda} R_{\theta_\lambda} \xi)\overline{\hat{d}_\nu(D_{a_\nu} R_{\theta_\nu} \xi)}$ such that $\hat{c}_\lambda$ and $\hat{d}_\nu$ satisfy (1) with $(R-2, M, N_1, N_2)$.*

*Proof.* This is the content of Lemma 5.5 of [3]. □

# B A Reconstruction Formula for Shearlets

In this last section we present a new construction of a shearlet family that admits a reconstruction formula. As we have stated in the previous section, having such a family is essential for a full exploitation of our framework. Our construction is motivated by [12, 10].

Let us start out by taking a mother shearlet $\psi$ with Fourier decay of order $L_1$ in the first variable, $M$ vanishing moments in $x_1$ direction, that is, $\psi = \left(\frac{\partial}{\partial x_1}\right)^M \vartheta$ with $\vartheta \in L^2(\mathbb{R}^2)$, and furthermore, we assume $\vartheta$ has Fourier decay of order $L_1$ in the second variable. Furthermore, $\psi$ has the following properties

$$\begin{aligned} \hat{\psi}^1(\boldsymbol{\xi}) &= \hat{\psi}_1(\xi_1)\hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right), \\ \operatorname{supp}\hat{\psi}_1 &\subset \left[-\frac{1}{4}, -\frac{1}{32}\right] \cup \left[\frac{1}{32}, \frac{1}{4}\right], \\ \operatorname{supp}\hat{\psi}_2 &\subset \left[-\frac{4}{3}, \frac{4}{3}\right]. \end{aligned}$$

Furthermore, let us define $\psi^2(x_1, x_2) = \psi^1(x_2, x_1)$. We can now define a shearlet system through

$$\psi^i_{asb}(\mathbf{x}) = a^{-\frac{3}{4}}\psi^i(T^i(\mathbf{x} - \mathbf{b})),$$

where $T^i = D^i_{1/a} S^i_s$ for $i = 1, 2$, with $D^1_{1/a} = \operatorname{diag}(1/a, 1/\sqrt{a})$, $D^2_{1/a} = \operatorname{diag}(1/\sqrt{a}, 1/a)$, and $S^1_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ and $S^2_s = (S^1_s)^\tau$.

Now we need to find partition functions which we can use to patch together these shearlet systems in such a way that they still form a system of parabolic molecules and more importantly, that they admit a useful reconstruction formula. The construction we are about to describe is somewhat cumbersome in notation. To begin, let us define the frequency cones the shearlets will be associated with. Horizontal cones are defined as follows

$$\mathcal{C}^1 : = \left\{ \boldsymbol{\xi} : |\xi_1| \geqslant \frac{1}{8}, \left|\frac{\xi_2}{\xi_1}\right| \leqslant \frac{4}{3} \right\}, \ \bar{\mathcal{C}}^1 : = \left\{ \boldsymbol{\xi} : |\xi_1| \geqslant \frac{1}{4}, \left|\frac{\xi_2}{\xi_1}\right| \leqslant \frac{5}{4} \right\}.$$

Vertical cones $\mathcal{C}^2$ and $\bar{\mathcal{C}}^2$ are obtained by rotating the corresponding horizontal cones through an angle of $\pi/2$. We also have two low-frequency boxes

$$\mathcal{C}^0 : = \{\|\boldsymbol{\xi}\|_\infty \leqslant 1\} \text{ and } \bar{\mathcal{C}}^0 : = \left\{\|\boldsymbol{\xi}\|_\infty \leqslant \frac{3}{4}\right\}.$$

Let us now define the partition functions we shall work with. Take

$$\begin{aligned} &\gamma^1(\boldsymbol{\xi}) = g_1(\xi_1) g_2\left(\frac{\xi_2}{\xi_1}\right), \quad \gamma^2(\boldsymbol{\xi}) = \gamma^1(\xi_2, \xi_2) \\ &\text{with } \operatorname{supp} g_1 \subset \left[\frac{1}{8}, \infty\right), \quad \operatorname{supp} g_2 \subset \left[-\frac{4}{3}, \frac{4}{3}\right], \end{aligned}$$

where $g_1$ and $g_2$ are smooth and real-valued, Similarly,

$$\chi^1(\xi) = h_1(\xi_1)h_2\left(\frac{\xi_2}{\xi_1}\right), \quad \chi^2(\xi) = h_1(\xi_2)\left(1 - h_2\left(\frac{\xi_2}{\xi_1}\right)\right), \quad \chi^0 = 1 - \chi^1 - \chi^2,$$

$$\text{with } \operatorname{supp} h_1 \subset \left[\frac{1}{4}, \infty\right), \quad h_1|_{[\frac{1}{2},\infty)} \equiv 1, \text{ and } \operatorname{supp} h_2 \subset \left[-\frac{5}{4}, \frac{5}{4}\right], \quad h_2|_{[-\frac{4}{5},\frac{4}{5}]} \equiv 1, \tag{28}$$

where $h_1$ and $h_2$ are non-negative and $h_1$ is non-decreasing.

It follows straight from the definitions that $\operatorname{supp} \gamma^i \subset \mathcal{C}^i$ and $\operatorname{supp} \chi^i \subset \bar{\mathcal{C}}^i$ for $i = 0, 1, 2$.

**Lemma B.1.** *For $i = 1, 2$ we have*

$$\left\| \chi^i \left( (D^i_{1/a} S^i_s)^\tau \xi \right) \right\|_{C^N(\operatorname{supp} \hat{\psi}^i)} \leqslant \gamma_N, \ N \in \mathbb{N},$$

*and*

$$\left\| \chi^0 \left( (D^i_{1/a} S^i_s)^\tau \xi \right) \right\|_{C^N(\operatorname{supp} \hat{\psi}^i)} \leqslant \tilde{\gamma}_N, \ N \in \mathbb{N}.$$

*Analogous estimates hold for $\gamma^i$ and $\frac{\chi^i}{\gamma^i}$.*

*Proof.* Take $i = 1$. We have

$$\chi^1 \left( (D^1_{1/a} S^1_s)^\tau \xi \right) = h_1(a^{-1}\xi_1) h_2 \left( s + a^{1/2} \frac{\xi_2}{\xi_1} \right).$$

Since $\xi \in \operatorname{supp} \hat{\psi}^1$ we have

$$\frac{a^{-1}}{32} \leqslant |a^{-1}\xi_1| \leqslant \frac{a^{-1}}{4}.$$

Therefore, if $a \leqslant 2^{-4}$ then $h_1(a^{-1}\xi_1) = 1$ and

$$\chi^1(a^{-1}\xi_1, a^{-1}s\xi_1 + a^{-1/2}\xi_2) = h_2\left( s + a^{1/2} \frac{\xi_2}{\xi_1} \right).$$

Hence, for a multi-index $\boldsymbol{\alpha}$ we have

$$\left| \partial^{\boldsymbol{\alpha}} h_2 \left( s + a^{1/2} \frac{\xi_2}{\xi_1} \right) \right| \lesssim a^{\frac{|\alpha|}{2}} \xi_1^{-1-|\boldsymbol{\alpha}|} \xi_2^{|\boldsymbol{\alpha}|} \sup_{\boldsymbol{\beta} \leqslant \boldsymbol{\alpha}} \left| \partial^{\boldsymbol{\beta}} h_2 \left( s + a^{1/2} \frac{\xi_2}{\xi_1} \right) \right| \lesssim 1,$$

which follows from the properties of $\operatorname{supp} h_2$. What is left is to address the case $a > 2^{-4}$. We have

$$\left| \partial_1^k \left( h_1(a^{-1}\xi_1) \right| = a^{-k} |\partial_1^k h_1(a^{-1}\xi_1)| \leqslant 2^{4k} \left| \partial_1^k h_1(a^{-1}\xi_1) \right| \lesssim 1,$$

where the last inequality follows from the properties of $\operatorname{supp} h_1$. Putting together the bounds for $h_1$ and $h_2$ yields the required estimate.

Now we turn our attention to the case $i = 2$. We have

$$\chi^2 \left( (D^2_{a^{-1}} S_s)^\tau \xi \right) = h_1(a^{-1}\xi_2) \left( 1 - h_2 \left( \frac{a^{-1}\xi_2}{a^{-1}s\xi_2 + a^{-1/2}\xi_1} \right) \right).$$

The question regarding $h_1(a^{-1}\xi_2)$ and its derivatives is readily addressed using the same arguments as we did in the case $i = 1$. Thus, due to the properties of $h_2$ and its support, we only need to consider the case

$$\frac{5}{4} \geqslant \left| a^{1/2}\frac{\xi_1}{\xi_2} + s \right| \geqslant \frac{4}{5} \tag{29}$$

Define functions $g(\boldsymbol{\xi}) = \xi_2$ and $h(\boldsymbol{\xi}) = (a^{1/2}\xi_1 + s\xi_2)^{-1}$ so that

$$(gh)(\boldsymbol{\xi}) = \frac{a^{-1}\xi_2}{a^{-1}s\xi_2 + a^{-1/2}\xi_1} = \frac{1}{a^{1/2}\frac{\xi_1}{\xi_2} + s}.$$

Consider a multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$. We have

$$\partial^{\boldsymbol{\alpha}}(gh)(\boldsymbol{\xi}) = \sum_{\boldsymbol{\beta} \leqslant \boldsymbol{\alpha}} C_{\boldsymbol{\beta}} \partial^{\boldsymbol{\alpha}-\boldsymbol{\beta}} g(\boldsymbol{\xi}) \partial^{\boldsymbol{\beta}} h(\boldsymbol{\xi}).$$

Bounding the derivatives of $g$ is trivial. Regarding the derivatives of $h$, we have

$$\partial^{\boldsymbol{\alpha}} h(\boldsymbol{\xi}) = (-1)^{|\alpha|} a^{\frac{\alpha_1}{2}} s^{\alpha_2} (h(\boldsymbol{\xi}))^{2^{|\alpha|}}.$$

Therefore, $\partial^{\boldsymbol{\alpha}} h$ is bounded from above as long as $h$ is, i.e., as long as $1/h(\boldsymbol{\xi}) = a^{1/2}\xi_1 + s\xi_2$ is bounded from below, but this is ensured by (29). In conclusion, $\chi^2$ and its derivatives are bounded on $\operatorname{supp}\hat{\psi}^2$.

Let us now find the bounds for $\chi^0$. Without loss of generality take $i = 1$. To show

$$\left\| \chi^0\left( (D^1_{1/a_\lambda} S^1_s)^\tau \boldsymbol{\xi} \right) \right\|_{C^N(\operatorname{supp}\hat{\psi}^1)} \leqslant \gamma_N,$$

it is sufficient to show that derivatives of $\chi^2\left((D^1_{1/a}S^1_s)^\tau\boldsymbol{\xi}\right)$ are uniformly bounded on $\operatorname{supp}\hat{\psi}^1$. We have

$$\chi^2\left((D^1_{1/a}S^1_s)^\tau\boldsymbol{\xi}\right) = h_1(a^{-1}s\xi_1 + a^{-1/2}\xi_2)\left(1 - h_2\left(\frac{a^{-1}s\xi_1 + a^{-1/2}\xi_2}{a^{-1}\xi_1}\right)\right).$$

Restrictions imposed by (28) suggest we only ought to consider

$$\frac{4}{5} \leqslant \left| \frac{a^{-1}s\xi_1 + a^{-1/2}\xi_2}{a^{-1}\xi_1} \right| \leqslant \frac{5}{4},$$

which gives

$$\frac{1}{40} \leqslant \left| s\xi_1 + a^{1/2}\xi_2 \right| \leqslant \frac{5}{16}. \tag{30}$$

On the other hand, looking at the support of $h_1$ suggests that we only need to consider $\boldsymbol{\xi}$ that satisfy

$$\frac{1}{4} \leqslant a^{-1}|s\xi_1 + a^{1/2}\xi_2|.$$

Combined with (30), we have that there is an $a_0$ such that if $a \leqslant a_0$ then

$$a^{-1}|s\xi_1 + a^{1/2}\xi_2| < \frac{1}{4},$$

that is, not in the support of $h_1$. Thus, using (29) for $a > a_0$ we have

$$\left|\partial^\alpha h_1(a^{-1}s\xi_1 + a^{-1/2}\xi_2)\right| \lesssim a^{-\alpha_1 - \frac{\alpha_2}{2}} s^{\alpha_2} \left|\partial^\alpha h_1(a^{-1}s\xi_1 + a^{-1/2}\xi_2)\right| \lesssim 1.$$

Bounding $h_2\left(s + a^{1/2}\frac{\xi_2}{\xi_1}\right)$ is analogous to the previously addressed case $i = 1$. □

Let us now write $\chi^0 = \chi^{01} + \chi^{02}$, and define $\tilde{\chi}^1 = \chi^1 + \chi^{01}$ and $\tilde{\chi}^2 = \chi^2 + \chi^{02}$. Using Lemma (B.1) we can show that the analogous statement holds for $\tilde{\chi}^1$ and $\tilde{\chi}^2$. We can now define our shearlet system, associated with high frequencies, in the Fourier domain

$$\hat{\sigma}^i_{as\mathbf{b}}(\xi) = \gamma^i(\xi)\hat{\psi}^i_{as\mathbf{b}}(\xi)$$

and its corresponding dual system

$$\hat{\tilde{\sigma}}^i_{as\mathbf{b}}(\xi) = \frac{\tilde{\chi}^i(\xi)}{\gamma^i(\xi)}\hat{\psi}^i_{as\mathbf{b}}(\xi),$$

for $i = 1, 2$.

**Theorem B.2.** *The families*

$$\Psi = \left\{\sigma^i_{as\boldsymbol{b}}, i \in \{1,2\}, a \in [0,1], s \in (-3/2, 3/2), \boldsymbol{b} \in \mathbb{R}^2\right\}$$

*and*

$$\tilde{\Psi} = \left\{\tilde{\sigma}^i_{as\boldsymbol{b}}, i \in \{1,2\}, a \in [0,1], s \in (-3/2, 3/2), \boldsymbol{b} \in \mathbb{R}^2\right\}$$

*constitute two systems of parabolic molecules.*

*Proof.* This follows directly from Lemma B.1 and the support properties of $\psi^i$. □

Let us now establish a reconstruction formula. We define

$$C_\psi = \int_{\mathbb{R}^2} \frac{\left|\hat{\psi}(\xi)\right|^2}{|\xi_1|^2} d\xi,$$

and

$$\Delta_\psi(\xi) = \int_{-3/2}^{3/2}\int_0^1 \left|\hat{\psi}(a\xi_1, a^{\frac{1}{2}}(\xi_2 - s\xi_1))\right|^2 a^{-\frac{3}{2}} da ds.$$

Lastly, we define functions $\varphi_0$ and $\varphi_1$ through

$$|\hat{\varphi}_0(\xi)|^2 = C_\psi - \Delta_{\psi_1} \text{ and } |\hat{\varphi}_1(\xi)|^2 = C_\psi - \Delta_{\psi_2}.$$

We have to show that these $\varphi_0$ and $\varphi_1$ are smooth, which we do using the standard arguments.

**Lemma B.3.** *We have*

$$\hat{\varphi}_0(\xi) = \mathcal{O}(|\xi|^{-N})$$

*for some $N \in \mathbb{N}$ and $\left|\frac{\xi_2}{\xi_1}\right| \leqslant \frac{4}{3}$, and an analogous statement holds for $\varphi_1$ in the cone $\left|\frac{\xi_1}{\xi_2}\right| \leqslant \frac{4}{3}$.*

*Proof.* We first notice that $C_\psi$ can be written as

$$C_\psi = \int_{\mathbb{R}} \int_0^\infty \left| \hat{\psi}\left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right|^2 a^{-3/2} da ds$$

it follows

$$|\hat{\varphi}_0(\boldsymbol{\xi})| = \int_{|s|>3/2} \int_0^\infty \left| \hat{\psi}\left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right|^2 a^{-3/2} da ds$$
$$+ \int_{|s|<3/2} \int_{a>a_0} \left| \hat{\psi}\left( a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)\right|^2 a^{-3/2} da ds.$$

We split the first of these integrals in two parts, one over $[0, a_0]$ and the other for $a > a_0$. To treat those we use the vanishing moments and Fourier decay in $\xi_2$. The second integral is treated by using the Fourier decay in the first variable. □

Let us now consider

$$f_{\mathsf{high}}(\mathbf{x}) = \int \langle f, \check{\gamma}^1 * \psi^1_{as\mathbf{b}} \rangle \left( \frac{\tilde{\chi}^1}{\gamma^1} \right)^{\vee} * \psi^1_{as\mathbf{b}}(\mathbf{x}) \frac{da ds d\mathbf{b}}{a^3} +$$
$$+ \int \langle f, \check{\gamma}^2 * \psi^2_{as\mathbf{b}} \rangle \left( \frac{\tilde{\chi}^2}{\gamma^2} \right)^{\vee} * \psi^2_{as\mathbf{b}}(\mathbf{x}) \frac{da ds d\mathbf{b}}{a^3}.$$

For the low-frequency case we take

$$f_{\mathsf{low}}(\mathbf{x}) = \int \langle f, \check{\gamma}^0 * T_{\mathbf{b}}\varphi_0 \rangle \left( \frac{\tilde{\chi}^1}{\gamma^0} \right)^{\vee} * T_{\mathbf{b}}\varphi_0(\mathbf{x}) d\mathbf{b} + \int \langle f, \check{\gamma}^0 * T_{\mathbf{b}}\varphi_1 \rangle \left( \frac{\tilde{\chi}^2}{\gamma^0} \right)^{\vee} * T_{\mathbf{b}}\varphi_1(\mathbf{x}) d\mathbf{b},$$

where $T_{\mathbf{b}}$ is the translation operator.

We now have all the required ingredients for the last missing piece, the reconstruction formula.

**Theorem B.4.** *We have*

$$f = \frac{1}{C_\psi}(f_{high} + f_{low}).$$

*Proof.* Taking the Fourier transform of $f_{\mathsf{high}}$ yields

$$\hat{f}_{\mathsf{high}}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}) \left( \tilde{\chi}^1(\boldsymbol{\xi}) \Delta_{\psi^1}(\boldsymbol{\xi}) + \tilde{\chi}^2(\boldsymbol{\xi}) \Delta_{\psi^2}(\boldsymbol{\xi}) \right),$$

whilst the Fourier transform of $f_{\mathsf{low}}$ yields

$$\hat{f}_{\mathsf{low}}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}) \left( \tilde{\chi}^1(\boldsymbol{\xi}) |\hat{\varphi}_0(\boldsymbol{\xi})|^2 + \tilde{\chi}^2(\boldsymbol{\xi}) |\hat{\varphi}_1(\boldsymbol{\xi})|^2 \right).$$

Therefore

$$\frac{1}{C_\psi}(\hat{f}_{\mathsf{high}}(\boldsymbol{\xi}) + \hat{f}_{\mathsf{low}}(\boldsymbol{\xi})) = \frac{\hat{f}(\boldsymbol{\xi})}{C_\psi} \left( \tilde{\chi}^1(\boldsymbol{\xi})(\Delta_{\psi^1}(\boldsymbol{\xi}) + |\hat{\varphi}_0(\boldsymbol{\xi})|^2) + \tilde{\chi}^2(\boldsymbol{\xi})(\Delta_{\psi^2}(\boldsymbol{\xi}) + |\hat{\varphi}_1(\boldsymbol{\xi})|^2) \right)$$
$$= \hat{f}(\boldsymbol{\xi})(\tilde{\chi}^1(\boldsymbol{\xi}) + \tilde{\chi}^2(\boldsymbol{\xi})) = \hat{f}(\boldsymbol{\xi})$$

which is precisely what we wanted to show.

□

# References


[1] E. J. Candès and D. L. Donoho. Continuous Curvelet Transform: I. Resolution of the Wavefront Set. *Appl. Comput. Harmon. Anal*, 19:162–197, 2003.

[2] E. J. Candès and L. Demanet. The Curvelet Representation of Wave Propagators is Optimally Sparse. *Comm. Pure Appl. Math*, 58:1472–1528, 2004.

[3] P. Grohs and G. Kutyniok. Parabolic Molecules. *Foundations of Computational Mathematics*, pages 299–337, 2013.

[4] P. Grohs, G. Kutyniok, S. Keiper, and M. Schaefer. $\alpha$-molecules. Technical report, ETH Zürich.

[5] H. F. Smith. A Parametrix Construction for Wave Equations with $C^{1,1}$ Coefficients. *Annales de l'institut Fourier*, pages 797–835, 1998.

[6] G. Kutyniok, D. Labate, and Editors. *Shearlets: Multiscale Analysis for Multivariate Data*. Birkhäuser, 2012.

[7] Y. Meyer and R. R. Coifman. *Wavelets, Calderón-Zygmund, and Multilinear Operators*. Cambridge University Press, 1997.

[8] P. Grohs. Continuous Shearlet Frames and Resolution of the Waffront Set. *Monatshefte für Mathematik*, 164:393–426, 2010.

[9] K. Guo and D. Labate. The Construction of Smooth Parseval Frames of Shearlets. *Math. Model. Nat. Phenom.*, 8(1):82–105, 2013.

[10] P. Grohs. Bandlimited Shearlet Frames with Nice Duals. *Journal of Computational and Applied Mathematics*, pages 139–151, 2013.

[11] K. Guo and D. Labate. Representation of Fourier Integral Operators Using Shearlets. *Journal of Fourier Analysis and Applications*, 14:327–371, 2008.

[12] P. Grohs. Continuous Shearlet Tight Frames. *Journal of Fourier Analysis and Applications*, pages 506–518, 2011.

# Recent Research Reports

| Nr. | Authors/Title |
|---|---|
| 2015-07 | J. Dick and F.Y. Kuo and Q.T. Le Gia and Ch. Schwab<br>Fast QMC matrix-vector multiplication |
| 2015-08 | P. Chen and Ch. Schwab<br>Adaptive Sparse Grid Model Order Reduction for Fast Bayesian Estimation and Inversion |
| 2015-09 | J.-L. Bouchot and B. Bykowski and H. Rauhut and Ch. Schwab<br>Compressed Sensing Petrov-Galerkin Approximations for Parametric PDEs |
| 2015-10 | A. Jentzen and P. Pusnik<br>Strong convergence rates for an explicit numerical |

approximation method for stochastic evolution equations with

non-globally Lipschitz continuous nonlinearities

| Nr. | Authors/Title |
|---|---|
| 2015-11 | R. Bourquin<br>Exhaustive search for higher-order Kronrod-Patterson Extensions |
| 2015-12 | R. Bourquin and V. Gradinaru<br>Numerical Steepest Descent for Overlap Integrals of Semiclassical Wavepackets |
| 2015-13 | A. Hiltebrand and S. Mishra<br>Entropy stability and well-balancedness of space-time DG for the shallow water equations with bottom topography |
| 2015-14 | B. Ayuso de Dios and R. Hiptmair and C. Pagliantini<br>Auxiliary Space Preconditioners for SIP-DG Discretizations of H(curl)-elliptic Problems with Discontinuous Coefficients |