FUNCTION RECONSTRUCTION BY THE WAVELET AND SHEARLET TRANSFORM

Jandre Snyman
Supervised by Dr Melissa Tacy

A dissertation submitted in partial fulfilment of the requirements of the degree of Bachelor of Science with Honours

Department of Mathematics and Statistics
University of Otago

October, 2019

Abstract

The ability to reformulate functions in a way that highlights their most important properties is a key tool in the study of partial differential equations and functional analysis. This dissertation explores the topic of function reconstruction by reviewing three common transforms and their associated inversion formulas. We cover, in detail, the Fourier transform, the wavelet transform and the shearlet transform. Once we've discussed these transforms we apply the shearlet transform to a problem in functional analysis. In particular, we study the L^2 to L^p behaviour of a certain family of convolution operators.

Acknowledgements

Throughout the writing of this dissertation I have received a great deal of support and assistance. First and foremost I would like to thank my supervisor, Dr. Melissa Tacy, whose expertise has been invaluable in the formulating of the research topic and methodology in particular. I am extremely grateful for her guidance and the knowledge she has passed down to me. I am proud to have been her student.

I would like to acknowledge my lecturers who have taught me everything I know. You have shown me the beauty in mathematics and helped to develop my passion for the subject.

Next I would like to thank my fellow honours students. You all ensured that there were plenty of laughs to be had in the honours room. Because of you the honours room felt like a second home.

Thank you to my parents and sister for supporting me throughout my years at university. Thank you to my parents, I will always be grateful to you for ensuring that I get a quality education that will prepare me for the future.

Lastly, but definitely not least I would like to thank my girlfriend Michaela Roy, for her continued love and support. When I was tired you gave me the energy to keep going.

Contents

Introduction

In the theory of partial differential equations and functional analysis we often derive abstract formulations of functions. These formulations often tell us very little about the behaviour of the function itself. Thus we need to derive properties of the function ourselves. We could use first principles, but we can also try to reformulate the function in such a way that it makes certain properties more obvious. A common way of doing this is by superposition. We pick a family G of simpler functions, we call these the analysing functions. Then under certain assumptions we can superimpose these analysing functions to recreate the original function f. This superposition can be discrete;

$$f(x) = \sum_{i \in I} a_i g_i(x) \quad a_i \in \mathbb{R}(\text{or } \mathbb{C}),\ g \in G, \tag{1}$$

where I is a countable index set. Alternatively the superposition can be continuous;

$$f(x) = \int_{\Omega} a(\omega) g(\omega, x) d\mu(\omega), \qquad \text{for all } x \in \mathbb{R}^n, \tag{2}$$

where g is a function on $\mathbb{R}^{n+m}$, Ω is some subset of $\mathbb{R}^m$, μ is some measure on Ω, and a is some function from $\mathbb{R}^m$ to $\mathbb{R}$(or $\mathbb{C}$). We call superpositions such as (1) and (2) *reconstruction Formulas.*

In this dissertation we focus on continuous reconstruction formulas for functions in $L^2(\mathbb{R}^n)$. However, for a given function f we obviously can't just take any superposition of the analysing functions and expect to get f back. We need to know how to weight each analysing function to exactly reconstruct f. That is, we need to know what the function a in (2) is. This is where the *transform* comes in. For our purposes we can define a transform as follows; Let g be an analysing function from $\mathbb{R}^{m+n}$ to $\mathbb{R}$(or $\mathbb{C}$) and W the set of all functions from $\mathbb{R}^m$ to $\mathbb{R}$(or $\mathbb{C}$). Then a transform T_g associated with g is an injective linear map,

$$T : L^2(\mathbb{R}^n) \to W, \quad g \mapsto T_g f$$

such that there exists a subspace $\mathcal{L}$ of $L^2(\mathbb{R}^n)$, a set $\Omega \subset \mathbb{R}^m$ and a measure μ on Ω satisfying

$$f(x) = \int_{\mathbb{R}^n} T_g f(\omega) g(\omega, x) d\mu(\omega) \qquad \text{for all } x \in \mathbb{R}^n.$$

Under reasonable assumptions we can show that $T_g f \in L^2(\mathbb{R}^m)$, in this case we have $W = L^2(\mathbb{R}^m)$. We then call the map,

$$T^{-1} : W \to L^2(\mathbb{R}^n), \quad w \mapsto \int_{\Omega} w(\omega) g(\omega, \cdot) d\mu(\omega)$$

the inverse transform.

We will be looking at three such transforms namely; The *Fourier* Transform, the *Wavelet* Transform and the *Shearlet* Transform. We will discuss these transforms in the stated order. Each transform will build on the one before. As mentioned, the first transform we will discuss is the Fourier transform. The Fourier transform, developed in the 1800's, is the oldest and most common of the three transforms. We denote the Fourier Transform by $\mathcal{F}$ and it is given by

$$\mathcal{F}[f](\xi) := \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x, \xi \rangle} dx.$$

The Fourier transform analyses the frequency content of a function f. Then under some assumptions, covered in Chapter 1, we can reconstruct f as a superposition of infinitely many waves;

$$f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(x) e^{i\langle x, \xi \rangle} dx.$$

The Fourier transform is extremely useful in functional analysis. From simple properties of $\hat{f}$ we can derive properties of f that would be quite difficult or tedious to derive from first principles. In particular the Fourier transform is very good at telling us about the *global* regularity of f and about the decay of f . However, the Fourier transform is only really good at analysing the global behaviour of f. The Fourier transform is the transform associated with the function $e^{i\langle \cdot, \cdot \rangle}$, which is not well localised in any of its variables. This means that the Fourier transform is not so sensitive to local properties of functions. For example the Fourier transform can detect global continuity of a functions, but it can't detect continuity at a single point.

Next we will look at the wavelet transform. The wavelet transform fixes some of the localisation issues that the Fourier transform has. The wavelet transform is a family of transforms associated with a family of analysing functions. We define the wavelet transform of a function $f \in L^2(\mathbb{R})$ as the 2-dimensional function

$$[T^{wav}f](a,a) := |a|^{-1/2} \int_{\mathbb{R}} f(x)\overline{\psi\left(\frac{x-b}{a}\right)}dx, \quad a > 0,\ b \in \mathbb{R}.$$

Here we are free to pick ψ, as long as it is chosen to be a localised wave-like function. How to pick ψ is discussed in Chapter 2. The wavelet transform dilates the wavelet and shifts it in space. This allows the wavelet transform to analyses regions of any size, anywhere in space. We can reconstruct any $f \in L^2(\mathbb{R})$ by the reconstruction formula;

$$f = C_\psi^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [T^{wav}f](a,b)\psi^{a,b}\frac{da\ db}{a^2}.$$

However, equality holds only in the sense that both sides have the same inner products with all other elements in $L^2(\mathbb{R})$. Equality does not hold in a pointwise sense. In 1-dimension the wavelet transform is good at detecting global and local regularity of functions. In particular, we can characterise global and local Hölder continuity of a function f by the asymptotic behaviour of $[T^{wav}f]$ in a. It is also possible to extend the wavelet transform to higher dimensions. However, in higher dimensions the wavelet transform lacks sensitivity to directionality. The lack of sensitivity to directionality means that the wavelet transform is not very effective at detecting local geometry in higher dimensions. This can be a problem since the local geometry of a higher dimensional function is often what we are most interested in.

Last we will look at the Shearlet Transform. Shearlets can be extended to arbitrary dimensions, but we will restrict to the 2-dimensional case. Like the wavelet transform fixes the Fourier transform's lack of localisation we will fix the wavelet's lack of sensitivity to directionality with the shearlet transform. We define the shearlet transform in the same way we define the wavelet transform, except we introduce directional variation. This is done by shearing and asymmetrically dilating the support of our analysing functions. The shearlet transform is then given by

$$[\mathcal{SH}_\psi f](a,s,t) = \int_{\mathbb{R}^2} f(x)\psi_{a,s,t}(x)dx. \quad a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2.$$

Reconstructing functions using the shearlet transform is possible and we will cover how to do so in detail in Chapter 3. However, it is a bit more complicated than using the Fourier transform or the wavelet transform. As with the wavelet, the shearlet reconstruction formulas hold only in an L^2 weak sense. While the shearlet transform has directional sensitivity, it also has directional bias. It can only analyse vertical frequency details of f after an infinite shear. To fix this we need to introduce the *Cone Adapted Shealet Transform*. This leads to a new reconstruction formula, that reconstructs the vertical and horizontal frequency parts of f separately.

Finally, once we have introduced all three transforms we will apply the shearlet transform to a problem in functional analysis. In particular we will look at convolution operators. We will pick specific multipliers $m \in L^\infty(\mathbb{R}^2)$ and the linear operator T_m on $L^2(\mathbb{R}^2)$ by

$$T_m f = \mathcal{F}^{-1}[m\hat{f}], \qquad f \in L^2(\mathbb{R}^2).$$

Showing that this is a bounded linear operator on $L^2(\mathbb{R}^2)$ is easy as it is the L^∞ bound that tells us about the L^2 to L^2 estimates. However we want to know if this operator defines a bounded linear operator from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ for any $p \in [1,\infty]$. When trying to prove that this is the case we will reconstruct m using a cone adapted shearlet reconstruction formula. We then consider a family of convolution operators generated by a family of smooth multipliers concentrated on the unit circle. We will show that all convolution operators in this family are bounded linear operators from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ for any $p \in [2,\infty]$. We also find estimates for their L^p norms.

Chapter 1

The Fourier Transform

While the main focus of this dissertation will be *Wavelets* and *Shearlets*, these transforms rely heavily on the Fourier transform. Almost all results about *Wavelets* and *Shearlets* are proved using the Fourier Transform. Many of the definitions and concepts relating to *Wavelets* and *Shearlets* are formulated using the Fourier transform. As such we will take the entire first chapter to review all aspects of the Fourier transform that will be necessary before progressing to Chapters 2 and 3. Because of this, this chapter will essentially be a summary of [15]. We will present many results without proving them since their proofs are not informative to our later work. However proofs and/or illustrative examples will be provided where deemed necessary.

1.1 The Fourier Transform of Integrable functions

We start off our discussion by defining the Fourier transform.

Definition 1.1.1 (Fourier Transform[15]**)**

Let $f \in L^1(\mathbb{R}^n)$ then we define the Fourier transform of f to be

$$\mathcal{F}[f](\xi) := \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,\xi\rangle} dx.$$

<u>*note:*</u> In order to ensure that the Fourier transform is well defined we have to assume that f is L^1 integrable. This ensure that the integral converges absolutely.

For a given function $f \in L^1(\mathbb{R}^n)$ deriving properties of its fourier transform $\hat{f}$ can be difficult, and the properties of $\hat{f}$ are closely related to the additional properties of f. However, there are some properties of $\hat{f}$ that only require $f \in L^1(\mathbb{R}^n)$.

Theorem 1.1.2 (adapted from Chapter 7 page 167 in[16]**)**

If $f \in L^1(\mathbb{R}^n)$ then we have;

1. *$\hat{f}$ is bounded with $|\hat{f}(\xi)| \leq ||f||_1$.*
2. *$\hat{f}$ is continuous on $\mathbb{R}^n$.*
3. $\lim_{\xi \to \pm\infty} \hat{f}(\xi) = 0$.

There are a few more general properties of the Fourier transform that we will provide below. These properties are especially useful when we manipulate expressions where we have to switch back and forth between f and $\hat{f}$. (What we mean by switching back and forth between f and $\hat{f}$ is made more clear later when we discuss the inverse transform).

Theorem 1.1.3 (adapted from [15] **and** [16] **chapter 19)**

Let $f \in L^1(\mathbb{R}^n)$ then the following statements hold.

1. *The map $\mathcal{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ is a bounded linear operator. Here $C_0(\mathbb{R}^n)$ is the set of all continuous functions that go to zero at infinity.*
2. *For any point $x_0 \in \mathbb{R}^n$ we have that*
$$\mathcal{F}[\tau_{x_0} f](\xi) = e^{-i\langle x_0, \xi\rangle} \hat{f}(\xi).$$
By τ_{x_0} we mean the translation operator $\tau_{x_0} f(x) = f(x - x_0)$.
3. *For any fixed point $\xi_0 \in \mathbb{R}^n$ we have that*
$$\mathcal{F}[e^{i\langle x, \xi_0\rangle} f](\xi) = \tau_{\xi_0} \hat{f}(\xi).$$
4. *For any constant real number $c > 0$ we have*
$$\mathcal{F}[f(cx)](\xi) = c^{-n} \hat{f}\left(\frac{\xi}{c}\right).$$

In much of our later work we will need to have a little more structure on the function space in which we are working. The function space $L^2(\mathbb{R}^n)$ can be viewed as a Hilbert space and it will turn out to have the necessary structure that we require. It will be useful to relate the norm of a function f with the that of its Fourier transform $\hat{f}$. The main result we will be using for this is known as Plancherel's Theorem.

Theorem 1.1.4 (Plancherel's Theorem adapted from Chapter 19 Page 226 of [14]**)**
Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ then $\hat{f} \in L^2(\mathbb{R}^n)$ and in particular

$$\int_{\mathbb{R}^n} |f(x)|^2 dx = \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 d\xi.$$

That is, the Fourier transform is an isometry from $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$ with respect to the L^2 norm.

In some cases the Fourier transform $\hat{f}$ of a function f is much easier to find and work with than f itself. In this case we would like to invert the Fourier transform to define f in terms of $\hat{f}$. Defining the inverse Fourier transform is not so straight forward. Since $f \in L^1(\mathbb{R}^n)$ does not imply that $\hat{f} \in L^1(\mathbb{R}^n)$, we are not guaranteed that the inverse Fourier transform (see Definition 1.1.10.2 below) is well defined on $L^1(\mathbb{R}^n)$. To see this, recall from Theorem 1.1.2 that for any $f \in L^1(\mathbb{R}^n)$ the function $\hat{f}$ is continuous. Therefore any integrable function f that is not continuous does not have an inverse Fourier transform in $L^1(\mathbb{R}^n)$. For example consider the square wave function given by a geometric series

$$g(\xi) = \sum_{k=-\infty}^{\infty} \left(\frac{-1}{2}\right)^{|k|} \chi_{I(k)}(\xi),$$

where $I(k) = \left[k - \frac{1}{2}, k + \frac{1}{2}\right)$, and $\chi_{I(k)}$ is the characteristic function of this interval. The graph of this function is given in figure 1.1 below. We then also have

$$\int_{\mathbb{R}} |g(\xi)| d\xi = 1 + 2\sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k = 3.$$

Therefore g is in $L^1(\mathbb{R})$ and it even satisfies $g(\xi) \to 0$ as $|\xi| \to \infty$, but it is not continuous. Thus there exists no $f \in L^1(\mathbb{R})$ such that $\hat{f}(\xi) = g(\xi)$.

In order to ensure that the inverse Fourier transform is invertible we have to consider a more specialised function space.

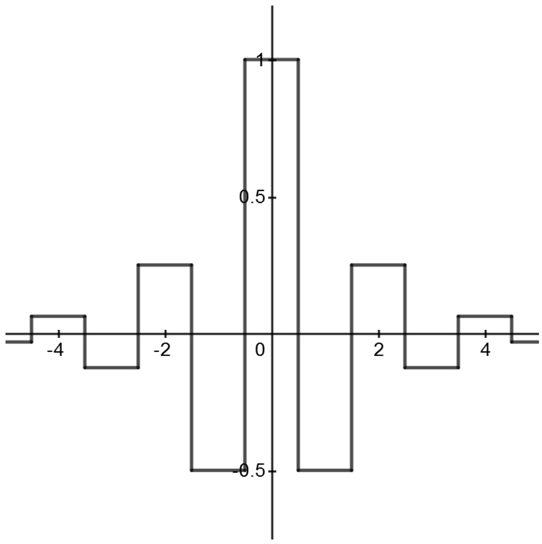


Figure 1.1: Geometric series square wave

Definition 1.1.5 (adapted from Chapter 19 Page 227 [14]**)**

1. *Let $a, b \in \mathbb{N}_0^n$ be multi-indices. Let $f \in C^\infty(\mathbb{R}^n)$ then we define*

$$||f||_{a,b} = \sup_{x \in \mathbb{R}^n} \left| x^a \partial^b f(x) \right|.$$

2. *We define the Schwartz Space, $\mathcal{S}(\mathbb{R}^n)$ as the space of rapidly decaying functions on $\mathbb{R}^n$. That is*

$$\mathcal{S}(\mathbb{R}^n) := \{ f \in C^\infty(\mathbb{R}^n) : ||f||_{a,b} < \infty \quad \forall a, b \in \mathbb{N}^n \}.$$

Remark 1.1.6

1. *If $a = (a_1, ..., a_n),\ b = (b_1, ..., b_n)$ and $x = (x_1, ..., x_k)$ then the multi-index notation used above is defined as.*

$$x^a := \prod_{j=1}^n x_j^{a_j} \quad \textit{and} \quad \partial^b := \frac{\partial^{b_1 + ... + b_n}}{\partial_{x_1}^{b_1} ... \partial_{x_n}^{b_n}}.$$

We define another differential operator D, which we will use later, by

$$D^a := \prod_{j=1}^n \left(\frac{1}{i} \frac{\partial}{\partial x_j} \right)^{a_j}.$$

2. *We denote the set of all smooth, compactly supported functions on $\mathbb{R}^n$ by $C_0^\infty(\mathbb{R}^n)$, we call this the space of Bump Functions. It can be shown that $C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$.* [15]

Since the Schwartz functions are all bounded and rapidly decaying functions, we would suspect that they should be included in all L^p spaces. This suspicion is in fact correct.

Propostion 1.1.7 (Adapted from [15]**)**

Suppose that $f \in \mathcal{S}(\mathbb{R}^n)$, the for any $p \in [1, \infty]$ we have that $f \in L^p(\mathbb{R}^n)$.

proof: Pick an arbitrary Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$. First we note that $f \in \mathcal{S}(\mathbb{R}^n)$ implies that

$$||f||_{a,b} = \sup_{x \in \mathbb{R}^n} \left| x^a \partial^b f(x) \right| < \infty$$

for all $a, b \in \mathbb{N}^n$. In particular, if $a = b = 0$ then we have that

$$||f||_\infty = \sup_{x \in \mathbb{R}^n} \left| x^0 \partial^0 f(x) \right| = ||f||_{0,0} < \infty$$

which means that $f \in L^\infty(\mathbb{R}^n)$. Next, Suppose that $p \in [1, \infty)$. Then we notice that

$$\int_{\mathbb{R}^n} |f(x)|^p = \int_{\mathbb{R}^n} ((1 + ||x||^2)|f(x)|)^p \frac{1}{(1 + ||x||^2)^p} dx.$$

The fact that $f \in \mathcal{S}$ means that

$$\left|(1 + ||x||^2) f(x)\right| = \left| f(x) + \sum_{j=1}^n x_j^2 f(x) \right| = \left| x^0 \partial^0 f(x) + \sum_{j=1}^n x^{2e_j} \partial^0 f(x) \right| < \infty,$$

where $\{e_j\}_{j=1}^n$ is the standard basis for $\mathbb{R}^n$. For $p \geq 1$ we also have that

$$\frac{1}{(1 + x^2)^p} \leq \frac{1}{1 + x^2} \in L^1(\mathbb{R}^n).$$

Now we finally compute that

$$\begin{aligned}
\int_{\mathbb{R}^n} |f(x)|^p dx &\leq ||(1 + ||x||^2) f(x)||_\infty^p \int_{\mathbb{R}^n} \frac{1}{1 + ||x||^2} dx \\
&= \pi^n ||(1 + ||x||^2) f(x)||_\infty^p \\
&\leq \pi^n (||f||_{0,0} + \sum_{j=1}^n ||f||_{2e_j, 0}) < \infty.
\end{aligned}$$

Therefore $f \in L^p(\mathbb{R}^n)$. $\square$

The Fourier Transform and Schwartz functions interact well and it turns out that they are exactly what we need to guarantee the invertibility of the Fourier transform. Before we define the inverse Fourier transform we mention two more useful general properties of the Fourier transform that hold on the Schwartz functions.

Propostion 1.1.8 (Adapted from [15]**)**

For any multi-index $b \in \mathbb{N}^n$ and $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\mathcal{F}[(-x)^b f](\xi) = D^b \hat{f}(\xi) \quad \textit{and} \quad \mathcal{F}[D^b f](\xi) = \xi^b \hat{f}(\xi).$$

The second result is an extension of Plancherel's Theorem that holds on the Schwartz functions.

Propostion 1.1.9 (Parseval's Theorem[Adapted from [15]**)**

Let $f, g \in \mathcal{S}$ *then we have that*

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx = \int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}dx.$$

Now we define the inverse Fourier transform by the following theorem.

Theorem 1.1.10 (Adapted from [15]**)**

1. *The Fourier transform maps Schwartz functions to Schwartz functions. That is, if* $f \in \mathcal{S}(\mathbb{R}^n)$ *then* $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$.

2. *If* $f \in \mathcal{S}(\mathbb{R}^n)$ *then we have that*

$$f(x) = \mathcal{F}^{-1}[\hat{f}](x) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\langle x,\xi\rangle} dx.$$

   For $g \in \mathcal{S}(\mathbb{R}^n)$ *we call* $\mathcal{F}^{-1}[g]$ *the inverse Fourier transform of* g *and denote it by* $\check{g}$.

In our earlier discussion about switching between a function and its Fourier transform in manipulations we were referring to the application of Plancherel's Theorem, Parseval's formula and $\mathcal{F}^{-1}$. As shown in the proof of Proposition 1.1.11, below.

Propostion 1.1.11

Let

$$f(\xi) = \frac{1 - e^{-2\pi i\xi}}{2\pi i\xi}.$$

Then f *is continuous, bounded and* $f \in L^2(\mathbb{R}^2)$. *In particular*

$$||f||_{L^2} = \sqrt{2\pi}$$

proof: First let $\chi_{[0,1]}$ be the indicator function of $[0, 1]$, that is

$$\chi_{[0,1]}(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\chi_{[0,1]} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, so we can take its Fourier transform. Then we compute

$$\begin{aligned}\mathcal{F}^{-1}[\chi_{[0,1]}](\xi) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \chi_{[0,1]} e^{-2\pi i x\xi} dx \\ &= \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-2\pi i x\xi} dx \\ &= \frac{1}{\sqrt{2\pi}}\left[\frac{e^{-2\pi x\xi}}{-2\pi i\xi}\right]_0^1 \\ &= \frac{1}{\sqrt{2\pi}}\frac{1-e^{-2\pi i\xi}}{2\pi i\xi} \\ &= \frac{1}{\sqrt{2\pi}} f(\xi).\end{aligned}$$

Thus we can say that

$$f = \sqrt{2\pi}\,\mathcal{F}^{-1}[\chi_{[0,1]}].$$

Then by Theorem 1.1.2 we know that f is bounded and continuous. Now since $\chi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, Plancherel's Theorem (Theorem 1.1.4) implies $f \in L^2(\mathbb{R})$ with

$$||f||_{L^2} = \sqrt{2\pi}||f^{-1}[\chi_{[0,1]}]||_{L^2} = \sqrt{2\pi}.$$

$\square$

The ability to switch back and forth between f and $\hat{f}$ is of interest to us since we will later be interested in functions defined by their Fourier transform. We will still want to derive properties of the original function. So we need to be able to switch back and forth. One last result we will be needing a lot in Chapter 4 relates to convolutions of functions.

Definition 1.1.12

Let $f, g \in L^1(\mathbb{R}^n)$ then we we call the function

$$f \star g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$$

the convolution of f with g.

The Fourier transform enables us to change convolutions of functions into the product of functions by the following result.

Propostion 1.1.13 (Adapted from [15]**)**

Suppose that $f, g \in \mathcal{S}(\mathbb{R}^n)$ then we have that $f \star g \in \mathcal{S}(\mathbb{R}^n)$ and

$$\mathcal{F}[f \star g](\xi) = (2\pi)^{n/2}\hat{f}(\xi)\hat{g}(\xi).$$

1.2 The Fourier Transform of Distributions

Not all functions that we will want to consider will be as well behaved as the usual trigonometric, polynomial and exponential functions. Sometimes we will consider functions such as the Heaviside step function; given by

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

The Heaviside step function creates complications when we try to differentiate it, since it's derivative is not well defined at $x = 0$. But we can integrate against $\frac{d}{dx}H(x)$ and still get a sensible answer. This is shown in Proposition 1.2.1 below.

Propostion 1.2.1

Let $f \in L^1(\mathbb{R})$ be a differentiable on all of $\mathbb{R}$, then we have that

$$\int_{\mathbb{R}} f(x)\frac{d}{dx}H(x)dx = f(0).$$

proof: The result follows by integration by parts as follows;

$$\int_{\mathbb{R}} f(x)\frac{d}{dx}H(x)dx = f(x)H(x)\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} \left(\frac{d}{dx}f(x)\right) H(x)dx,$$

since $f \in L^1(\mathbb{R})$ we know that $f(x) \to 0$ as $|x| \to \infty$. Therefore

$$\int_{\mathbb{R}} f(x)\frac{d}{dx}H(x)dx = -\int_0^{\infty} \frac{d}{dx}f(x)dx = -f(x)\Big|_0^{\infty} = f(0).$$

□

Proposition 1.2.1 shows that while the derivative of the Heaviside step function doesn't make sense as a conventional pointwise function, it does make sense to integrate certain functions against it. We can then generalise $\frac{d}{dx}H(x)$ to a linear functional.

Definition 1.2.2

Let $C := \{f \in L^1(\mathbb{R}) : |f(0)| < \infty\}$, then C is a vector space of functions with the usual pointwise addition and scalar multiplication. Define the Linear functional

$$T_\delta : C \to \mathbb{R}, \quad f \mapsto f(0).$$

If $f \in C$ is differentiable on all of $\mathbb{R}$ then

$$T_\delta f = \int_{\mathbb{R}} f(x)\frac{d}{dx}H(x)dx.$$

The linear functional T_δ is an example of what is known as a *Distribution*, and we call T_δ the *delta distribution*. Distributions can be seen as generalised functions. Distributions appear frequently in the theory of partial differential equations, one of the main fields where the Fourier transform is applied. To deal with this we extend the Fourier transform to a special type of distribution, known as a *Tempered Distribution*. However before we can define this version of distributions we will need a topological result related to $\mathcal{S}(\mathbb{R}^n)$.

Propostion 1.2.3 (adapted from [15]**)**

We define the function $d : \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n) \to [0, \infty)$ *by*

$$d(f,g) := \sum_{a,b \in \mathbb{N}^n} 2^{-|a|-|b|} \frac{||f-g||_{a,b}}{1 + ||f-g||_{a,b}}.$$

Then d *is well defined on* $\mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ *and the pair* $(\mathcal{S}(\mathbb{R}^n), d)$ *is a complete metric space.*

We can use the metric d to define a space known as the Tempered distributions.

Definition 1.2.4 (Adapted from [15]**)**

Let $\mathcal{S}'(\mathbb{R}^n)$ *be the collection of all continuous linear functionals on* $\mathcal{S}(\mathbb{R}^n)$*, that is*

$$\mathcal{S}'(\mathbb{R}^n) = \{T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C} : T \text{ linear and } T(u_j) \to 0 \text{ whenever } d(u_j, 0) \to 0 \text{ in } \mathcal{S}(\mathbb{R}^n)\}.$$

We call the space $\mathcal{S}'(\mathbb{R}^n)$ *the space of Tempered distributions.*

Corollary 1.2.4.1 below shows that all L^p functions can be associated with a unique tempered distribution. So tempered distributions can be viewed as a generalisation of the familiar notion of a function.

Corollary 1.2.4.1 (Adapted from [15] **and chapter 6 page 161 of** [9]**)**

1. *Suppose that* $T, S \in \mathcal{S}'(\mathbb{R}^n)$ *and* $\alpha \in \mathbb{C}$ *then we define*

$$(T+S)(u) := T(u) + S(u) \quad \text{and} \quad (\alpha T)(u) := \alpha T(u) \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^n)$$

   then $(T + \alpha S) \in \mathcal{S}'(\mathbb{R}^n)$*. This means that we can view* $\mathcal{S}'(\mathbb{R}^n)$ *as a vector space.*

2. *For all* $p \in [1, \infty]$ *every function* $f \in L^p(\mathbb{R}^n)$ *generates a unique tempered distribution by*

$$T_f(u) := \langle f, u \rangle_{L^2} := \int_{\mathbb{R}^n} f(x)u(x)dx, \quad u \in \mathcal{S}(\mathbb{R}^n).$$

3. For any $p \in [1, \infty]$ the collection

$$\mathcal{S}'_p := \{T_f : f \in \mathcal{S}(\mathbb{R}^n)\} \subset \mathcal{S}'(\mathbb{R}^n)$$

is a subspace of $\mathcal{S}'(\mathbb{R}^n)$.

Part *1* and *2* of the corollary above are both proven in [15], however part *3* is not. That part *3* holds can be easily seen from the point wise addition and scalar multiplication defined in *1* and the fact that all L^p spaces are vector spaces.

For an example of a tempered distribution defined by an L^p function consider

$$f(x) = \chi_{[0,1]}(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f \in L^p$ for any $p \in [1, \infty]$. Now define the linear the map

$$T_f : \mathcal{S}(\mathbb{R}) \to \mathbb{R}, \quad u \mapsto \int_{\mathbb{R}} f(x)u(x)dx = \int_0^1 u(x)dx.$$

Now we claim that T_f is a tempered distribution. Since T_f is a well defined linear map from $\mathcal{S}(\mathbb{R})$ to $\mathbb{C}$ all we need to do is show that $T_f(u_j) \to 0$ if $d(u_j, 0) \to 0$. So consider

$$T_f(u_j) = \int_0^1 u_j(x)dx,$$

then the triangle inequality for integrals implies that

$$|T_f(u_j)| \leq \int_0^1 |u_j(x)|dx \leq ||u_j||_\infty = ||u_j||_{0,0}.$$

Then the fact that $d(u_j, 0) \to 0$ as $j \to \infty$ implies that $||u_j||_{0,0} \to 0$ as $j \to \infty$. Thus we can conclude that $T_f(u_j) \to 0$ if $d(u_j, 0) \to 0$, therefore T_f is a tempered distribution.

The family of subspaces $\mathcal{S}'_p$ is certainly a large collection of examples of Tempered distributions, but the actual space $\mathcal{S}'(\mathbb{R}^n)$ is much larger still. Repeating the argument above, replacing $\int_0^1 u_j(x)dx$ with $u_j(0)$ one can show that the delta distribution, T_δ, is a tempered distribution. This can be viewed as the distribution generated by the derivative of the Heavyside step function, which is certainly not an L^p function.

Much like Schwartz functions played well with the Fourier transform the tempered distributions also work well with the Fourier transform. For this to happen we must define the Fourier transform on $\mathcal{S}'(\mathbb{R}^n)$ in a sensible way. Before we can do this we need to consider some operations on the tempered distributions. We extend operations from functions to tempered distributions by applying the operation to an L^p function f. We then try to find a general form for the Tempered distribution generated by the image of f under the operation. Using this general form we define the operation on the tempered distributions to match. This method is illustrated using Proposition 1.2.9 below.

The first operations that we wish to carry over from L^p spaces to the tempered distributions are conjugation and multiplication. The function spaces L^p are closed under conjugation and multiplication by smooth functions (These operations are defined in a point-wise fashion). These operations are not extendable to the whole of $\mathcal{S}'(\mathbb{R}^n)$ but can be extended to a certain subspace.

Definition 1.2.5 (adapted from [15]**)**

We define $C^\infty_{pb}(\mathbb{R}^n)$ to be the collection of smooth functions that grow at most polynomially. That is $f : \mathbb{R}^n \to \mathbb{C}$ is in $C^\infty_{pb}(\mathbb{R}^n)$ if and only if f is smooth and for every multi-index $b \in \mathbb{Z}^n$ there exists a constant $C_b > 0$ and an $N \in \mathbb{N}$ so that

$$|D^b f(x)| \leq C_b(1 + |x|^n)^N.$$

Remark 1.2.6

From the definition of the Schwartz space we can see that all Schwartz functions belong to $C^\infty_{pb}(\mathbb{R}^n)$. However $C^\infty_{pb}(\mathbb{R}^n)$ does not contain all smooth functions. Functions such as a^x where $a > 0$ are smooth but grow much faster than any polynomial.

Now we can define multiplication and conjugation on the tempered distributions by the following proposition.

Propostion 1.2.7 (Adapted from [15]**)**

Fix $T \in \mathcal{S}'(\mathbb{R}^n)$ and $f \in C^\infty_{pb}(\mathbb{R}^n)$ then we define the conjugate of T to be the distribution

$$\overline{T}(u) = \overline{T(u)} \quad \textit{for all } u \in \mathcal{S}'(\mathbb{R}^n)$$

and the multiplication of f and T by

$$fT(u) = T(fu) \quad \text{for all } u \in \mathcal{S}'(\mathbb{R}^n).$$

In this case we have that $\overline{T}, fT \in \mathcal{S}'(\mathbb{R}^n)$.

Differentiation is one of our main tools for analysing Schwartz functions. We would like to define a concept similar to differentiation on the tempered distributions. In order to define derivatives on the tempered distributions we consider first the distribution associated with the derivative of a Schwartz function. So we consider a function $f \in \mathcal{S}(\mathbb{R}^n)$ and any multi-index $a \in \mathbb{N}_0^n$. Then recall that

$$D^a = \prod_{j=1}^{n} \left(\frac{1}{i} \frac{\partial}{\partial x_j} \right)^{a_j}.$$

Now we consider the distribution $T_{D^a f}$ generated by $D^a f$. Pick a test function $u \in \mathcal{S}(\mathbb{R}^n)$ then we have

$$T_{D^a f} u = \int_{\mathbb{R}^n} D^a f(x) u(x) dx.$$

We integrate by parts for each $\frac{1}{i}\frac{\partial}{\partial x_j}$ and then the fact that $u \in \mathcal{S}(\mathbb{R}^n)$ implies that the boundary terms go to zero at infinity. Then we have that

$$T_{D^a f} u = (-1)^{|a|} \int_{\mathbb{R}^n} f(x) D^a u(x) dx = (-1)^{|a|} T_f(D^a u),$$

where

$$|a| = \prod_{j=1}^{n} |a_j|.$$

Thus we conclude that

$$T_{D^a f}(u) = (-1)^{|a|} T_f(D^a u) \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^n).$$

In the same way, we can also show that

$$T_{\partial^a f}(u) = (-1)^{|a|} T_f(\partial^a u) \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^n).$$

These results then generalise exactly to all Tempered distributions.

Propostion 1.2.8 (Adapted [15]**)**

For any tempered distribution T we define the distributional derivatives.

$$D^a T(u) = (-1)^{|a|} T(D^a u) \quad \text{and} \quad \partial^a T(u) = (-1)^{|a|} T(\partial^a u) \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^n).$$

Then we have that $D^a T, \partial^a T \in \mathcal{S}'(\mathbb{R}^n)$.

Finally, since the Fourier transform on the Schwartz functions related differentiation and convolutions we also extend the idea of convolutions to the tempered distributions. As we did with derivatives we first check the distribution associated with the convolutions of Schwartz functions. To understand the intuition behind this definition of distributional convolution we will derive it on the Tempered distributions generated by Schwartz functions. Pick any $f, g \in \mathcal{S}(\mathbb{R}^n)$, then we have that

$$(f \star g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy.$$

Due to Cauchy-Scwartz we know that $(f \star g) \in L^1(\mathbb{R})$ and therefore it can generate a tempered distribution. Let $T_{f\star g}$ be the tempered distribution generated by $(f \star g)$. To determine the action of $T_{f\star g}$ pick a test function $u \in \mathcal{S}(\mathbb{R}^n)$, then we compute

$$\begin{aligned}
T_{f\star g}u &= \int_{\mathbb{R}^n} (f \star g)(x)u(x)dx \\
&= \int_{\mathbb{R}^n} u(x) \left(\int_{\mathbb{R}^n} f(x-y)g(y)dy \right) dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} u(x)f(x-y)g(y)dxdy && \text{[Interchange integrals by Fubini]} \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} u(\tilde{x}+y)f(\tilde{x})g(y)d\tilde{x}dy && \text{[Substitute } \tilde{x} = x-y\text{]} \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} u(\tilde{x}-\tilde{y})f(\tilde{x})g(-\tilde{y})d\tilde{x}d\tilde{y} && \text{[Substitute } \tilde{y} = -y\text{]} \\
&= \int_{\mathbb{R}^n} f(\tilde{x}) \left(\int_{\mathbb{R}^n} u(\tilde{x}-\tilde{y})g(-\tilde{y})d\tilde{y} \right) d\tilde{x} \\
&= \int_{\mathbb{R}^n} f(\tilde{x})(u \star g^*)(\tilde{x})d\tilde{x} \\
&= T_f(u \star g^*),
\end{aligned}$$

where $g^*(y) = g(-y)$. Then we conclude that

$$T_{f\star g}(u) = T_f(g^* \star u) \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^n).$$

We can then extend this result to all tempered distributions by the following proposition.

Propostion 1.2.9 ([15])

We define the convolution of $T \in \mathcal{S}'(\mathbb{R}^n)$ with $g \in \mathcal{S}(\mathbb{R}^n)$ by

$$T \star g(u) = T(g^* \star u) \quad \textit{for all } u \in \mathcal{S}(\mathbb{R}^n)$$

where $g^*(x) = g(-x)$. Then we have that $T \star g \in \mathcal{S}'(\mathbb{R}^n)$.

Now we are ready to extend the Fourier transform from Schwartz functions to the tempered distributions. As we did before we first consider the distributions associated with the Fourier transform of Schwartz functions. Let $f \in \mathcal{S}(\mathbb{R}^n)$ then we consider the distributions generated by $\mathcal{F}[f]$ and $\mathcal{F}^{-1}[f]$. We start with $\mathcal{F}[f]$, so pick a test function $u \in \mathcal{S}(\mathbb{R}^n)$. Now we compute

$$\begin{aligned}
T_{\mathcal{F}[f]}(u) &= \int_{\mathbb{R}^n} u(\xi)\, \mathcal{F}[f](\xi) d\xi \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} u(\xi) \left(\int_{\mathbb{R}^n} f(x) e^{-i\langle x,\xi\rangle} dx \right) d\xi \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) \left(\int_{\mathbb{R}^n} u(\xi) e^{-i\langle x,\xi\rangle} d\xi \right) dx \quad \text{[Interchange integrals by Fubini]} \\
&= \int_{\mathbb{R}^n} f(x)\, \mathcal{F}[u](x) dx \\
&= T_f\left(\mathcal{F}[u]\right).
\end{aligned}$$

Since $u \in \mathcal{S}(\mathbb{R}^n)$ we know that $\mathcal{F}[u] \in \mathcal{S}(\mathbb{R}^n)$. Therefore the last line is well defined for any tempered distribution, even if it is not generated by an L^p function.

We can also find a tempered distribution associated with the inverse Fourier transform.

$$\begin{aligned}
T_{\mathcal{F}[f]}(u) &= \int_{\mathbb{R}^n} u(\xi)\, \mathcal{F}^{-1}[f](\xi) d\xi \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} u(x) \left(\int_{\mathbb{R}^n} f(\xi) e^{\langle x,\xi\rangle} d\xi \right) dx \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(\xi) \left(\int_{\mathbb{R}^n} u(x) e^{i\langle x,\xi\rangle} dx \right) d\xi \quad \text{[Interchange integrals by Fubini]} \\
&= \int_{\mathbb{R}^n} f(\xi)\, \mathcal{F}^{-1}[u](\xi) dx \\
&= T_f\left(\mathcal{F}^{-1}[u]\right).
\end{aligned}$$

Again, the last line is well defined for any distribution if $u \in \mathcal{S}(\mathbb{R}^n)$. Then since $u \in \mathcal{S}(\mathbb{R}^n)$ was arbitrary we conclude that

$$T_{\mathcal{F}[f]}(u) = T_f(\mathcal{F}[u]) \quad \text{and} \quad T_{\mathcal{F}^{-1}[f]}(u) = T_f(\mathcal{F}^{-1}[u]) \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^n).$$

This result is then extended to all tempered distributions by the following definition;

Definition 1.2.10 ([15])

Let $T \in \mathcal{S}'(\mathbb{R}^n)$ we define

$$\mathcal{F}[T](u) = \hat{T}(u) = T(\hat{u}) = T(\mathcal{F}[u]),$$
$$\mathcal{F}^{-1}[T](u) = \check{T}(u) = T(\check{u}) = T(\mathcal{F}^{-1}[u])$$

for all $u \in \mathcal{S}(\mathbb{R}^n)$.

The following two results from [15] are analogous to Theorem 1.0.3, Theorem 1.0.9 and Proposition 1.0.7. We view this as confirmation that the Fourier Transform on the tempered distributions does in fact extend the Fourier transform on the Schwartz functions.

Propostion 1.2.11 ([15])

1. *For all $T \in \mathcal{S}'(\mathbb{R}^n)$ we have that $\mathcal{F}[T], \mathcal{F}^{-1}[T] \in \mathcal{S}'(\mathbb{R}^n)$.*

2. *For all $T \in \mathcal{S}'(\mathbb{R}^n)$ we have that*

$$\mathcal{F}^{-1}[\mathcal{F}[T]] = T$$
$$\mathcal{F}[\mathcal{F}^{-1}[T]] = T.$$

The next result extends the interaction of the Fourier transform with operations such as translation and differentiation from functions to distributions.

Propostion 1.2.12 ([15])

Fix the points $x_0, \xi \in \mathbb{R}^n$ and multi-indices $a, b \in \mathbb{Z}^n$. Then the following identities hold.

1. $\mathcal{F}[\tau_{x_0} T] = e^{-i\langle x_0, \xi\rangle}\hat{T}$.

2. $\mathcal{F}[e^{i\langle x, \xi_0\rangle} T] = \tau_{\xi_0}\hat{T}$.

3. $\mathcal{F}[D^a T] = \xi^a \hat{T}$.

4. $\mathcal{F}[(-x)^b T] = D^b \hat{T}$.

Now we have reviewed most of the properties of the Fourier transform that will be necessary for our future work. We are ready to start looking at the Wavelet transform in the following chapter. The results stated in this chapter will be used throughout the rest of this dissertation without much further explanation.

Chapter 2

The Wavelet Transform

In this chapter we will be following [5] loosely. Our main aim will be defining the *Wavelet* transform, as well as looking at the underlying intuition and implications. This chapter will not be an exhaustive overview of wavelet theory but rather a discussion of its main applications. We will be looking at other uses of wavelets such as regularity detection, but most of our time will be spent on function reconstruction formulas based on wavelets. The ideas in this chapter can also be extended to multiple dimensions but we will leave part of this extension to Chapter 3. Because of this, we will only work with functions of one variable. Before we look at wavelets we will look at a simpler predecessor of the wavelet Transform known as the *windowed Fourier* transform.

2.1 The windowed Fourier transform

Often times when we are applying the Fourier transform it is because we are specifically interested in the frequency content of a given signal (a time dependent function). However, we are usually not so interested in the global frequency content but rather the frequency content at a given moment in time. An applied example of this would be the extraction of noise from a signal. If we can figure out the largest frequency contribution to the noise we can remove this frequency from our signal by methods from Fourier analysis and hopefully obtain a clearer signal. However this noise might only have occurred once or twice. So we wouldn't want to remove the necessary frequency everywhere in time but rather at a few given instances. This is where the usual Fourier transform fails. It can be difficult to read off local frequency information from the global Fourier transform. The first fix for this is known as the *windowed Fourier* transform.

The idea of the *windowed Fourier* transform is to replace the global Fourier multiplier $e^{-i\langle x,\xi\rangle}$ by a multiplier that has compact support or at least rapid decay outside of a compact set. This new multiplier must also be sensitive to different frequencies. So we define the *windowed Fourier* transform by;

Definition 2.1.1 (Continuous windowed Fourier transform[5, Chapter 1 page 2]**)** *Let $f \in L^1(\mathbb{R}^n)$ and let g be a function with compact support then we define the continuous windowed Fourier transform of f as the two-dimensional function;*

$$[T^{win}f](\omega, t) := \frac{1}{2\pi}\int f(s)g(s-t)e^{-i\omega s}ds.$$

In the *windowed Fourier* transform ω again selects the frequency that we are analysing and t shifts the region which we analyse. By making the support of g larger or smaller we can also control the size of the region in time that we are focussing on. There also exists a discrete form of the *windowed Fourier* transform.

Definition 2.1.2 (Discrete windowed Fourier transform[5, Chapter 1 page 2]**)** *Let $f \in L^1(\mathbb{R}^n)$ and let g be a function with compact support. Also fix $\omega_0, t_0 > 0$ then we define the discrete windowed Fourier transform as*

$$T^{win}_{m,n}(f) := \frac{1}{2\pi}\int f(s)g(s-nt_0)e^{-im\omega_0 s}ds \quad m, n \in \mathbb{Z}.$$

The discrete transform is simply a special case of the continuous transform where we only shift the window and frequency between regularly spaced values. For fixed n the numbers $T^{win}_{m,n}(f)$ generated as m varies correspond to the Fourier coefficients of the function $f(\cdot)g(\cdot - nt_0)$.

As for the choice of the window function g in the windowed Fourier transform the most popular choice is the Gaussian Bump function. In all uses of the *windowed Fourier* transform we require the window function to be well localised in both the time domain and the

Fourier domain. Thus we want g and $\hat{g}$ to be concentrated on a bounded region. In the case that g and $\hat{g}$ are both concentrated around the origin we can interpret $[T^{\text{win}} f](\omega, t)$ as the frequency content of the function f around time t and frequency ω. The idea of the windowed Fourier transform is illustrated below in figure 2.1

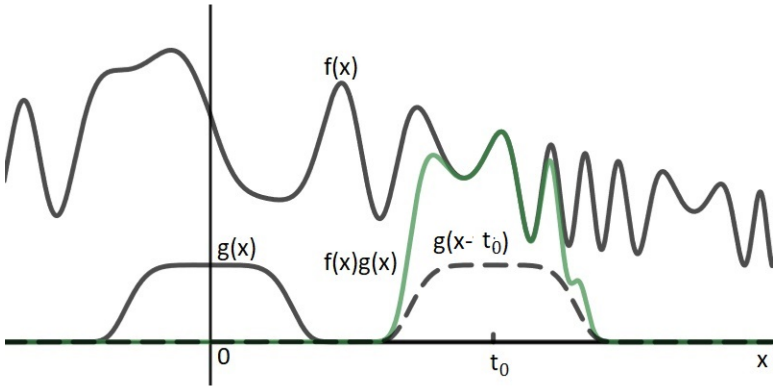


Figure 2.1: Illustration of windowed Fourier transform.

2.2 Continuous Wavelet transform

As with the Fourier transform in Chapter 1 we start our discussion of the *wavelet* transform by first defining it since the intuition is more clear when the definition is known.

Definition 2.2.1 (Adapted from [5] **chapter 1 page 3)**

For a given function $f : \mathbb{R} \to \mathbb{R}$ we define its wavelet transform to be the two-dimensional function

$$(T^{wav} f)(a, b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\psi \left(\frac{t-b}{a} \right)} dt, \quad a, b \in \mathbb{R}. \tag{2.1}$$

Where $\psi : \mathbb{R} \to \mathbb{R}$ is a function satisfying

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \tag{2.2}$$

While the Wavelet transform is defined in terms of a complex wavelet, for our purposes it will be sufficient and easier to consider real valued wavelets. In this case we will be able to drop the conjugation. We insist on 2.2 as it enforces some form of decay on our wavelet, as we discuss at the start of section 2.2.1. Notice that the wavelet transform of f can be viewed as the L^2 inner product of f and the doubly indexed family of functions

$$\psi^{a,b}(t) := |a|^{1/2} \psi \left(\frac{t-b}{a} \right). \tag{2.3}$$

Here we call the function ψ the "mother wavelet" and the functions $\psi^{a,b}$ are called "wavelets". We take the normalisation $|a|^{1/2}$ so that the L^2 norm of the wavelet is preserved, that is $||\psi^{a,b}|| = ||\psi||$. Preserving the L^2 norm of the wavelet simplifies calculations a little bit.

Common choices for ψ are; the Haar wavelet given by

$$\psi(t) = \chi_{(0,1/2)}(t) - \chi_{[1/2,1)}(t),$$

where χ_R is the indicator function for the region R. The Haar wavelets have compact support, but are certainly not smooth. Another common choice is the second derivative of the Gaussian, also known as the mexican hat function;

$$\psi(t) = (1 - t^2)e^{-\frac{t^2}{2}}.$$

The mexican hat is smooth, but has infinite support. An example of a wavelet that is smooth with compact support is the first derivative of the Gaussian bump function;

$$w(x) = \frac{d}{dx}e^{\frac{-1}{1-x^2}} = \frac{-2x}{(1-x^2)^2}e^{\frac{-1}{1-x^2}}.$$

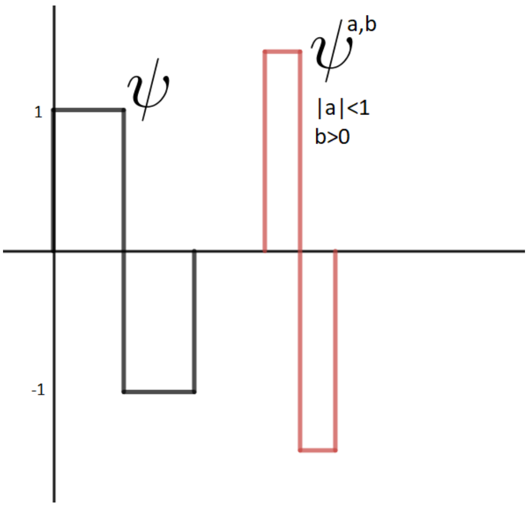
then the associated wavelet is

$$\psi(x) = \begin{cases} w(x) & \text{if } |x| < 1, \\ 0 & \text{otherwise.} \end{cases}$$

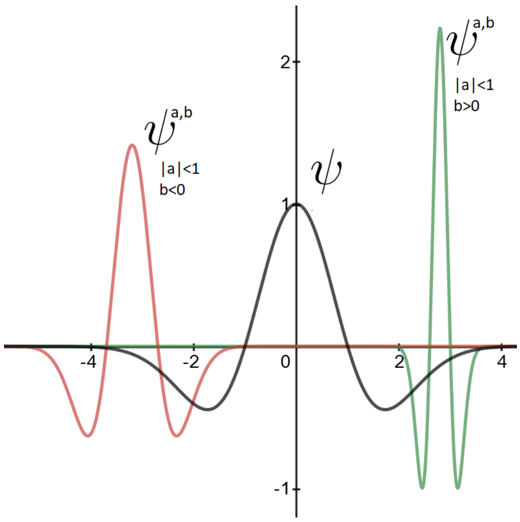
The graphs of these wavelets are shown in figure 2.2 below.

Also notice the *wavelet* transform and *windowed Fourier* transform are similar in the sense that both can be expressed as an L^2 inner product of f with some shifted function. However in figure 2.2 we see that the wavelet transform allows us to dilate (stretch and squeeze) and shift the wavelet. This dilation allows us to analyse smaller and smaller regions. However, the *windowed Fourier* transform has a fixed window size that gets shifted along. This is the main difference between the *windowed Fourier* transform and the *wavelet* transform.

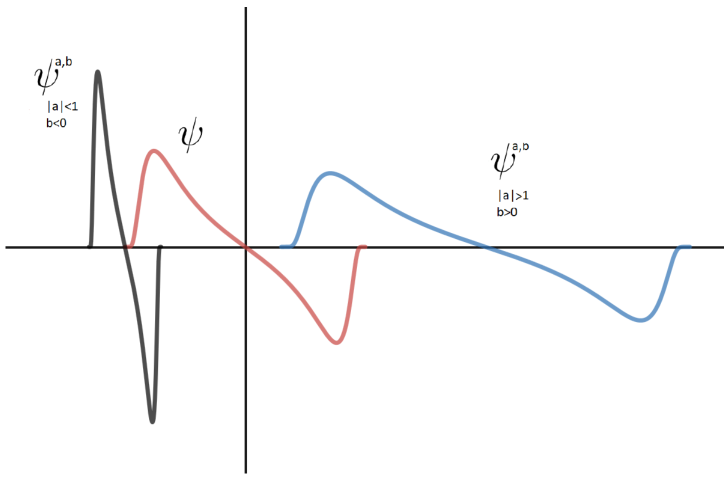
Figure 2.2: Three common wavelets.



(a) The Haar wavelet



(b) The mexican hat wavelet



(c) The derivative of the Gaussian Bump.

The *wavelet* transform in (2.1) also provides a time-frequency description of the function f. To see this we notice that changing the parameter a dilates the domain of ψ.
The smaller a is the higher the frequency the transform analyses, and the larger a is the lower the frequency is that the transform analyses. We can then also analyse the frequency content of f in different regions by changing the parameter b which shifts ψ.

One nice property of the *wavelet* transform is that it is relatively easy to prove reconstruction formulae relating a function f to its wavelet transform $[T^{wav} f](a,b)$.

2.2.1 Reconstruction Formulas

This subsection will follow Chapter 2.4 and 2.6 in [5] closely and the proofs presented here are adapted from the same chapter. Before we start stating results there are two more assumptions that we must place on our wavelet. The first assumption is that ψ is a unit vector in $L^2(\mathbb{R})$, that is $||\psi||_{L^2} = 1$. The reason for this assumption is that it makes the statements and proofs presented in [5] simpler. This is also not an unreasonable assumption since it is simply a case of normalization. The definition of $\psi^{a,b}$ in (2.3) also ensures that $||\psi^{a,b}||_{L^2} = ||\psi||_{L^2} = 1$. In this case Hölder's inequality ensures that $|[T^{wav}f](a,b)| \leq ||f||_{L^2}$. This means that we can view the wavelet transform as a bounded linear transform from $L^2(\mathbb{R})$ to $L^\infty(\mathbb{R}^2)$.

Our next assumption is a bit more stringent and will require some discussion. We define the admissibility condition;

$$C_\psi = 2\pi \int_{\mathbb{R}} |\xi|^{-1} |\hat{\psi}(\xi)|^2 d\xi < \infty. \tag{2.4}$$

This assumption is practically equivalent to (2.2) as we will explain below. First, if we assume that $\psi \in L^1(\mathbb{R})$ then the Lebesgue Dominated Convergence Theorem tells us that $\hat{\psi}$ is continuous and (2.4) implies that $\hat{\psi}(0) = 0$ which is equivalent to the assumption $\int_{\mathbb{R}} \psi(x)dx = 0$. The converse implication does not quite hold, we need one extra assumption. If we have that 2.2 holds then we also insist that for some $\alpha > 0$ we have;

$$\int_{\mathbb{R}} (1 + |x|)^\alpha |\psi(x)| dx < \infty. \tag{2.5}$$

From this it follows that there exists $C > 0$ such that $|\hat{\psi}(\xi)| \leq C|\xi|^{-\beta}$ with $\beta = \min(\alpha, 1)$. Then we have that (2.4) is satisfied. This shows that (2.4) and (2.2) are close to equivalent. However when it comes to actually using the wavelet transform we will often require much stricter decay behaviour. For example in Chapter 4 we will use a wavelet that is a smooth function with compact support.

Our first interesting result relating to wavelets comes in the form of an identity that can be viewed as some form of a reconstruction formula.

Propostion 2.2.2

For any $f, g \in L^2(\mathbb{R})$ we have that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(T^{wav}f)(a,b)\overline{(T^{wav}g)(a,b)}\frac{da\ db}{a^2} = C_\psi\langle f, g\rangle. \tag{2.6}$$

proof:(adapted from [5] Chapter 2 page 24)

Fix some $f, g \in L^2(\mathbb{R})$. Then we compute

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(T^{wav}f)(a,b)\overline{(T^{wav}g)(a,b)}\frac{da\ db}{a^2} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle f, \psi^{a,b}\rangle\overline{\langle g, \psi^{a,b}\rangle}\frac{da\ db}{a^2}\\ &\text{[Definition of } L^2 \text{ inner product and } (T^{wav}\cdot)(a,b).]\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle \hat{f}, \widehat{\psi^{a,b}}\rangle\overline{\langle \hat{g}, \widehat{\psi^{a,b}}\rangle}\frac{da\ db}{a^2}\\ &\text{[By Proposition 1.1.9]}\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\hat{f}(\xi)|a|^{1/2}e^{-ib\xi}\hat{\psi}(a\xi)d\xi\right]\\ &\qquad\left[\int_{-\infty}^{\infty}\overline{\hat{g}(\xi')}|a|^{1/2}e^{ib\xi'}\hat{\psi}(a\xi')d\xi'\right]\frac{da\ db}{a^2}\\ &\text{[By Theorem 1.1.3]}\end{aligned} \tag{2.7}$$

The ξ integral in (2.7) above can be interpreted as $(2\pi)^{1/2}$ times the Fourier transform (with frequency variable b) of the function

$$F_a(\xi) = |a|^{1/2}\hat{f}(\xi)\hat{\psi}(a\xi).$$

Similarly the ξ' integral can be viewed as $(2\pi)^{1/2}$ times the complex conjugate of the Fourier transform of

$$G_a(\xi') = |a|^{1/2}\hat{g}(\xi')\overline{\hat{\psi}(a\xi')}.$$

It then follows from proposition (1.1.9) that

$$\begin{aligned}(2.7) &= 2\pi\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F_a(\xi)\overline{G_a(\xi)}d\xi\frac{da}{a^2}\\ &= 2\pi\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(\xi)\overline{\hat{g}(\xi)}|\hat{\psi}(a\xi)|^2 d\xi\frac{da}{|a|}\\ &= 2\pi\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(\xi)\overline{\hat{g}(\xi)}|\hat{\psi}(a\xi)|^2\frac{da}{|a|}d\xi\end{aligned}$$

$$\begin{aligned}
&[\text{Fubini's Theorem allows the interchange}]\\
&= 2\pi \int_{-\infty}^{\infty} \hat{f}(\xi)\overline{\hat{g}(\xi)} \int_{-\infty}^{\infty} |\hat{\psi}(a\xi)|^2 \frac{da}{|a|} d\xi \qquad (2.8)\\
&= C_\psi \int_{-\infty}^{\infty} \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi\\
&[\text{Substitute } u = a\xi]\\
&= C_\psi \langle \hat{f}, \hat{g} \rangle\\
&= C_\psi \langle f, g \rangle.
\end{aligned}$$

Now (2.8) justifies why we insisted on (2.4). Without this we would not be able to make the identity in (2.6) hold. □

Recall that we called the result of Proposition 2.2.2 a reconstruction formula. What we mean by this needs some unpacking as it is clearly not a reconstruction formula in a point-wise sense. The way we need to interpret this identity is that

$$f = C_\psi^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (T^{wav} f)(a,b) \psi^{a,b} \frac{da\ db}{a^2} \qquad (2.9)$$

where equality holds in a weak sense. The particular weak equality in which (2.9) holds is defined in Definition 2.2.3 below

Definition 2.2.3

Let $f, g \in L^2(\mathbb{R}^n)$ then we say that $f = g$ in a weak sense if for all $h \in L^2(\mathbb{R}^n)$ we have that $\langle f, h \rangle_{L^2} = \langle g, h \rangle_{L^2}$.

The equality in (2.9) holds in this weak sense because of proposition 2.2.2. Equation (2.9) has a convergence that holds in a slightly stronger sense, given by the next proposition.

Propostion 2.2.4

For any $f \in L^2(\mathbb{R})$ we have

$$\lim_{\substack{A_1 \to 0 \\ A_2, B \to \infty}} \left\| f - C_\psi^{-1} \iint_{\substack{A_1 \le |a| \le A_2 \\ |b| \le B}} (T^{wav} f)(a,b) \psi^{a,b} \frac{da\ db}{a^2} \right\|_{L^2} = 0. \qquad (2.10)$$

proof:(Adapted and expanded from [5] Chapter 2 page 25-26)

In (2.10) we view the integral as a representative for the element $F_{A_1,A_2,B} \in L^2(\mathbb{R})$ that satisfies

$$\langle F_{A_1,A_2,B}, g\rangle = \iint\limits_{\substack{A_1\le|a|\le A_2\\ |b|\le B}} (T^{wav}f)(a,b)\langle\psi^{a,b}, g\rangle \frac{da\ db}{a^2}. \tag{2.11}$$

for all $g \in L^2(\mathbb{R})$. This element $F_{A_1,A_2,B}$ is unique due to the positive-definiteness and linearity of inner products. Now recall that the wavelet transform is a bounded linear operator with $|[T^{wav}f](a,b)| \le ||f||_{L^2}$. This along with the Triangle inequality for integrals and the Cauchy-Schwartz inequality then imply that the absolute value of (2.11) is bounded above by

$$\iint\limits_{\substack{A_1\le|a|\le A_2\\ |b|\le B}} ||f||_{L^2}\ ||\psi^{a,b}||_{L^2}\ ||g||_{L^2} \frac{da\ db}{a^2} = 4B\left(\frac{1}{A_1} - \frac{1}{A_2}\right)||f||_{L^2}||g||_{L^2}.$$

By the Riesz-Fréchet representation theorem we can then view

$$f - C_\psi^{-1} \iint\limits_{\substack{A_1\le|a|\le A_2\\ |b|\le B}} (T^{wav}f)(a,b)\psi^{a,b}\frac{da\ db}{a^2}$$

as a representative of a unique linear functional on L^2. The Riesz-Fréchet representation theorem also says that $||F_{A_1,A_2,B}||_{L^2} = ||F_{A_1,A_2,B}||_{op}$, which allows us to compute;

$$\left|\left| f - C_\psi^{-1} \iint\limits_{\substack{A_1\le|a|\le A_2\\ |b|\le B}} (T^{wav}f)(a,b)\psi^{a,b}\ \frac{da\ db}{a^2}\right|\right|_{op}$$

$$= \sup_{||g||=1}\left|\left\langle f - C_\psi^{-1} \iint\limits_{\substack{A_1\le|a|\le A_2\\ |b|\le B}} (T^{wav}f)(a,b)\psi^{a,b}\ \frac{da\ db}{a^2},\ g\right\rangle\right|$$

$$\le \sup_{||g||=1}\left| C_\psi^{-1} \iint\limits_{\substack{|a|\ge A_2\\ \text{or } |a|\le A_1\\ \text{or } |b|\ge B}} (T^{wav}f)(a,b)\overline{(T^{wav}g)(a,b)}\ \frac{da\ db}{a^2}\right|$$

$$\le \sup_{||g||=1}\left[C_\psi^{-1} \iint\limits_{\substack{|a|\ge A_2\\ \text{or } |a|\le A_1\\ \text{or } |b|\ge B}} |(T^{wav}f)(a,b)|^2\ \frac{da\ db}{a^2}\right]^{1/2} \left[C_\psi^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|(T^{wav}g)(a,b)|^2\ \frac{da\ db}{a^2}\right]^{1/2}.$$

By Proposition 2.2.2 the second term in the last line is $||g||_{L^2} = 1$ and the integral in the first term converges to zero as $A_1 \to 0$ and $A_2, B \to \infty$. To see this note that

$$C_\psi^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |(T^{wav} f)(a,b)|^2 \, \frac{da \, db}{a^2} = C_\psi^{-1} \iint\limits_{\substack{|a| \geq A_2 \\ \text{or } |a| \leq A_1 \\ \text{or } |b| \geq B}} |(T^{wav} f)(a,b)|^2 \, \frac{da \, db}{a^2} + C_\psi^{-1} \iint\limits_{\substack{A_1 \leq |a| \leq A_2 \\ |b| \leq B}} (T^{wav} f)(a,b) \psi^{a,b} \, \frac{da \, db}{a^2}.$$

Since the first integral on the right converges to the integral on the left as $A_1 \to 0$ and $A_2, B \to \infty$, we must have that the second integral on the right converges to zero. From this it follows that (2.10) holds. $\square$

From Proposition 2.2.4 and equation (2.10) we see that any function $f \in L^2(\mathbb{R})$ can be approximated arbitrarily close by a superposition of wavelets. However, there is a problem in that statement that needs to be addressed. The problem is that we are saying that a superposition of functions with zero integral can approximate a function with potentially non-zero integral as closely as we like. But if we take a superposition of functions with zero integral then the superposition still has zero integral. This can be seen by integrating equation (2.9) over $\mathbb{R}$ and commuting the integrals, allowed by Fubini's theorem. Does Proposition 2.2.4 then suggest that f also has zero integral and do we have a contradiction? No, we do not have a contradiction, but we do need to take a closer look at why that is.

The first step in resolving this apparent contradiction is by noticing that the limit in equation (2.10) is not meant to hold in the L^1 norm but rather in the L^2 norm. We know that these two spaces are not equivalent, which gives us some hope that that a rigorous explanation exists. This rigorous explanation is given in [5] which we now follow and expand on.

Consider a function $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and let us also assume that $\psi \in L^1(\mathbb{R})$. Then by Fubini's theorem we can commute integrals and apply the Cauchy-Schwartz inequality to show that

$$C_\psi^{-1} \iint\limits_{\substack{A_1 \le |a| \le A_2 \\ |b| \le B}} (T^{wav} f)(a,b)\psi^{a,b} \frac{da\ db}{a^2} \tag{2.12}$$

is in $L^1(\mathbb{R})$ for all $A_1, A_2, B > 0$. By doing so one would also find that the L^1 norm of (2.12) is bounded above by

$$8C_\psi^{-1}||f||_{L^2}||\ \psi||_{L^2}\ ||\psi||_{L^1} B \left(\frac{1}{A_1^{1/2}} - \frac{1}{A_2^{1/2}} \right).$$

By simply commuting the integrals again one can also check that (2.12) has zero integral. Now we can explain away the contradiction that equation (2.10) seems to imply. Consider the function

$$f - C_\psi^{-1} \iint\limits_{\substack{A_1 \le |a| \le A_2 \\ |b| \le B}} (T^{wav} f)(a,b)\psi^{a,b} \frac{da\ db}{a^2}.$$

As $A_1, \to 0$ and $A_2, B \to \infty$ this function becomes very flat and spread out. However, the function does not becomes flat enough or it does not decay fast enough to induce convergence in the L^1 norm. The function still has the same integral as f due to the linearity of integration, but it has arbitrarily small L^2 norm. So we really should not use or understand equation (2.10) in a point wise sense or as an L^1 approximation, but only understand it as a statement about L^2 norms. With this view the apparent contradiction implied by Proposition 2.2.4 is no longer an issue.

In Chapter 2 of [5] on pages 27 to 28 Daubechies states a few variations on equation (2.9) that still hold under certain hypotheses. However we do not wish to state and discuss all of these variations here. We will discuss two variations that will be relevant to our later work.

In some applications we will be concerned with wavelets that are symmetric over their domain and because of this we will only be interested in the size of the dilation a. Therefore we will be considering strictly positive a. It is then useful to adapt equation (2.9) to something that still holds for $a > 0$. In [5] Daubechies suggest a stricter admissibility condition than (2.4), namely

$$C_\psi = 2\pi \int_0^\infty |\xi|^{-1} |\hat{\psi}(\xi)|^2 d\xi = 2\pi \int_{-\infty}^0 |\xi|^{-1} |\hat{\psi}(\xi)|^2 d\xi < \infty. \tag{2.13}$$

The equality follows if we assume that ψ is real since $\hat{\psi}(-\xi) = \overline{\hat{\psi}(\xi)}$. Then she finds that

$$f = C_\psi^{-1} \int_0^\infty \int_{-\infty}^\infty (T^{wav} f)(a,b) \psi^{a,b} \frac{db\ da}{a^2}. \tag{2.14}$$

And we can interpret this identity in the weak sense of Definition 2.2.3 and slightly stronger sense of Proposition 2.2.4. The proof of this resolution identity is exactly the same as that of (2.9). The only difference is that we now need to show that

$$\lim_{\substack{A_1 \to 0 \\ A_2, B \to \infty}} \left\| f - C_\psi^{-1} \iint_{\substack{A_1 \le a \le A_2 \\ |b| \le B}} (T^{wav} f)(a,b) \psi^{a,b} \frac{da\ db}{a^2} \right\|_{L^2} = 0. \tag{2.15}$$

But as we said, this is done in exactly the same way by just replacing the necessary integrals.

Another resolution identity that Daubechies states in [5] is for analytical signals. An analytical signal f has supp $\hat{f} \subset [0, \infty)$. If we assume both f and ψ are analytical signals then by Proposition 1.1.9 we have that (2.10) immediately reduces to

$$f = C_\psi^{-1} \int_0^\infty \int_{-\infty}^\infty (T^{wav} f)(a,b) \psi^{a,b} \frac{db\ da}{a^2}. \tag{2.16}$$

Note here C_ψ is as in (2.4) not (2.13) and the resolution is again interpreted in the weak sense of Definition 2.2.3 and slightly stronger sense of Proposition 2.2.4.

There is one more variation which interests us since we will use it in Section 2.2.2. It is shown in [5] that if we reconstruct f with a different wavelet than we used to deconstruct it then we can derive a form of equation (2.9) that holds even pointwise under some circumstances. Suppose we have two functions ψ_1 and ψ_2 that satisfy the new admissibility condition

$$\int_{-\infty}^\infty |\xi|^{-1}\ |\hat{\psi}_1(\xi)|\ |\hat{\psi}_2(\xi)| d\xi < \infty. \tag{2.17}$$

We then also define the new constant C_{ψ_1,ψ_2} by

$$C_{\psi_1,\psi_2} := 2\pi \int_{-\infty}^{\infty} \xi^{-1}\ \overline{\hat{\psi}_1(\xi)}\ \hat{\psi}_2(\xi) d\xi. \tag{2.18}$$

Then by following the proof of Proposition 2.6 line by line one can show that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \langle f, \psi_1^{a,b}\rangle\ \langle \psi_2^{a,b}, g\rangle\ \frac{db\ da}{a^2} = C_{\psi_1,\psi_2}\langle f, g\rangle. \tag{2.19}$$

In this case we can rewrite (2.9) as

$$f = C_{\psi_1,\psi_2}^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \langle f, \psi_1^{a,b}\rangle\ \psi_2^{a,b}\ \frac{db\ da}{a^2} \tag{2.20}$$

to be interpreted in the usual weak sense. Here we have some more flexibility in our selection of ψ_1 and ψ_2. Neither have to be admissible wavelets, as long as their product is. Notice that if we can write $\hat{\psi}_1(\xi) = \xi\tilde{\psi}_1(\xi)$ and $\tilde{\psi}_1, \psi_2 \in L^2(\mathbb{R})$ then by Cauchy-Schwartz we don't need $\hat{\psi}_2(0) = 0$ in order to satisfy (2.17). We don't even need $\tilde{\psi}_1, \psi_2 \in L^2(\mathbb{R})$. By Hölder's inequality it is enough to have $\tilde{\psi}_1 \in L^q(\mathbb{R})$ and $\psi_2 \in L^p(\mathbb{R})$ if $q^{-1} + p^{-1} = 1$ and $p, q \in [1.\infty]$. Also, even if we only have $\hat{\psi}_1 = O(\xi)$ as $\xi \to 0$ then it is not necessary to have $\hat{\psi}_2(0) = 0$. So we see that deconstructing and reconstructing with different functions ψ_1 and ψ_2 allows us a lot more freedom in our selection of our analysing functions. This freedom will be used in the next section to derive some interesting results about local regularity of certain functions.

In [5] Daubechies states and proves a result from [8] which gives conditions under which (2.20) holds even pointwise. Below we restate the result and expand on the proof presented in [5].

Propostion 2.2.5 (taken from chapter 2 page 28 of [5]**)**

Suppose that $\psi_1, \psi_2 \in L^1(\mathbb{R})$, that ψ_2 is differentiable with $\psi_2' \in L^2(\mathbb{R})$, that $x\psi_2 \in L^1(\mathbb{R})$, and that $\hat{\psi}_1(0) = \hat{\psi}_2(0) = 0$. Then if $f \in L^2(\mathbb{R})$ and bounded then for every x where f is continuous equation (2.20) holds point wise. That is;

$$f(x) = C_{\psi_1,\psi_2}^{-1} \lim_{\substack{A_1\to 0\\ A_2\to\infty}} \int_{A_1\le|a|\le A_2}\int_{-\infty}^{\infty} \langle f, \psi_1^{a,b}\rangle\ \psi_2^{a,b}(x)\ \frac{db\ da}{a^2}. \tag{2.21}$$

proof:(Adapted from [5] Chapter 2 page 29- 30)

This proof will be broken up into 4 parts progressively building up to the main result.

1) We start by noticing that before taking the limit we can write (2.21) as

$$f_{A_1,A_2}(x) = C^{-1}_{\psi_1,\psi_2} \int_{A_1 \le |a| \le A_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y)|a|^{-1} \overline{\psi_1\left(\frac{y-b}{a}\right)}\, \psi_2\left(\frac{x-b}{a}\right)\, dy \frac{db\ da}{a^2}$$

$$= C^{-1}_{\psi_1,\psi_2} \int_{-\infty}^{\infty} \int_{A_1 \le |a| \le A_2} \int_{-\infty}^{\infty} f(y)|a|^{-1} \overline{\psi_1\left(\frac{y-b}{a}\right)}\, \psi_2\left(\frac{x-b}{a}\right)\, \frac{db\ da}{a^2} dy \tag{2.22}$$

$$= \int_{-\infty}^{\infty} M_{A_1,A_2}(x-y) f(y) dy. \tag{2.23}$$

Because of (2.17) these integrals above converge absolutely and so Fubini's theorem allows us to change the order of integrals as we wish. The function M_{A_1,A_2} is defined by

$$M_{A_1,A_2}(x) = C^{-1}_{\psi_1,\psi_2} \int_{A_1 \le |a| \le A_2} \int_{-\infty}^{\infty} f(y)|a|^{-1} \overline{\psi_1\left(-\frac{b}{a}\right)}\, \psi_2\left(\frac{x-b}{a}\right)\, \frac{db\ da}{|a|^3}. \tag{2.24}$$

This definition is derived by substituting $\tilde{b} = b - y$ in the inner double integral of (2.22).

2) By writing M_{A_1,A_2} in the form of (2.24), we can view the b integral in M_{A_1,A_2} as convolution. Then by commuting integrals (allowed by Fubini's theorem) we can use proposition (1.1.13) to easily find that the Fourier transform of M_{A_1,A_2} is given by

$$\hat{M}_{A_1,A_2}(\xi) = (2\pi)^{1/2} C^{-1}_{\psi_1,\psi_2} \int_{A_1 \le |a| \le A_2} \hat{\psi}_1(a\xi)\ \overline{\hat{\psi}_1}(a\xi) \frac{da}{|a|} \tag{2.25}$$

$$= \hat{M}(A_1\xi) - \hat{M}(A_2\xi). \tag{2.26}$$

In 2.26 we defined $\hat{M}$ by

$$\hat{M}(\xi) = (2\pi)^{1/2} C^{-1}_{\psi_1,\psi_2} \int_{|a| \ge |\xi|} \hat{\psi}_2(a)\ \overline{\hat{\psi}_1}(a) \frac{da}{|a|}. \tag{2.27}$$

This definition follows from (2.25) by making the change of variables $\tilde{a} = a\xi$. Now recall that we have assumed that $a\hat{\psi}_2(a) \in L^2(\mathbb{R})$ and $\hat{\psi}_1$ is bounded. This allows us to apply the Cauchy-Schwartz inequality to find

$$|\hat{M}(\xi)| \le C \left(\int_{|a| \ge |\xi|} |\hat{\psi}_1(a) \frac{da}{|a|^4}| \right)^{1/2} \left(\int_{-\infty}^{\infty} |a|^2 |\hat{\psi}_2(a)|^2 \right)^{1/2} \tag{2.28}$$

$$\le C' ||a\hat{\psi}_2||_{L^2}\ |\xi|^{-3/2}. \tag{2.29}$$

By (2.17) and the triangle inequality for integrals we also have that $\hat{M}$ is bounded. From this it follows that

$$|\hat{M}(\xi)| \le C(1 + |\xi|)^{-3/2}. \tag{2.30}$$

So we have that $\hat{M} \in L^1(\mathbb{R})$ and then from Theorem 1.1.2 we have that M, the inverse

Fourier transform of $\hat{M}$, is well defined, bounded and continuous.

3) Next we try to understand the decay of M, this is inevitably related to the regularity of $\hat{M}$. By breaking the definition of $\hat{M}$ into $a < 0$ and $a \geq 0$ one finds that for $\xi \neq 0$ the function $\hat{M}$ is differentiable with respect to ξ with derivative given by

$$\frac{d}{d\xi}\hat{M}(\xi) = (2\pi)^{1/2} C^{-1}_{\psi_1,\psi_2} \frac{-1}{\xi} \left[\hat{\psi}_2(\xi)\overline{\hat{\psi}_1(\xi)} + \hat{\psi}_2(-\xi)\overline{\hat{\psi}_1(-\xi)}\right]. \tag{2.31}$$

From the Riemann-Lebesgue Theorem (Theorem 19.13 in Chapter 19 on page 222 in [14]) since $x\psi_2 \in L^1(\mathbb{R})$ we know that $\hat{\psi}_2$ is differentiable on all of $\mathbb{R}$. Then for $\xi = 0$ we have that

$$\frac{d}{d\xi}\hat{M}\bigg|_{\xi=0} = -(2\pi)^{1/2}\ C^{-1}_{\psi_1,\psi_2} 2\ \overline{\hat{\psi}_1(0)}\ \hat{\psi}_2{}'(0) = 0.$$

From this it follows that $\hat{M}$ is differentiable on all of $\mathbb{R}$. Furthermore, from $x\psi_2 \in L^1(\mathbb{R})$ it follows that

$$\begin{aligned}
|\hat{\psi}_2(\xi)| &= |\hat{\psi}_2(\xi) - \hat{\psi}_2(0)| \\
&\leq C \int_{-\infty}^{\infty} |e^{-i\xi x} - 1||\psi_2(x)|dx \\
&\leq C|\xi| \int_{-\infty}^{\infty} |x\psi_2(x)|dx \\
&\leq C'|\xi|.
\end{aligned}$$

We can then conclude (2.31) that

$$\left|\frac{d}{d\xi}\hat{M}(\xi)\right| \leq C' \left[|\hat{\psi}_1(\xi)| + |\hat{\psi}_1(-\xi)|\right]. \tag{2.32}$$

Since we assumed that $\psi_1 \in L^2(\mathbb{R})$ it then follows that $\frac{d}{d\xi}\hat{M} \in L^2(\mathbb{R})$. The Cauchy-Schwartz inequality now implies that

$$\begin{aligned}
\int_{-\infty}^{\infty} |M(x)|dx &= \int_{-\infty}^{\infty} (1+x^2)^{1/2}(1+x^2)^{-1/2}|M(x)|dx && (2.33)\\
&\leq \left[\int_{-\infty}^{\infty} (1+x^2)|M(x)|^2 dx\right]^{1/2} \left[\int_{-\infty}^{\infty} (1+x^2)^{-1} dx\right]^{1/2} && (2.34)\\
&\leq C\left[\int_{-\infty}^{\infty} \left(|\hat{M}(\xi)|^2 + \left|\frac{d}{d\xi}\hat{M}(\xi)\right|^2\right) d\xi\right]^{1/2} && (2.35)\\
&< \infty. && (2.36)
\end{aligned}$$

This implies that $M \in L^1(\mathbb{R})$ and in particular this means that $\hat{M}$ is well defined and

$$\begin{aligned}
\hat{M}(0) &= (2\pi)^{1/2} C_{\psi_1,\psi_2}^{-1} \int_{-\infty}^{\infty} \hat{\psi}_2(a)\ \overline{\hat{\psi}_1}(a) \frac{da}{|a|} \\
&= (2\pi)^{1/2} \left(2\pi \int_{-\infty}^{\infty} \hat{\psi}_2(a)\ \overline{\hat{\psi}_1}(a) \frac{da}{|a|} \right)^{-1} \int_{-\infty}^{\infty} \hat{\psi}_2(a)\ \overline{\hat{\psi}_1}(a) \frac{da}{|a|} \\
&= (2\pi)^{-1/2}.
\end{aligned}$$

By the way we normalized the Fourier transform it then follows that

$$\int_{-\infty}^{\infty} M(x)dx = 1.$$

4.) Now using (2.26) we can rewrite (2.23) as

$$f_{A_1,A_2}(x) = \int_{-\infty}^{\infty} \frac{1}{A_1} M\left(\frac{x-y}{A_1}\right) f(y)dy - \int_{-\infty}^{\infty} \frac{1}{A_2} M\left(\frac{x-y}{A_2}\right) f(y)dy.$$

If f is bounded and continuous in x then the fact that M is continuous, integrable and has integral 1 means, by the dominated convergence theorem, that the first term tends to $f(x)$ as A_1 tends to zero. We can bound the second term by

$$\begin{aligned}
\left| \int_{-\infty}^{\infty} \frac{1}{A_2} M\left(\frac{x-y}{A_2}\right) f(y)dy \right| &\leq \left[\int_{-\infty}^{\infty} \frac{1}{A_2^2} \left| M\left(\frac{x-y}{A_2}\right)\right|^2 dy \right]^{1/2} \left[\int_{-\infty}^{\infty} |f(y)|^2 dy \right]^{1/2} \\
&\leq A_2^{-1/2} ||M||_{L^2}\ ||f||_{L^2} \\
&< C A_2^{-1/2}.
\end{aligned}$$

Where the last inequality follows from the fact that $M \in L^2(\mathbb{R})$. This implies that the second term tends to zero as A_2 tends to infinity. Therefore $f_{A_1,A_2} \to f$ (point-wise) as $A_1 \to 0$ and $A_2 \to \infty$. □

2.2.2 Characterising local regularity.

In the previous section we saw how wavelets can be used to reconstruct a given function f. However, these reconstructions don't tell us much about the behaviour of f. In this section we take a look at some ways in which wavelets can be used to analyse local behaviour of functions. In particular we will see how wavelets can characterise local regularity of functions. Our discussion will follow Section 9 of chapter 2 in [5], which is itself taken from [8]. The first result relates the continuity of a given function to the growth of its wavelet transform.

Theorem 2.2.6 (Taken from Chapter 2 page 45 of [5]**)**

Suppose that $\int_{-\infty}^{\infty}(1+|x|)|\psi(x)| < \infty$ *and* $\hat{\psi}(0) = 0$. *Let* f *be a bounded function that is Hölder continuous with Hölder exponent* $0 < \alpha \leq 1$, *that is;*

$$|f(x) - f(y)| \leq C|x-y|^{\alpha}.$$

Then its wavelet transform can be estimated by

$$|T^{wav}(a,b)| = \langle f, \psi^{a,b}\rangle| \leq C'|a|^{\alpha+1/2}.$$

proof: Since we have that $\hat{\psi}(0) = 0$ we also have

$$\int_{-\infty}^{\infty} \psi(x)dx = 0.$$

This simply means that ψ is a wavelet. It then follows, by adding zero, that;

$$[T^{wav}f](a,b) = \int_{-\infty}^{\infty} |a|^{-1/2}\psi\left(\frac{x-b}{a}\right)[f(x) - f(b)]dx.$$

Then by Hölder continuity substituting $y = a^{-1}(x-b)$ we find

$$\begin{aligned} |[T^{wav}f](a,b)| &\leq \int_{-\infty}^{\infty} |a|^{-1/2}\psi\left(\frac{x-b}{a}\right) C|x-b|^{\alpha}dx \\ &\leq C|a|^{\alpha+1/2}\int_{-\infty}^{\infty} |\psi(y)||y|^{\alpha}dx \\ &\leq C'|a|^{\alpha+1/2}. \end{aligned}$$

□

Next we will see that using a compactly supported wavelet allows us to determine the continuity type of functions in a subspace of $L^2(\mathbb{R})$.

Theorem 2.2.7 (Taken from Chapter 2 page 46 of [5]**)**

Suppose that ψ *is compactly supported. Suppose also that* $f \in L^2(\mathbb{R})$ *is bounded and continuous. If for some* $\alpha \in (0,1)$ *the wavelet transform of* f *satisfies*

$$|[T^{wav}f]((a,b))| \leq C|a|^{\alpha+1/2} \tag{2.37}$$

then f *is Hölder continuous with exponent* α.

proof: The proof of this result is again broken up into four smaller parts building up to the main result.

1) We start by choosing a second wavelet ψ_2 so that it is compactly supported and

continuously differentiable with $\int_{-\infty}^{\infty} \psi_2(x)dx = 0$. Also normalise ψ_2 so that $C_{\psi,\psi_2} = 1$, where C_{ψ,ψ_2} is defined as in (2.18). Then by Proposition 2.2.5 we have that

$$f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(x) \frac{db\ da}{a^2}.$$

Next, we split the integral over a into two parts, $|a| \leq 1$ and $|a| > 1$. Then define

$$f_{SS}(x) = \int_{|a|\leq 1} \int_{-\infty}^{\infty} \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(x) \frac{db\ da}{a^2},$$
$$f_{LS}(x) = \int_{|a|>1} \int_{-\infty}^{\infty} \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(x) \frac{db\ da}{a^2}.$$

Here f_{SS} corresponds to an analysis of small scale details of f and f_{LS} to an analysis of large scale details of f.

2) First we notice that

$$\begin{aligned} |f_{LS}(x)| &\leq \int_{|a|\geq 1} \int_{-\infty}^{\infty} |\psi_2^{a,b}(x)|\ ||f||_{L^2}\ ||\psi||_{L^2} \frac{db\ da}{a^2} \\ &\leq C \int_{|a|\geq 1} \int_{-\infty}^{\infty} \left| \psi_2 \left(\frac{x-b}{a} \right) \right| \frac{db\ da}{a^2} \\ &\leq C ||\psi_2||_{L^1} \int_{|a|\leq 1} |a|^{-3/2} < \infty. \end{aligned}$$

This implies that f_{LS} is bounded uniformly in x. Next we consider the difference $|f_{LS}(x+h) - f_{LS}(x)|$ for any $|h| \leq 1$. Then we see that

$$\begin{aligned} &|f_{LS}(x+h) - f_{LS}(x)| \\ &\leq \int_{|a|\geq 1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(y)| \left| \psi \left(\frac{y-b}{a} \right) \right| \left| \psi_2 \left(\frac{x+h-b}{a} \right) - \psi_2 \left(\frac{x-b}{a} \right) \right| dy\ db\ da. \quad (2.38) \end{aligned}$$

The fact that ψ_2 is continuously differentiable implies that $|\psi_2(z+t) - \psi_2(z)| \leq C|t|$. Since ψ and ψ_2 are compactly supported we have $\operatorname{supp} \psi, \operatorname{supp} \psi_2 \subset [-R, R]$ for some finite R. Now we can bound (2.38) as follows;

$$\begin{aligned} |f_{LS}(x+h) - f_{LS}(x)| &\leq C'|h| \int_{|a|\geq 1} \iint_{\substack{|x-b|\leq|a|R+1 \\ |y-b|\leq|a|R}} |f(y)| \frac{dy\ db\ da}{a^4} \\ &\leq C''|h| \int_{|a|\geq 1} \int_{|y-x|\leq 2|a|R+1} |f(y)| \frac{dy\ da}{|a|^3} \\ &\leq C''|h|\ ||f||_{L^2} \int_{|a|\geq 1} |a|^{-3}(4|a|R+2)^{1/2} da \\ &\leq C'''|h|. \end{aligned}$$

Note that h was arbitrary, meaning that this bound holds for all $|h| \leq 1$. This bound and the fact that f_{LS} is uniformly bounded means we can conclude that $|f_{LS}(x+h) - f_{LS}(x)| \leq C|h|$ for all $|h| \leq 1$, uniformly in x. This means that f_{LS} is not just continuous but also Lipschitz continuous. This implies that f_{LS} is also Hölder continuous with exponent α, since $|h| \leq 1$. Note that we didn't use the bound (2.37), thus the large scale part f_{LS} of f is always continuous.

3) We can also see that the small scale part f_{SS} of f is uniformly bounded:

$$\begin{aligned}
|f_{SS}(x)| &\leq C \int_{|a|\leq 1} \int_{-\infty}^{\infty} |a|^{\alpha+1/2}|a|^{-1/2} \left|\psi_2\left(\frac{x-b}{a}\right)\right| \frac{db\ da}{a^2} \\
&\leq C||\psi_2||_{L^1} \int_{|a|\leq 1} |a|^{-1+\alpha} da \\
&= C' < \infty.
\end{aligned}$$

4) Finally, since f_{SS} is uniformly bounded we only need to check $|f_{SS}(x+h) - f_{SS}(x)|$ for small h, say $|h| \leq 1$. Using $|\psi_2(z+t) - \psi_2(z)| \leq C|t|$ again we find that

$$\begin{aligned}
|f_{SS}(x+h) - f_{SS}(x)| \leq &\int_{|a|\leq|h|} \int_{-\infty}^{\infty} |a|^{\alpha} \left(\left|\psi_2\left(\frac{x-b}{a}\right)\right| + \left|\psi_2\left(\frac{x+h-b}{a}\right)\right|\right) \frac{db\ da}{a^2} \\
&+ \int_{|h|\leq|a|\leq 1} \int_{|x-b|\leq|a|R+|h|} |a|^{\alpha} \left(\left|\psi_2\left(\frac{x-b}{a}\right)\right| + \left|\psi_2\left(\frac{x+h-b}{a}\right)\right|\right) \frac{db\ da}{a^2} \\
&\qquad\qquad\qquad\qquad [\text{Since } \operatorname{supp}\psi_2 \subset [-R,R].] \\
\leq &\int_{|a|\leq|h|} \int_{-\infty}^{\infty} |a|^{\alpha} \left(\left|\psi_2\left(\frac{x-b}{a}\right)\right| + \left|\psi_2\left(\frac{x+h-b}{a}\right)\right|\right) \frac{db\ da}{a^2} \\
&+ \int_{|h|\leq|a|\leq 1} \int_{|x-b|\leq|a|R+|h|} |a|^{\alpha} C \left|\frac{h}{a}\right| \frac{db\ da}{a^2} \\
&\qquad\qquad\qquad\qquad [\text{Since } |\psi_2(z+t) - \psi_2(z)| \leq C|t|] \\
\leq &\, C'\left[||\psi_2||_{L^2} \int_{|a|\leq|h|} |a|^{-1+\alpha} da + |h| \int_{|h|\leq|a|\leq 1} |a|^{-3+\alpha}(|a|R + |h|)\right] \\
= &\, C''|h|^{\alpha}.
\end{aligned}$$

Then since the sum of two Hölder continuous functions is another Hölder continuous function it follows that f is Hölder continuous with exponent α. □

Together, Theorems 2.2.6 and 2.2.7 show that the global Hölder continuity of a bounded function $f \in L^2(\mathbb{R})$ is completely characterised by the decay in a of its wavelet transform. However, if f is only Lipschitz continuous ($\alpha = 1$) we can only make conclusions about the decay of the wavelet transform. In this case the wavelet transform doesn't tell us much about the global continuity of f.

Recall that before Proposition 2.2.5 we noted that our wavelet ψ need not be regular. The same is true for Theorems 2.2.6 and 2.2.7. We mainly used that ψ has zero integral and some decay conditions such as $\psi \in L^2(\mathbb{R})$, $\int_{-\infty}^{\infty}(1+|x|)|\psi(x)| < \infty$ or that ψ is compactly supported. Such wavelets are easy to come by, for example the Haar wavelet, the Mexican hat wavelet or the derivative of the Guassian bump. We never explicitly assumed that ψ has zero integral, however it is possible for the estimate (2.37) to fail if this is not the case. We can also characterise the differentiability of f and Hölder continuity of f by the decay of it's wavelet transform in a. In [5] Daubechies states Theorem 2.2.8 below without proving it.

Theorem 2.2.8 (Adapted from Chapter 2 page 48 of [5]**)**
Let ψ be a wavelet such that $\int_{\mathbb{R}} x^m \psi(x)dx = 0$ for all $m = 0, ..., n$ for some $n \in \mathbb{N}$. Also let $\alpha \in (0,1)$ then the following two statements are equivalent.

- *$f \in C^n$ with all $f^{(m)}$, $m = 0, ..., n$ bounded and square integrable, and $f^{(n)}$ Hölder continuous with exponent α.*
- *$|\langle f, \psi^{a,b}\rangle| \leq C|a|^{n+1/2+\alpha}$ uniformly in a.*

Note, as before we require no regularity for the wavelet ψ.

From the three previous theorems we see that wavelets are quite effective at analysing the global regularity of functions. However, since wavelets can be shifted and dilated they can deal with intervals of any size centred at any point. This localised nature of wavelets allows us to analyse local regularity of functions as well. We can see this clearly in the next two results stated in [5] and originally proven in [8].

Theorem 2.2.9 (Taken from Chapter 2 page 49 from [5])

Suppose we have that $\int_{\mathbb{R}}(1+|x|)|\psi(x)|dx < \infty$ *and* $\int_{\mathbb{R}} \psi(x)dx = 0$. *If a function* f *is bounded and Hölder continuous at* x_0 *with exponent* $\alpha \in (0,1)$, *that is,*

$$|f(x_0+h) - f(x_0)| \leq C|h|^{\alpha},$$

then we have

$$|\langle f, \psi^{a,x_0+b}\rangle| \leq C|a|^{1/2}(|a|^{\alpha} + |b|^{\alpha}).$$

proof: Since Hölder continuity is completely preserved under translation we may, without loss of generality, assume that $x_0 = 0$. Just as in the proof of Theorem 2.2.6 since $\int_{\mathbb{R}} \psi(x)dx = 0$ we have that

$$\begin{aligned}
|\langle f, \psi^{a,b}| &\leq \int |f(x) - f(0)||a|^{-1/2}\left|\psi\left(\frac{x-b}{a}\right)\right| dx \\
&\leq C\int_{\mathbb{R}} |x|^{\alpha}|a|^{-1/2}\left|\psi\left(\frac{x-b}{a}\right)\right| dx && \text{[Hölder continuity at } x=0\text{]} \\
&\leq C|a|^{\alpha+1/2}\int \left|y + \frac{b}{a}\right|^{\alpha} |\psi(y)|dy && \text{[Substitute } y = (x-b)a^{-1}\text{]} \\
&\leq C'|a|^{1/2}(|a|^{\alpha} + |b|^{\alpha}).
\end{aligned}$$

By repeating this calculation with extra substitution $u = x - x_0$ one finds the main result. □

The next result is not a fully converse statement to Theorem 2.2.9, but it does allow us to deduce Hölder continuity from conditions similar to the conclusion of Theorem 2.2.9 .

Theorem 2.2.10 (Taken from Chapter 2 page 49 from [5])

Suppose that ψ *is a compactly supported wavelet. Also let* $f \in L^2(\mathbb{R})$ *be bounded and continuous. If, for some* $\gamma > 0$ *and* $\alpha \in (0,1)$ *we have*

$$|\langle f, \psi^{a,b}\rangle| \leq C|a|^{\gamma+1/2} \quad \textit{uniformly in } b, \tag{2.39}$$

and

$$|\langle f, \psi^{a,b+x_0}\rangle| \leq C|a|^{1/2}\left(|a|^{\alpha} + \frac{|b|^{\alpha}}{|\ln|b||}\right), \tag{2.40}$$

then f *is Hölder continuous at* x_0 *with exponent* α.

proof: This proof is again broken into successive parts building up to the main result.

1. The first part of this proof is really three parts, however the first three parts are exactly the first three parts of Theorem 2.2.7. The only difference is that in part three of that proof we now replace α with γ. So we take the integral representation of f

$$f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(x) \frac{db\ da}{a^2}$$

and split it into a small scale part f_{SS} for $|a| \leq 1$ and a large scale part f_{LS} for $|a| \geq 1$. The regularity of f_{LS} is completely dealt with in the first two parts in the proof of Theorem 2.2.7.

2. Boundedness of f_{SS} is dealt with in part three in the proof of Theorem 2.2.7. So we only need to check $|f_{SS}(x_0 + h) - f_{SS}(x_0)|$. In the same way as in the proof of Theorem 2.2.9 we may assume without loss of generality that $x_0 = 0$. Then we find by (2.39) and (2.40) that

$$\begin{aligned}
|f_{SS}(h) - f_{SS}(0)| &= \left| \int_{|a|\leq 1} \int_{-\infty}^{\infty} \langle f, \psi^{a,b} \rangle \left(\psi_2^{a,b}(h) - \psi_2^{a,b}(0) \right) \frac{db\ da}{a^2} \right| \\
&\leq \int_{|a|\leq |h|^{\alpha/\gamma}} \int_{-\infty}^{\infty} \left| \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(h) \right| \frac{db\ da}{a^2} + \int_{|h|^{\alpha/\gamma}\leq |a| \leq |h|} \int_{-\infty}^{\infty} \left| \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(h) \right| \frac{db\ da}{a^2} \\
&\quad + \int_{|a|\leq |h|} \int_{-\infty}^{\infty} \left| \langle f, \psi^{a,b} \rangle \psi_2^{a,b}(0) \right| \frac{db\ da}{a^2} \\
&\quad + \int_{|h|\leq |a| \leq 1} \int_{-\infty}^{\infty} \left| \langle f, \psi^{a,b} \rangle \left(\psi_2^{a,b}(h) - \psi_2^{a,b}(0) \right) \right| \frac{db\ da}{a^2} \\
&\leq \underbrace{\int_{|a|\leq |h|^{\alpha/\gamma}} \int_{-\infty}^{\infty} |a|^{\gamma} \left| \psi_2 \left(\frac{h-b}{a} \right) \right| \frac{db\ da}{a^2}}_{=:T_1} \qquad \text{[By (2.39)]} \\
&\quad + \underbrace{\int_{|h|^{\alpha/\gamma}\leq |a| \leq |h|} \int_{-\infty}^{\infty} \left(|a|^{\alpha} + \frac{|b|^{\alpha}}{|\ln |b||} \right) \left| \psi_2 \left(\frac{h-b}{a} \right) \right| \frac{db\ da}{a^2}}_{=:T_2} \qquad \text{[By (2.40)]} \\
&\quad + \underbrace{\int_{|a|\leq |h|} \int_{-\infty}^{\infty} \left(|a|^{\alpha} + \frac{|b|^{\alpha}}{|\ln |b||} \right) \left| \psi_2 \left(-\frac{b}{a} \right) \right| \frac{db\ da}{a^2}}_{=:T_3} \qquad \text{[By (2.40)]} \\
&\quad + \underbrace{\int_{|h|\leq |a| \leq 1} \int_{-\infty}^{\infty} \left(|a|^{\alpha} + \frac{|b|^{\alpha}}{|\ln |b||} \right) \left| \psi_2 \left(\frac{h-b}{a} \right) - \psi_2 \left(-\frac{b}{a} \right) \right| \frac{db\ da}{a^2}}_{=:T_4}, \qquad \text{[By (2.40)]}
\end{aligned}$$

where we assumed that $\alpha > \gamma$. If $\alpha \leq \gamma$ then $T_2 = 0$, which adds nothing to the working that follows below. So by assuming $\alpha > \gamma$ we lose no generality. To finish the proof we will show that each term is bounded (up to a constant) by $|h|^{\alpha}$ from which Hölder continuity follows.

3. Boundedness of T_1 follows from direct computation;

$$T_1 \leq \int_{|a|\leq |h|^{a/\gamma}} |a|^{\alpha-1}||\psi_2||_{L^1} \leq C|h|^{\alpha}.$$

4. Recall that ψ_2 is has compact support. This means that supp $\psi_2 \subset [-R, R]$ for some $R > 0$. Then we find

$$\begin{aligned} T_2 &\leq \int_{|a|\leq|h|} |a|^{\alpha-1}||\psi_2||_{L^1} da \\ &+ \int_{|h|^{\alpha/\gamma}\leq|a|\leq|h|} |a|^{-1}||\psi_2||_{L_1} \frac{(|a|R+h)^{\alpha}}{|\ln(|a|R+|h|)|} da \\ &\leq C|h|^{\alpha}\left[1 + \frac{1}{|\ln|h||}\int_{h^{\alpha/\gamma}\leq|a|\leq|h|} |a|^{-1} da\right] \\ &\leq C'|h|^{\alpha}. \end{aligned}$$

5. For sufficiently small $|h|$ we see that

$$\begin{aligned} T_3 &\leq \int_{|a|\leq h} |a|^{-1+\alpha}||\psi_2||_{L^1} da + \int_{|a|\leq|h|} |a|^{-1}||\psi_2||_{L^2} \frac{(|a|R)^{\alpha}}{|\ln(|a|R)|} da \\ &\leq C|h|^{\alpha}. \end{aligned}$$

6. Finally we compute

$$\begin{aligned} T_4 &\leq C|h| \int_{|h|\leq|a|\leq 1} |a|^{-3}\left[|a|^{\alpha} + \frac{(|a|R+|h|)^{\alpha}}{|\ln(|a|R+|h|)|}\right](|a|R+|h|) da \\ &\leq C'[1 + |h|^{-1+\alpha} + |h|(1+|h|^{\alpha-2})] \\ &\leq C''|h|^{\alpha}. \end{aligned}$$

We can summarise these last five parts (2.−6.) in the statement

$$|f_{SS}(h) - f_{SS}(0)| \leq C|h|^{\alpha}.$$

It then follows that f is Hölder continuous at $x_0 = 0$. By performing the appropriate variable substitutions in the working above we can show Hölder continuity for arbitrary x_0 satisfying the hypothesis of the theorem. □

These last two theorems show us how the localised nature of wavelets allows us to analyse functions locally in order to extract regularity information. This regularity information would not necessarily be so easy to extract from a function definition either. In particular, determining from first principles whether a function is Hölder continuous can be difficult or even impossible, but now wavelets provide us with one more tool to deal with such problems.

In [5] Daubechies uses these last two theorems to justify nicknaming wavelets a "mathematical microscope". Before we end this chapter we state one more theorem that further justifies the name "mathematical microscope". We state Theorem 2.2.11 without proof since we are not so interested in what we can learn about wavelets from the proof of the theorem. However, we wish to highlight the effectiveness of wavelets to analyse local behaviour of functions. The proof of this theorem can be found in the textbook from which it is adapted.

Theorem 2.2.11 (Adapted from page 160 of [13]**)**

Let ψ be a wavelet such that $\psi \in C^n$, with compact support and $\psi = (-1)^n \Psi^{(n)}$ where $\int_{\mathbb{R}} \Psi(x)dx \neq 0$. Suppose that f is locally integrable, that is $f \in L^1[c,d]$ for some $c < d$. If there exists some $a_0 > 0$ such that $|\langle f, \psi^{a,b} \rangle|$ has no local maximum for $b \in [c,d]$ and $a < a_0$ then for any $\varepsilon > 0$ the following holds.

There exists a constant $K > 0$ such that for every $x \in [c+\varepsilon, d-\varepsilon]$ there exists a polynomial p_x of degree n such that

$$\forall t \in \mathbb{R}, \qquad |f(t) - p_x(t)| \leq K|t-x|^n.$$

The conclusion of Theorem 2.2.11 essentially says that at every point in $[c,d]$ there is a small neighbourhood on which f can be arbitrarily closely approximated by a degree n polynomial. So wavelets don't just enable us to determine the local regularity of a function of one variable but also its general shape around a point. In [13] many more interesting theorems of this form are stated and proven. Sadly, discussing these theorems in detail would take more time and space than we have available here in this dissertation. Instead, in the next chapter we take a look at another transform that relies on the sensitivity to local regularity of the wavelet transform. This transform is known as the *Shearlet Transform*

Chapter 3

The Shearlet Transform

In the previous chapter we focussed on the analysis of one-dimensional functions, but obviously this is only a small subset of all possible functions that we want to understand. The wavelet transform can be easily extended to arbitrary finite dimensions and we will start this chapter by doing so. However, when we extend the wavelet transform to higher dimensions it also fails in some ways. In particular, wavelets cannot detect local geometry very effectively. Like the wavelet transform fixes the Fourier transform's lack of localisation we will fix the wavelet's lack of geometric sensitivity with a new transform called the the *Shearlet Transform*.

3.1 Multi-dimensional Wavelet Transform

Extending the wavelet transform to $n-$dimensions is done in a way similar to the extension of the Fourier transform to $n-$dimensions. That is, we simply replace the variables and measures in the definitions with $n-$dimensional vectors and measures, and we normalise the wavelets appropriately. Then we define the $n-$dimensional wavelet transform as follows;

Definition 3.1.1 (Adapted from Chapter 2 of [5] and page 13 of [11])
For a given function $f : \mathbb{R}^n \to \mathbb{R}$ we define its wavelet transform to be the $(n+1)-$dimensional function

$$[T^{wav} f](a,t) := |a|^{-n/2} \int_{\mathbb{R}^n} f(t) \overline{\psi\left(\frac{x-t}{a}\right)} dx \tag{3.1}$$

where $a > 0$, $t \in \mathbb{R}^n$ and $\psi : \mathbb{R}^n \to \mathbb{C}$ is a function satisfying

$$\int_{\mathbb{R}^n} \psi(x) dx = 0. \tag{3.2}$$

With this the admissibility condition (2.4) becomes

$$C_\psi = \int_{\mathbb{R}^n} |\xi|^{-1} |\hat{\psi}(\xi)|^2 d\xi < \infty.$$

Notice in (3.1) that the dilation of the wavelet support is isotropic, that is, it is dilated by exactly the same amount in all directions. When working with functions of one variable this isotropic dilation is fine. Since the phenomena seen in functions of one variable are one dimensional, they posses no directionality. When we work with a multi-variable function f it is well known that f may be smooth in one direction while being highly irregular in another. The lack of directionality in the wavelet transform prevents it from detecting such anisotropic phenomena (phenomena with intrinsic directionality).

Theorem 3.1.2
Let ψ be a continuously differentiable real-valued $2-$dimensional wavelet with bounded derivatives and supp $\psi = B_1(0) = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 < 1\}$. Let $\gamma : U \to \mathbb{R}^2$, with $U \subseteq \mathbb{R}$ be a positively oriented differentiable curve that does not self-interesect. Let δ_γ be the delta distribution along γ. Then the following statements hold

1. *If $t_0 \notin \gamma(U)$ then as a function of a the wavelet transform $[T^{wav}\delta_\gamma](a,t_0)$ has rapid decay as $a \to 0$*

2. *If $t_0 \in \gamma(U)$ then $[T^{wav}\delta_\gamma](a,t_0) = O(1)$ as $a \to 0$.*

proof: The proof of this theorem is not terribly difficult and follows from direct computations which we outline below. First we note that for any $a > 0$ and $t \in \mathbb{R}^2$ we have that supp $\psi^{a,t} = B_a(t)$. Then we assume that γ is parametrised by arclength and we find

$$\begin{aligned}[T^{wav}\delta_\gamma](a,t) &= a^{-1}\int_{\mathbb{R}^2}\delta_\gamma\psi\left(\frac{x-t}{a}\right)dx\\ &= a^{-1}\int_U \psi\left(\frac{\gamma(u)-t}{a}\right)du\\ &= a^{-1}\int_{U_{a,t}} \psi\left(\frac{\gamma(u)-t}{a}\right)du,\end{aligned}$$

where $S_{a,t}\subseteq S$ is defined by

$$\gamma(u)\in B_a(t) \iff u\in U_{a,t}.$$

1. Suppose first that $t_0\notin\gamma(U)$, then since we can view $\gamma(S)$ as a closed subset of $\mathbb{R}^2$, we have that $\gamma(U)^c$ is open. So there exists $a_0>0$ so that $U_{a,t_0}=\emptyset$ for all $a\leq a_0$. So the wavelet transform becomes an integral over an empty set, this implies that
$[T\delta_\gamma](a,t_0)=0$ for all $a\leq a_0$, from which it follows that $[T\delta_\gamma](a,t)$ has rapid decay as $a\to 0$.

2.Next suppose that $t_0\in\gamma(U)$, and assume that $\gamma(0)=t_0$. Then for a small enough, since γ does not self intersect we have that $U_{a,t_0}=(u_1,u_2)$ for some $u_1<u_2$ in U. Then we have that

$$[T^{wav}\delta_\gamma](a,t_0) = a^{-1}\int_{u_1}^{u_2}\psi\left(\frac{\gamma(u)-t_0}{a}\right)du.$$

Integration by parts then gives

$$\begin{aligned}[T^{wav}\delta_\gamma](a,t_0) =& a^{-1}u\psi\left(\frac{\gamma(u)-t_0}{a}\right)\Bigg|_{u_1}^{u_2}\\ &- a^{-2}\int_{u_1}^{u_2} u\left(\frac{\partial\psi}{\partial x_1}\left(\frac{\gamma(u)-t_0}{a}\right)\frac{d\gamma_1}{du}(u)+\frac{\partial\psi}{\partial x_2}\left(\frac{\gamma(u)-t_0}{a}\right)\frac{d\gamma_2}{ds}(u)\right)du.\end{aligned}$$

Now suppose at $t_0=(t_1,t_2)$, the tangent line of γ has gradient m. Then for small a we have that

$$\gamma(s)\approx(u+t_1, mu+t_2)$$

and

$$\frac{d\gamma_1}{du}(u)\approx 1 \quad\text{and}\quad \frac{d\gamma_m}{du}(u)\approx m$$

Since γ is parametrised in terms of arclength we have that $u_1 \approx -a$ and $u_2 \approx a$. Then we have that for small a

$$\begin{aligned}
[T^{wav}\delta_\gamma](a,t_0) \approx &\psi\left((1,m)\right) + \psi\left(-(1,m)\right) \\
& - a^{-2}\int_{-a}^{a} s\left(\frac{\partial\psi}{\partial x_1}\left(\frac{(u,mu)}{a}\right)\frac{d\gamma_1}{ds}(u) + \frac{\partial\psi}{\partial x_2}\left(\frac{(u,mu)}{a}\right)\frac{d\gamma_2}{du}(u)\right)du. \\
& \lesssim \psi\left((1,m)\right) + \psi\left(-(1,m)\right) - a^{-2}C\int_{-a}^{a} u(1+m)du \\
& = \psi\left((1,m)\right) + \psi\left(-(1,m)\right) + O(1) \qquad (3.3) \\
& = O(1).
\end{aligned}$$

□

In part 2. of the proof of theorem 3.1.2 we can see that the wavelet transform would be able to distinguish between curves with varying gradients at t_0. Changing the gradient of γ at t_0 would change the value of m in (3.3). This means that the wavelet transform will not be able to tell the difference between curves through t_0 if they have the same gradient at t_0. Thus the wavelet transform can tell us that there exists a singularity at t_0, but it can't tell us much about the local geometry of the singularity curve beyond its gradient at t_0. This inability of the wavelet transform to detect local geometry is due to its isotropic nature, it has no directional variation. In order to detect local geometry of curves effectively, we will need a transform that has directional sensitivity. One such transform is the *Shearlet* Transform.

3.2 The Continuous Shearlet Transform

The continuous shearlet transform is analogous to the wavelet transform in the sense that we shift and dilate the support of a given analysing function. However, shearlets try to capture anisotropic detail by taking asymmetric dilations and rotating the support of the analysing functions. We use matrix operators for these transformations.

Definition 3.2.1

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function and let $D \in GL_2(\mathbb{R})$, that is D is an invertible 2×2 matrix. Then we define the linear operator T_D by;

$$T_D f(x) = |detD|^{-1} f(D^{-1}x), \quad x \in \mathbb{R}^2.$$

We call T_D the transformation operator associated with the matrix D.

Remark 3.2.2

If we have multiple matrices $A_1, ..., A_n \in GL_2(\mathbb{R})$ recall from linear algebra that

$$det(A_1...A_n) = \prod_{i=1}^{n} det(A_i)$$

moreover,

$$(A_1...A_n)^{-1} = A_n^{-1}...A_1^{-1} \quad and \quad det[(A_1...A_n)^{-1}] = \prod_{i=1}^{n} det(A_i)^{-1}$$

Now for any function $f : \mathbb{R}^2 \to \mathbb{R}$ we have

$$T_{A_n}...T_{A_1} f(x) = \left[\prod_{i=1}^{n} det(A_i)^{-1}\right] f(A_n^{-1}...A_1^{-1}x) = det[(A_1...A_n)^{-1}] f((A_1...A_n)^{-1}x) = T_{A_1...A_n} f(x).$$

From this we see that the composition of matrix operators results in the matrix operator induced by the product of the respective inducing matrices. Also notice that the matrix operators inherit the lack of commutativity from the underlying matrix structure.

Using these transformation operators we can write the wavelet $\psi^{a,b}$ as

$$\psi^{a,b} = T_{A_a} \tau_b \psi$$

where

$$A_a = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

and τ_b is the translation operator defined in Theorem 1.1.3.2. However this dilation is symmetric along the $x, y-$axis, so we define a new matrix

$$A_{a,\alpha} = \begin{pmatrix} a & 0 \\ 0 & a^{\alpha} \end{pmatrix}$$

where $\alpha \in (0, 1)$ with associated transformation operator $T_{A_{a,\alpha}}$. This transform scales the x-axis by a factor of a and the $y-$axis by a factor of a^{α}. The value of α determines the order of the dilation. If $\alpha = \frac{1}{2}$ then we say that $A_{a,\alpha}$ induces parabolic scaling, if $\alpha = \frac{1}{3}$ then we say that $A_{a,\alpha}$ induces cubic scaling, and so on and so forth. When working with shearlets the most common choice is $\alpha = \frac{1}{2}$ as this has nice properties when one wants to discretise the shearlet transform. See [10] and [12] for more on the Discreet Shearlet Transform.

Next we consider how to rotate the domain of our analysing function. The most obvious choice would be a rotation by some $\theta \in (0, 2\pi)$. The matrix that induces this transformation would be

$$A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Unfortunately rotations do not preserve the integer lattice unless $\theta = \pm\frac{k\pi}{2}$ for some integer k. We want to preserve the integer lattice to make it easier to transition from a continuous shearlet transform to a discrete transform. For this reason we choose to shear the support of the analysing function along the $x-$axis. The shearing matrix is given by

$$S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

for some $s \in \mathbb{R}$. We translate the shearlet in the same way we translated the wavelet, using the translation operator τ_t for any $t \in \mathbb{R}^2$. Now we can define the shearlet transform

Definition 3.2.3

Let $\psi \in L^2(\mathbb{R}^2)$ then we define the continuous shearlet system $SH(\psi)$ generated by ψ to be the collection

$$SH(\psi) := \{\psi_{a,s,t} = a^{-3/4}\tau_t T_{A_a} T_{S_s}\psi \;|a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2\},$$

where A_a is a parabolic dilation matrix. For $f \in L^2(\mathbb{R}^2)$ we define the shearlet transform by

$$[\mathcal{SH}_\psi f](a,s,t) = \langle f, \psi_{a,s,t}\rangle = \int_{\mathbb{R}^2} f(x)\psi_{a,s,t}(x)dx. \quad a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2.$$

As with wavelets we also define an admissibility condition for shearlets. The admissibility condition is defined by restricting the growth of $\hat{\psi}$ in order to ensure ψ has certain regularity and decay properties.

Definition 3.2.4

We say that $\psi \in L^2(\mathbb{R}^2)$ is an admissible shearlet if

$$\int_{\mathbb{R}^2} \frac{|\hat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1 < \infty.$$

Notice that constructing an admissible shearlet is quite easy. Any function ψ such that $\hat{\psi}$ is compactly supported away from $\xi_1 = 0$ is an admissible shearlet. For example, shifting the the first derivative of the Gaussian bump function. If we take ψ to have

$$\hat{\psi}(\xi) = \begin{cases} \frac{-2(||\xi||-1)}{(1-(||\xi||-1)^2)^2} e^{\frac{-1}{1-(||\xi||-1)^2}} & \text{if } ||\xi|| < 1, \\ 0 & \text{otherwise.} \end{cases}$$

then ψ is an admissible shearlet. An important example of an admissible shearlet is called a classical shearlet. Classical shearlets have a little more structure which makes them useful in calculations and application.

Definition 3.2.5 (Adapted from [11]**)**

Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where $\psi_1 \in L^2(\mathbb{R})$ is a discrete wavelet in the sense that it satisfies

$$\sum_{j \in \mathbb{Z}} |\hat{\psi}_1(2^{-j}\xi)|^2 = 1 \quad \text{for a.e. } \xi \in \mathbb{R}.$$

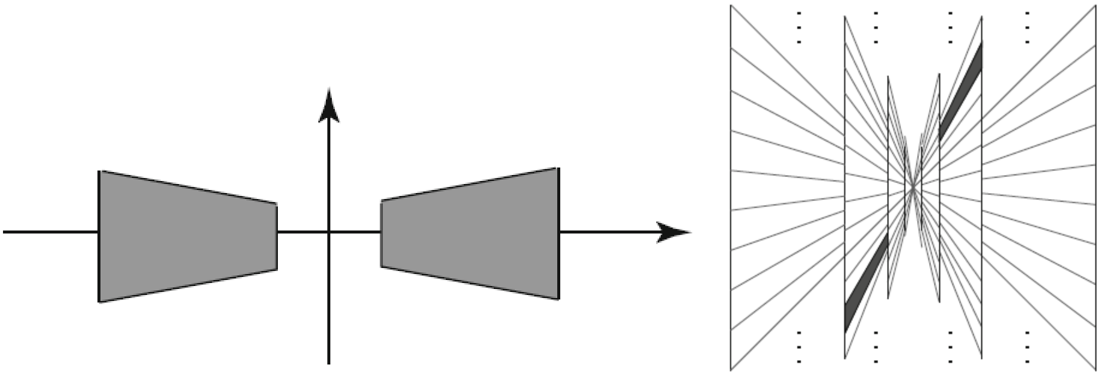
This is known as the Calderón condition. Furthermore $\hat{\psi}_1$ is smooth and compactly supported with supp $\hat{\psi}_1 \subseteq \left[\frac{-1}{2}, \frac{-1}{16}\right] \cup \left[\frac{1}{16}, \frac{1}{2}\right]$. We also require ψ_2 to be a bump function in the sense that

$$\sum_{k=-1}^{1} |\hat{\psi}_2(\xi + k)|^2 = 1 \quad \text{for a.e. } \xi \in [-1, 1]$$

with $\hat{\psi}_2$ smooth and compactly supported, in particular, supp $\hat{\psi}_2 \subseteq [-1, 1]$. If ψ satisfies all of these we call ψ a classical shearlet.

Notice that any classical shearlet has fourier tranform compactly supported away from zero, which means that it is an admissible shearlet. Classical shearlets are wavelet-like along one axis and bump-like along another. This wavelet-like nature of shearlets can be very useful to us in since it will allows us to apply our knowledge of wavelets to shearlets.

Figure 3.1: Classical shearlet.(Taken from [11])



(a) Support of the Fourier transform of a classical shearlet

(b) Support of the Fourier transform of several elements of a classical shearlet system for different values of a and s.

In figure 3.1b we can see that the support in the fourier domain of $\hat{\psi}$ is aligned with the ξ_1 axis for $s = 0$ and is rotated and stretched as s increases. However, notice that the support is only aligned with the ξ_2 axis in the limit as s goes to infinity. Even when the support is aligned with the ξ_2 axis it will be infinitely long and infinitely narrow. So all frequency information off of the ξ_2 axis will be undetected. This is a problem since this means that the classical shearlet system will be unable to detect functions or distributions that are mostly concentrated on the ξ_2-axis. This behaviour can clearly be problematic in practice. If we have a function or distribution that is simple and smooth away from the ξ_2-axis but highly irregular with a lot of directional variation close to the ξ_2-axis the classical shearlet system will not be a very effective analysis tool. One solution to this problem would be to swap the x_1- and x_2-axis in the spatial domain. This will also swap the ξ_1- and ξ_2-axis in the frequency domain.But, we keep the classical shearlets aligned as in Definition 3.2.5. However, we still can't deal with functions and distribution that have a lot of directional variation everywhere. One fix for this directional bias is the Cone-Adapted Shearlet System.

3.3 Cone-Adapted Shearlets

Cone-Adapted Shearlet systems are created by breaking up the frequency domain into five regions. We separate the low-frequency region from the high-frequency region by placing a square around the origin and by partitioning the high-frequency region into four cones. Two cones aligned with the ξ_1-axis and two cones aligned with the ξ_2-axis.

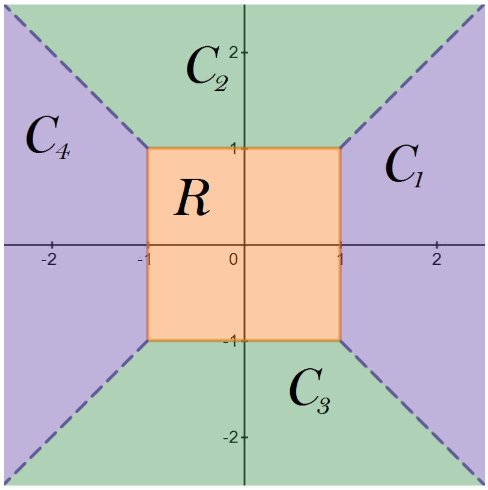


Figure 3.2: Resolving directional bias of continuous shearlet systems by splitting the frequency domain into four cones $C_i,\ i = 1, ..., 4$ and the low-frequency square $R = \{(\xi_1, \xi_2) \in \mathbb{R}^2 | \max(|\xi_1|, |\xi_2| \leq 1)\}$

.

Definition 3.3.1

For $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ we define the continuous shearlet system

$$SH(\phi, \psi, \tilde{\psi}) = \Phi(\phi) \cup \Psi(\psi) \cup \tilde{\Psi}(\tilde{\psi}),$$

where

$$\Phi(\phi) = \{\phi_t = \phi(\cdot - t) : t \in \mathbb{R}^2\},$$

$$\Psi(\psi) = \{\psi_{a,s,t} = a^{-3/4}\psi(A_a^{-1}S_s^{-1}(\cdot - t)) : a \in (0,1],\ |s| \leq 1 + a^{1/2}, t \in \mathbb{R}^2\}$$

$$\tilde{\Psi}(\tilde{\psi}) = \{\tilde{\psi}_{a,s,t} = a^{-3/4}\tilde{\psi}(\tilde{A}_a^{-1}\left(S_s^{-1}\right)^T(\cdot - t)) : a \in (0,1],\ |s| \leq 1 + a^{1/2}, t \in \mathbb{R}^2\}$$

where

$$\tilde{A}_a = \begin{pmatrix} a^{1/2} & 0 \\ 0 & a \end{pmatrix}.$$

In the same way we defined a shearlet transform associated with the continuous shearlet system we can now define a transform associated to the cone-adapted continuous shearlet system.

Definition 3.3.2

First define

$$\mathbb{S}_{cone} := \{(a,s,t) : a \in (0,1],\ |s| \leq 1 + a^{1/2},\ t \in \mathbb{R}^2\}$$

Then, for $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ with associated cone-adapted continuous shearlet system we define the cone-adapted continuous shearlet transform of $f \in L^2(\mathbb{R}^2)$ to be the mapping

$$f \mapsto [\mathcal{SH}_{\phi,\psi.\tilde{\psi}} f](t', (a,s,t), (\tilde{a}, \tilde{s}, \tilde{t})) := (\langle f, \phi_{t'} \rangle, \langle f, \psi_{a,s,t} \rangle, \langle f, \tilde{\psi}_{\tilde{a},\tilde{s},\tilde{t}} \rangle),$$

where

$$(t', (a,s,t), (\tilde{a}, \tilde{s}, \tilde{t})) \in \mathbb{R}^2 \times \mathbb{S}^2_{cone}.$$

It is worth taking some time to understand what each function ϕ, ψ and $\tilde{\psi}$ is associated with. Usually we choose ϕ to have compact support around the origin. This ensures that ϕ is associated with the low-frequency region $R = \{(\xi_1, \xi_2) : |\xi_1|, |\xi_2| \leq 1\}$. Then we associate the system $\Psi(\psi)$ with the horizontal cones by choosing a classical shearlet ψ, that is ψ satisfies Definition 3.2.5. Finally we associate the system $\tilde{\Psi}(\tilde{\psi})$ with the vertical cones by choosing $\tilde{\psi}$ to be a vertical classical shearlet, that is $\tilde{\psi}$ satisfies Definition 3.2.5 if we reverse the roles of ξ_1 and ξ_2.

Using the cone-adapted continuous shearlet transform we are now able to deal with functions or distribution with a lot of directionality . The horizontal cones can now detect directional variation along the ξ_1-axis at high frequencies while the vertical cones now deal with directional variation along the ξ_2-axis at high-frequencies. The low-frequency region is added since classical shearlets miss this region since they are supported away from the origin in the frequency domain.

3.4 Reconstruction formulas

Much like we were able to represent functions on $\mathbb{R}$ as a superposition of wavelets in Chapter 2, we can reconstruct functions on $\mathbb{R}^2$ as superpositions of shearlets. These reconstruction formulas will be central to the work we do in Chapter 4. Due to the extra shearing variable s that now appears in the transform and the added dimension of $\mathbb{R}^2$

many of these results will be a little bit more involved. But, they are very useful nonetheless.

We start by considering the continuous shearlet transform. We can view the continuous shearlet transform as a linear transformation between function spaces. In particular, the shearlet transform takes a function $f \in L^2(\mathbb{R}^2)$ to a function $[\mathcal{SH}_\psi f](a,s,t)$ defined on a subset of $\mathbb{R}^4$. Knowing when the shearlet transform is an isometry is important since it immediately leads to a reconstruction formula, given in Theorem 3.4.5 below.

Propostion 3.4.1 (Adapted from [11]**)**

First define the set

$$\mathcal{S} := (0,\infty) \times \mathbb{R} \times \mathbb{R}^2$$

and a function space

$$L^2_{\mathcal{S}} := L^2\left(\mathcal{S}, \frac{da}{a^{-3}}dsdt\right) := \left\{f : \mathcal{S} \to \mathbb{R} : \int_{\mathcal{S}} |f(a,s,t)|^2 \frac{da}{a^{-3}}dsdt < \infty\right\}.$$

For an admissible shearlet $\psi \in L^2(\mathbb{R})$ we define

$$C^+_\psi := \int_0^\infty \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi_1,\xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1 \quad \text{and} \quad C^-_\psi := \int_{-\infty}^0 \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi_1,\xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1.$$

If $C^+_\psi = C^-_\psi = 1$ then the mapping $\mathcal{SH}_\psi : L^2(\mathbb{R}^2) \to L^2_{\mathcal{S}}$ is a well defined isometry.

proof:[Proof adapted from [7]]

This result follows from calculation, but first we need to notice one general fact about the shearlet transform. For any $(a,s,t) \in \mathcal{S}$ we have

$$\begin{aligned}
[\mathcal{SH}_\psi f](a,s,t) &= \int_{\mathbb{R}^2} a^{-3/4}\psi(A_a^{-1}S_s^{-1}(x-t))f(x)dx \\
&= \int_{\mathbb{R}^2} f(x)\psi_{a,s,0}(x-t)dx \\
&= \int_{\mathbb{R}^2} f(x)\psi_{a,s,0}(-(t-x))dx \\
&= (f \star \psi^*_{a,s,0})(t)
\end{aligned}$$

where $\psi^*(x) = \psi(-x)$. In fact the same can actually be done for wavelets as well.

Now we compute

$$
\begin{aligned}
||\mathcal{SH}_\psi f||_{L^2_{\mathcal{S}}} &= \int_{\mathcal{S}} |[\mathcal{SH}_\psi f](a,s,t)|^2 \frac{da}{a^{-3}} ds dt \\
&= \int_{\mathcal{S}} |f \star \psi_{a,s,0}(t)|^2 dt ds \frac{da}{a^{-3}} \\
&= \int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}^2} |\hat{f}(\xi)||\widehat{\psi^*_{a,s,0}}|^2 d\xi ds \frac{da}{a^{-3}} \qquad \text{[By Theorem 1.1.4 (Plancherel)]} \\
&= \int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}^2} |\hat{f}(\xi)| a^{-3/2} |\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 + s\xi_1))|^2 ds d\xi da.
\end{aligned}
$$

Then we substitute $\omega_1 = a\xi_1$ and $\omega_2 = a^{1/2}(\xi_2 + s\xi_1)$ we find that

$$
\begin{aligned}
||\mathcal{SH}_\psi f||_{L^2_{\mathcal{S}}} = &\int_{\mathbb{R}} \int_0^\infty |\hat{f}(\xi)|^2 d\xi_1 d\xi_2 \int_0^\infty \int_{\mathbb{R}} \frac{|\psi(\omega_1 \hat{,} \omega_2)|}{\omega_1^2} d\omega_1 d\omega_2 \\
&+ \int_{\mathbb{R}} \int_{-\infty}^0 |\hat{f}(\xi)|^2 d\xi_1 d\xi_2 \int_{-\infty}^0 \int_{\mathbb{R}} \frac{|\psi(\omega_1 \hat{,} \omega_2)|}{\omega_1^2} d\omega_1 d\omega_2 \\
&= C_\psi^+ \int_{\mathbb{R}} \int_0^\infty |\hat{f}(\xi)|^2 d\xi_1 d\xi_2 + C_\psi^- \int_{\mathbb{R}} \int_{-\infty}^0 |\hat{f}(\xi)|^2 d\xi_1 d\xi_2 \\
&= \int_{\mathbb{R}^2} |\hat{f}(\xi)|^2 d\xi_1 d\xi_2 \qquad [C_\psi^+ = C_\psi^- = 1] \\
&= ||\hat{f}||_{L^2} \\
&= ||f||_{L^2}
\end{aligned}
$$

and therefore

$$
||\mathcal{SH}_\psi f||_{L^2_{\mathcal{S}}} = ||f||_{L^2}
$$

for all $f \in L^2(\mathbb{R}^2)$, which means $\mathcal{SH}$ is an isometry. □

As always, for Proposition 3.4.1 to be meaningful and useful we need to find an admissible shearlet ψ to which it applies, in other wods we need a ψ that satisfies $C_\psi^+ = C_\psi^- = 1$. Proposition 3.4.2 provides whole class of shearlets for which this is true.

Propostion 3.4.2

Let $\psi \in L^2(\mathbb{R}^2)$ be a classical shearlet, then $C_\psi^+ = C_\psi^- = 1$ and the shearlet transform associated with ψ is an isometry.

The proof of this follows directly from the definition of a classical shearlet. Since we won't really need Proposition 3.4.2, we ommit it's proof.

Remark 3.4.3

Any isometry between two inner product spaces where the norm is induced by the inner product preserves inner products. Therefore, for any admissible shearlet $\psi \in L^2(\mathbb{R}^2)$ we have that

$$\langle f, g \rangle_{L^2} = \langle \mathcal{SH}f, \mathcal{SH}g \rangle_{L^2_{\mathbb{S}}}.$$

Proposition 3.4.1 immediately allows us to prove a simple shearlet reconstruction formula in Theorem 3.4.5. However the proof of Theorem 3.4.5 uses a concept of from functional analysis known as an *approximate identity.* So we define a special case of an *approximate identity* in $L^2(\mathbb{R}^2)$.

Definition 3.4.4 (Adapted from [3]**)**

Let $\phi \in L^1(\mathbb{R}^2)$ with

$$\int_{\mathbb{R}^2} \phi(x)dx = 1.$$

For each $t > 0$ define the function ϕ_t by

$$\phi_t(x) = t^{-n}\phi\left(\frac{x}{t}\right)$$

then it can be shown by a change of variables that $||\phi_t||_{L^1} = ||\phi||_{L^1}$. The sequence $\{\phi_t\}$ is referred to as an approximate identity. It can then be shown that for $1 \le p < \infty$, the sequence $\{\phi_t \star f\}$ converges to f in $L^p(\mathbb{R}^2)$, that is;

$$\lim_{t\to 0} ||\phi_t \star f - f||_{L^p} = 0.$$

As a consequence, a subsequence $\{\phi_{t_k} \star f\}$ converges pointwise to f almost everywhere.

Using this new concept of an approximate identity we can prove Theorem 3.4.5 below.

Theorem 3.4.5 (Taken from [4]**)**

Suppose $\psi \in L^2(\mathbb{R}^2)$ is an admissible shearlet with $c_\psi^+ = c_\psi^- = 1$. Let $\{\rho_n\}_{n=1}^\infty$ be an approximate identity such that $\rho_n \in L^2(\mathbb{R}^2)$ and $\rho_n(-x) = \rho_n(x)$ for all x. Then for all $f \in L^2(\mathbb{R}^2)$

$$\lim_{n\to\infty} ||f - f_n||_{L^2} = 0,$$

where

$$f_n(x) = \int_{\mathbb{R}^2}\int_{\mathbb{R}}\int_0^1 [\mathcal{SH}f](a, s, t)(\rho_b \star \psi_{a,s,t})(x)a^{-3}dadsdt.$$

proof:[Adapted from [4]]

First recall that for $x \in \mathbb{R}^2$ we define the translation operator τ_x by $\tau_x f(y) = f(x-y)$.

Now we compute

$$
\begin{aligned}
(f \star \rho_n)(x) &= \int_{\mathbb{R}^2} f(y)\rho_n(x-y)dy \\
&= \langle f, \tau_x \rho \rangle_{L^2} \\
&= \langle \mathcal{SH}f, \mathcal{SH}\tau_x\rho_n \rangle_{L^2_{\mathcal{S}}} && \text{[By remark 3.4.3]} \\
& \int_{\mathbb{R}^2}\int_{\mathbb{R}}\int_0^1 [\mathcal{SH}f](a,s,t)\langle \phi(\cdot - x), \psi_{a,s,t}\rangle a^{-3} dadsdt \\
& \int_{\mathbb{R}^2}\int_{\mathbb{R}}\int_0^1 [\mathcal{SH}f](a,s,t)\langle \phi(x - \cdot), \psi_{a,s,t}\rangle a^{-3} dadsdt && [\rho_n \text{ is even}] \\
& \int_{\mathbb{R}^2}\int_{\mathbb{R}}\int_0^1 [\mathcal{SH}f](a,s,t)(\rho_n \star \psi_{a,s,t})(x) a^{-3} dadsdt.
\end{aligned}
$$

Therefore we can say that

$$f_n(x) = (f \star \rho_n)(x),$$

but since $\{\rho_n\}_{n=1}^{\infty}$ is an approximate identity we have

$$\lim_{n\to\infty} ||f_n - f||_{L^2} = \lim_{n\to\infty} ||(f \star \rho_n) - f||_{L^2} = 0,$$

which proves the theorem. □

In [6] they note that Theorem 3.4.5 implies the reconstruction formula given in Theorem 3.4.6 below.

Theorem 3.4.6

If $\psi \in L^2(\mathbb{R}^2)$ is an admissible shearlet with $C_\psi^+ = C_\psi^- = 1$ then for all $f \in L^2(\mathbb{R}^2)$ we have

$$f = \int_{\mathbb{R}^2}\int_{\mathbb{R}}\int_0^\infty [\mathcal{SH}f](a,s,t)\psi_{a,s,t} a^{-3} dadsdt, \tag{3.4}$$

where equality holds in the usual weak sense.

The reconstruction formula in (3.4) is concise and simple, however it has some problematic aspects. First it contains a term that is singular at $a = 0$ while integrating from $a = 0$. Second, both the a and s integrals run over infinite domains. This can cause problems with convergence at infinity. To fix the second problematic aspect we will need cone-adapted continuous shearlet systems. The rest of this chapter will be spent proving another reconstruction formula, Theorem 3.4.8. This reconstruction formula utilises cone-adapted shearlets and resolves the issue of integrating a and s over infinite domains.

We will be keeping the notation from Definition 3.3.1 and Definition 3.3.2.However, we also need to introduce a little more notation. First we define the frequency cones

$$\mathcal{C}_h := \{\xi \in \mathbb{R}^2 : |\xi_2|/|\xi_1| \leq 3/2\}$$

$$\mathcal{C}_v := \{\xi \in \mathbb{R}^2 : |\xi_1|/|\xi_2| \leq 2/3\}$$

where $\mathcal{C}_h$ corresponds to the horizontal frequency cone and $\mathcal{C}_v$ corresponds to the vertical frequency cone. We will also need the idea of directional vanishing moments.

Definition 3.4.7

A function f has N directional vanishing moments in the x_i-direction if

$$\frac{\hat{f}(\xi)}{\xi_i^N} \in L^2(\mathbb{R}^2).$$

We will assume that the shearlet ψ has infinitely many directional vanishing moments in all directions.

Next we state the main Theorem of this chapter, but we don't prove it just yet. Before we prove it we will have to prove a few preliminary results.

Theorem 3.4.8 (Taken from [7])

Let $\psi \in L^2(\mathbb{R}^2)$ be an admissible shearlet with $C_\psi = C_\psi^+ = C_\psi^-$, where these are the constants from Proposition 3.4.1. Then there exists a smooth function ϕ such that

$$C_\psi f = \int_{\mathbb{R}^2} \int_{-2}^{2} \int_0^1 a^{-3} [\mathcal{SH}_\psi f](a,s,t) \psi_{a,s,t} da ds dt + \int_{\mathbb{R}^2} \langle f, \phi(\cdot - t) \rangle \phi(\cdot - t) dt \qquad (3.5)$$

for any function f such that $\hat{f} \in L^2(\mathbb{R}^2)$ and $\operatorname{supp} \hat{f} \subset \mathcal{C}_h$.

Remark 3.4.9

The equality in (3.5) is to be understood in the same weak sense as in Chapter 2.2.1. We also have that exactly the same formula holds on the vertical frequency cone $\mathcal{C}_v$ if we replace $\mathcal{SH}_\psi$ with $\mathcal{SH}_{\tilde{\psi}}$ and ψ with $\tilde{\psi}$.

The first result that we need to build up to proving Theorem 3.4.8 is proposition 3.4.10.

Propostion 3.4.10 (Adapted from [7]**)**

Define the integral function

$$\Delta_\psi(\xi) := \int_{-2}^{2}\int_{0}^{1} |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds. \tag{3.6}$$

Under the hypotheses of Theorem 3.4.8 the reconstruction formula (3.5) holds if and only if

$$\Delta_\psi(\xi) + |\hat{\phi}(\xi)|^2 = C_\psi \quad \text{for all } \xi \in \mathcal{C}_h. \tag{3.7}$$

proof:[Proof adapted from [7]]

First we define

$$F(\phi) = \int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} a^{-3}[\mathcal{SH}_\psi f](a,s,t)\psi_{a,s,t} dadsdt + \int_{\mathbb{R}^2} \langle f, \phi(\cdot - t)\rangle \phi(\cdot - t)dt,$$

for any $f, \phi \in L^2(\mathbb{R}^2)$. Then for any $g \in L^2(\mathbb{R}^2)$ we have

$$\langle F(\phi), g\rangle = \int_{\mathbb{R}^2} \overline{g(x)} \left(\int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} a^{-3}[\mathcal{SH}_\psi f](a,s,t)\psi_{a,s,t}(x) dadsdt + \int_{\mathbb{R}^2} \langle f, \phi(\cdot - t)\rangle \phi(x - t)dt \right) dx \tag{3.8}$$

$$= \int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1}\int_{\mathbb{R}^2} a^{-3}\langle f, \psi_{a,s,t}\rangle \overline{g(x)}\psi_{a,s,t}(x) dxdadsdt + \int_{\mathbb{R}^2} \langle f, \phi(\cdot - t)\rangle \overline{g(x)}\phi(x - t)dxdt \tag{3.9}$$

$$= \int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} \langle f, \psi_{a,s,t}\rangle\langle \psi_{a,s,t}, g\rangle a^{-3} da\ ds\ dt + \int_{\mathbb{R}^2} \langle f, \phi(\cdot - t)\rangle\langle \phi(\cdot - t), g\rangle dt. \tag{3.10}$$

Now if we take the fourier transform in each of the L^2 inner products on both sides then we find that Plancherel's Theorem implies that

$$\langle \widehat{F(\phi)}, \hat{g}\rangle = \int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} \langle \hat{f}, \hat{\psi}_{a,s,t}\rangle\langle \hat{g}, \hat{\psi}_{a,s,t}\rangle a^{-3} da\ ds\ dt + \int_{\mathbb{R}^2} \langle \hat{f}, \mathcal{F}[\phi(\cdot - t)]\rangle\langle \hat{g}, \mathcal{F}[\phi(\cdot - t)]\rangle dt$$

Using the definition of the L^2 inner-product and the translation formula from Theorem 1.1.3.2 we find that

$$\langle \widehat{F(\phi)}, \hat{g}\rangle = T_1 + T_2,$$

where

$$T_1 = \int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} a^{-3}\left[\int_{\mathbb{R}^2} \hat{f}(\xi)\overline{a^{-3/4}e^{-i\langle t,\xi\rangle}\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1))}d\xi\right] \times \left[\int_{\mathbb{R}^2} \overline{\hat{g}(\eta)} a^{-3/4}e^{-i\langle t,\eta\rangle}\hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1))d\eta\right] dadsdt$$

$$= \int_{-2}^{2}\int_{0}^{1}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} a^{-3}\left[\hat{f}(\xi)\overline{a^{-3/4}e^{-i\langle t,\xi\rangle}\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1))}\right]$$
$$\times\left[\overline{\hat{g}(\eta)}a^{-3/4}e^{-i\langle t,\eta\rangle}\hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1))\right] d\eta d\xi dt da dt$$
$$= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1}\left[a^{-3/2}e^{-i\langle n-\xi,t\rangle}\hat{f}(\xi)\hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1)\overline{\hat{g}(\eta)\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)}\right] d\eta d\xi dt da ds$$
$$= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_{0}^{1} e^{-i\langle n-\xi,t\rangle}\hat{f}(\xi)\overline{\hat{g}(\eta)}\left[a^{-3/2}\hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1)\overline{\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)}\right] d\eta d\xi dt da ds.$$

The term T_2 is given by

$$T_2 = \int_{\mathbb{R}^2}\left[\int_{\mathbb{R}^2}\hat{f}(\xi)\overline{e^{-i\langle t,\xi\rangle}\hat{\phi}(\xi)}d\xi\right]\left[\int_{\mathbb{R}^2}\overline{\hat{g}(\eta)}e^{-i\langle t,\eta\rangle}\hat{\phi}(\eta)d\eta\right] dt$$
$$= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\left[\hat{f}(\xi)\overline{e^{-i\langle t,\xi\rangle}\hat{\phi}(\xi)}\right]\left[\overline{\hat{g}(\eta)}e^{-i\langle t,\eta\rangle}\hat{\phi}(\eta)\right] d\xi d\eta dt$$
$$= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{-i\langle n-\xi,t\rangle}\hat{f}(\xi)\overline{\hat{g}(\eta)}\left[\overline{\hat{\phi}(\xi)}\hat{\phi}(\eta)\right] d\xi d\eta dt.$$

We combine T_1 and T_2 to find

$$\langle\widehat{F(\phi)},\hat{g}\rangle = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{-i\langle n-\xi,t\rangle}\hat{f}(\xi)\overline{\hat{g}(\eta)}I(\xi,\eta)d\xi d\eta dt \tag{3.11}$$
$$= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\delta_{\eta-\xi}\hat{f}(\xi)\overline{\hat{g}(\eta)}I(\xi,\eta)d\xi d\eta dt \tag{3.12}$$
$$= \int_{\mathbb{R}^2}\hat{f}(\xi)\overline{\hat{g}(\xi)}I(\xi,\xi)d\xi, \tag{3.13}$$

with

$$I(\xi,\eta) = \int_{-2}^{2}\int_{0}^{1}\overline{\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)}\hat{\psi}(a\eta_1, a^{1/2}(\eta_2 - s\eta_1))a^{-3/2}da ds + \overline{\hat{\phi}(\xi)}\hat{\phi}(\eta).$$

Then if Theorem 3.4.8 holds we have that Parseval's theorem (Theorem 1.1.9) implies that

$$C_\psi\langle\hat{f},\hat{g}\rangle = \langle\widehat{F(\phi)},\hat{g}\rangle.$$

It then follows from (3.13) that

$$0 = \int_{\mathbb{R}^2}(C_\psi - I(\xi,\xi))\hat{f}(\xi)\hat{g}(\xi)d\xi,$$

for all $g \in L^2(\mathbb{R}^2)$. From this we can conclude that

$$||C_\psi - I(\xi,\xi)||_{L^2} = 0,$$

which means in in an $L^2(\mathbb{R}^2)$ sense we have

$$C_\psi = \int_{-2}^{2}\int_{0}^{1}\left|\hat{\psi}(a\xi_1, a^{1/2}(\xi_2 - s\xi_1))\right|^2 a^{-3/2}da ds + \left|\hat{\phi}(\xi)\right|^2 \tag{3.14}$$

for all ξ. Conversely if (3.14) holds then we have that $I(\xi,\xi) = C_\psi$ for all ξ which means that

$$\begin{aligned}\langle \widehat{F(\phi)}, \hat{g}\rangle &= \int_{\mathbb{R}^2} \hat{f}(\xi)\overline{\hat{g}(\xi)} I(\xi,\xi) d\xi \\ &= C_\psi \int_{\mathbb{R}^2} \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi \\ &= C_\psi \langle \hat{f}, \hat{g}\rangle.\end{aligned}$$

Now applying Parseval's theorem we have

$$\langle F(\phi), g\rangle = C_\psi \langle f, g\rangle$$

for all $g \in L^2(\mathbb{R}^2)$. Thus we conclude that $F(\phi) = C_\psi f$ in the weak sense of Definition 2.2.3. and Theorem 3.4.8 holds. This means that we can conclude that Theorem 3.4.8 and (3.7) are equivalent. □

Proposition 3.4.10 is quite powerful since it implies that all we need to do to prove Theorem 3.4.8 is show that any function ϕ defined by (3.7) is smooth. To show that such a ϕ is smooth all we need to do is show that $\hat{\phi}$ has decay faster than any polynomial. In particular we will show that for all $N \in \mathbb{N}$

$$|\hat{\phi}(\xi)|^2 = O(||\xi||^{-N}), \quad \text{for } \xi \in \mathcal{C}_h, \xi \to \infty$$

This implies that ϕ is smooth since it means that $|\phi(\xi)|$ also satisfies this type of decay, which means that ϕ is smooth. Before we can show that ϕ has the needed fourier decay we need to prove Proposition 3.4.11 which tells us more about the behaviour of the function Δ_ψ.

Propostion 3.4.11 (Taken from [7]**)**

$$C_\psi = \int_{\mathbb{R}} \int_0^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds. \tag{3.15}$$

proof:[Expanded proof from [7]]

Consider the coordinate transform

$$\eta(a,s) = (\eta_1(a,s), \eta_2(a,s)) = \left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)$$

the Jacobian of this transform is then

$$J(a,s) = \begin{pmatrix} \xi_1 & 0 \\ \frac{1}{2}a^{-1/2}((\xi_2 - s\xi_1)) & -a^{1/2}\xi_1 \end{pmatrix}$$

with $|\det J(a,s)| = a^{1/2}\xi_1 = (\eta_1/a)^2 = a^{-3/2}\eta_1^2$. Then we find that

$$\begin{aligned}\int_{\mathbb{R}}\int_0^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds &= \int_{\mathbb{R}}\int_0^\infty |\hat{\psi}\left(\eta_1(a,s), \eta_2(a,s)\right)|^2 |\det J(a,s)| \eta_1^{-2} da\ ds \\ &= \int_{\mathbb{R}}\int_0^\infty |\hat{\psi}(\eta_1,\eta_2)|^2 \eta_1^{-2} d\eta_1 d\eta_2 \\ &= C_\psi\end{aligned}$$

□

Now, using Propostion 3.4.11 we can prove that ϕ has sufficient Fourier decay for it to be smooth. Once we have proven Proposition 3.4.12 below then ϕ is smooth which, as noted earlier, will prove Theorem 3.4.8.

Propostion 3.4.12 (Adapted from [7]**)**

Suppose a function ϕ satisfies

$$\Delta_\psi(\xi) + |\hat{\phi}(\xi)|^2 = C_\psi$$

for all $\xi \in \mathcal{C}_h$. Then we have that

$$|\hat{\phi}(\xi)|^2 = O(||\xi||^{-N}),$$

for all $N \in \mathbb{N}$ and $\xi \in \mathcal{C}_h$.

proof: Using (3.15) and 3.7 we dind that

$$\begin{aligned}|\hat{\phi}(\xi)|^2 &= \int_{\mathbb{R}}\int_0^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds \\ &\quad - \int_{-2}^{2}\int_0^1 |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds \\ &= \int_{|s|>2}\int_0^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds \qquad (3.16) \\ &\quad + \int_{-2}^{2}\int_1^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds. \qquad (3.17)\end{aligned}$$

Next, consider each of these integrals separately, starting with (3.17). The fact that ψ is smooth means that

$$|\hat{\psi}(\xi)| = O(||\xi||^{-N}),$$

for all $N \in \mathbb{N}$ and $\xi \in \mathcal{C}_h$.

Now since s only varies over a compact set and a is large, up to irrelevant constants we can estimate

$$\begin{aligned}\int_{-2}^{2}\int_{1}^{\infty} |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds &\lesssim \int (a|\xi_1|)^{-N} a^{-3/2} da \\ &\lesssim |\xi_1|^{-N} \\ &\lesssim ||\xi||^{-N}.\end{aligned}$$

The last estimate follows from the fact that $\xi \in \mathcal{C}_h$ and therefore $|\xi_2|/|\xi_1| \leq 3/2$ which means that we can always estimate $||\xi_1||^{-1}$ from above by $||\xi||^{-1}$. As mentioned before, in the working above we ignore constants since we are only interested in asymptotic behaviour at the moment.

Finally we consider (3.16). We start by breaking (3.16) into two integrals,first we integrate over $a > 1$ and second we integrate over $a < 1$. Now since both a and s are large we have, up to irrelevant constants;

$$\begin{aligned}\int_{|s|>2}\int_{1}^{\infty} |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds &\lesssim \int_{|s|>2}\int_{1}^{\infty} a^{-N}|\xi_2 - s\xi|^{-2N} a^{-3/2} da\ ds \\ &= \int_{|s|>2}\int_{1}^{\infty} a^{-N}|\xi_1|^{-N}\left|\frac{\xi_2}{\xi_1} - s\right|^{-2N} a^{-3/2} da\ ds \\ &\leq \int_{|s|>2}\int_{1}^{\infty} a^{-N}|\xi_1|^{-N}\left|\frac{3}{2} - |s|\right|^{-2N} a^{-3/2} da\ ds \\ &\lesssim ||\xi||^{-N}.\end{aligned}$$

In the next estimate we will finally need the fact that we assumed that ψ has infinitely many directional vanishing moments. This means that in the second coordinate ψ is smooth. So

$$\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right) \leq a^{1/2}(\xi_2 - s\xi_1)^{-L}. \tag{3.18}$$

Since ψ has infinitely directional vanishing moments in all directions we have that

$$\frac{\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)}{(a\xi_1)^N} \in L^2(\mathbb{R}^2)$$

and therefore we have that

$$\frac{\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)}{(a\xi_1)^M} \leq (a\xi_1)^{-1/2}.$$

This means that

$$\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right) \leq (a\xi_1)^{M-1/2}. \tag{3.19}$$

Now taking the goemetric mean of (3.18) and (3.19) we can compute

$$\begin{aligned}
\int_{|s|<2}\int_0^1 |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds &\lesssim \int_{|s|<2}\int_0^1 a^M|\xi_1|^M a^{-L}|\xi_2 - s\xi_1|^{-2L} a^{-3/2} da\ ds \\
&= \int_{|s|<2}\int_0^1 a^M|\xi_1|^{M-2L} a^{-L}\left|\frac{\xi_2}{\xi_1} - s\right|^{-2L} a^{-3/2} da\ ds \\
&\leq \int_{|s|<2}\int_0^1 a^M|\xi_1|^{M-2L} a^{-L}\left|\frac{3}{2} - |s|\right|^{-2L} a^{-3/2} da\ ds.
\end{aligned}$$

The last line holds for all $L, M \in \mathbb{N}$, in particular for $L = N + 2$ and $M = L + 4$.

Substituting in this L and M we find

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

Now if we sum up all three estimate we find that

$$\begin{aligned}
|\hat{\phi}(\xi)|^2 &= \int_{|s|>2}\int_0^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds \\
&+ \int_{-2}^2\int_1^\infty |\hat{\psi}\left(a\xi_1, a^{1/2}(\xi_2 - s\xi_1)\right)|^2 a^{-3/2} da\ ds \\
&\lesssim ||\xi||^{-N}
\end{aligned}$$

and therefore

$$|\hat{\phi}(\xi)|^2 = O(||\xi||^{-N})$$

as we claimed. □

Since we were able to prove Proposition 3.4.12 we now know not only that ϕ is smooth but also that Theorem 3.4.8 also holds.

We now have two shearlet reconstruction formulas that hold in the usual weak sense. So we now have two ways of involving the shearlet transform in the calculation of L^2 inner products. However the two representations are not equally useful or reliable. Of the two reproduction formulas that we proved (3.5) from Theorem 3.4.8 is the most reliable reproduction, but it isn't very simple. Theorem 3.4.8 cannot be directly applied to a function $f \in L^2$ that isn't concentrated in the horizontal frequency cone. We would need to split f into two parts, f_h with $\hat{f}_h$ concentrated in $\mathcal{C}_h$ and f_v with $\hat{f}_v$ concentrated in $\mathcal{C}_v$. Then reconstruct f_h using (3.5) and f_v using the analogous formula in $\mathcal{C}_v$. This method is more reliable in the sense that it is more likely to preserve convergence of integrals since a, s vary over compact domains. However the fact that we need to project onto frequency cones can make the reconstruction become complicated and extremely involved. Since we then view f as the superposition of two functions we can't directly use the shearlet coefficients of f, which might be simple, but we need to use the shearlet coefficients of f_h and f_v with respect to potentially different shearlets. This means that we have to do a lot more calculations involving potentially unruly functions.

On the other hand (3.4) is a lot simpler and it allows us to reconstruct all of f using a single shearlet. However (3.4) has one major draw back, a and s vary over infinite domains. When using this representation in practice, the integrals are rarely convergent due to the fact that a and s vary over infinite domains.

While both formulas (3.4) and (3.5) have their own strengths and weaknesses, they do share one common weakness. Both involve integrating against a^{-3}. Since both formulas allow a to go to zero it can be difficult to find convergence of both formulas. In general we cannot expect all functions f to have $[\mathcal{SH}f](a, s, t)$ decaying sufficiently fast as a approaches zero. So both representation formulas can present problems when it comes to convergence of the integrals they involve. Thus we should be careful when using these formulas when trying to derive estimates for L^2 inner products.

Chapter 4

L^p norms of Convolution Operators

In this chapter we will derive and L^2 to L^p bound for a certain family of convolution operators for any $p \in [2, \infty]$. Our main strategy to prove this will be to apply the shearlet reconstruction formulas from Chapter 3.

Definition 4.0.1

Let $m \in L^2(\mathbb{R}^2)$ then we define the linear operator T_m by

$$T_m f = \mathcal{F}^{-1}[m\hat{f}], \qquad f \in L^2(\mathbb{R}^2).$$

By the Fourier inversion theorem we can write this as

$$T_m f = (2\pi)^{-1} \hat{m}(-\cdot) \star f.$$

Therefore we call this this operator a Convolution operator.

If $m \in L^\infty(\mathbb{R}^2)$ it immediately follows from Plancherel's Theorem that T_m is a bounded linear operator on $L^2(\mathbb{R}^2)$.

Propostion 4.0.2

Let $m \in L^\infty(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$ then we have

$$||T_m f||_{L^2} \leq (2\pi)^{-1} ||m||_\infty|| \ ||f||_{L^2} \quad \text{for all } f \in L^2(\mathbb{R}^2).$$

This means that T_m is a bounded linear operator on $L^2(\mathbb{R}^2)$.

Hölder's inequality also quickly implies a result similar to Proposition 4.0.2.

Propostion 4.0.3

Let $m \in L^2(\mathbb{R}^2)$ then we have

$$||T_m f||_\infty \leq (2\pi)^{-1}||m||_{L^2}||f||_{L^2} \quad \textit{for all } f \in L^2(\mathbb{R}^2).$$

This means that T_m is a bounded linear operator from $L^2(\mathbb{R}^2)$ to $L^\infty(\mathbb{R}^2)$.

proof: Pick $m, f \in L^2(\mathbb{R}^2)$, then we compute for any x

$$\begin{aligned}
|T_m f(x)| &= |\mathcal{F}^{-1}[m\hat{f}](x)| \\
&= \left|(2\pi)^{-1}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}m(\xi)\hat{f}(\xi)d\xi\right| \\
&\leq (2\pi)^{-1}||m||_{L^2}||\hat{f}||_{L^2} \quad \text{[Hölder's inequality.]} \\
&= (2\pi)^{-1}||m||_{L^2}||f||_{L^2} \quad \text{[Plancherel's Theorem.]}
\end{aligned}$$

□

We can also show a similar result for a smaller subspace of $L^2(\mathbb{R}^2)$

Propostion 4.0.4

Suppose that $m \in L^2(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ then we for any $f \in L^2(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ have that

$$||T_m f||_\infty \leq (2\pi)^{-1}||m||_{L^1}||f||_{L^1}$$

Thus T_m is a bounded linear operator from $L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ to $L^\infty(\mathbb{R}^2)$, with respect to the L^1 and L^∞ norms respectively.

proof: This result follows directly from computation. Since $m \in L^1(\mathbb{R}^2)$ we have that

$$|\hat{m}(x)| \leq (2\pi)^{-1}||m||_{L^1}, \quad \text{for all } \xi \in \mathbb{R}^2.$$

For any $f \in L^2(\mathbb{R}^2)$ we have;

$$\begin{aligned}
|T_m f(x)| = (2\pi)^{-1}&\left|\int_{\mathbb{R}^2} \hat{m}(t-x)f(x)dx\right| \\
&\leq (2\pi)^{-1}||m||_{L^1}\int_{\mathbb{R}^2} |f(x)dx| \\
&\leq (2\pi)^{-1}||m||_{L^1}||f||_{L^1}.
\end{aligned}$$

□

In order to prove more about the L^p norms of convolution operators we will need to restrict the class of multipliers. First we define the exact multiplier m that we will be using. We consider a multiplier that tends towards an L^2 normalised version of the delta distribution along the unit circle. We would like to be able to show that the delta distribution defines a convolution operator that is a bounded linear operator from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ for any $p \in [1, \infty]$. However, doing so directly is not so easy, but we can sneak up on it by approaching with smooth functions concentrated on the unit circle.

Definition 4.0.5

Let $\chi : \mathbb{R} \to \mathbb{R}$ be a bounded, smooth function with supp $\psi = [-1, 1]$, such that $||\chi||_{L^2} = 1$. Then we define the smooth 2-dimensional function

$$m_R : \mathbb{R}^n \to \mathbb{R}$$
$$m_R(\xi) := R^{1/2}\chi(R||\xi|| - R).$$

From now on we will also denote the convolution operator associated with m_R by T_{m_R}.

Remark 4.0.6

The way we defined m_R means that

$$\mathcal{M} := \text{supp } \ m_R = \{\xi \in \mathbb{R}^2 : 1 - R^{-1} \leq ||x|| \leq 1 + R^{-1}\}.$$

So m_R is a smooth, bounded, compactly supported function, that is m_R is a Schwartz function. This means that $\widehat{m_R}$ is rapidly decaying, in particular

$$|\widehat{m_R}(x)| = O_R(||x||^{-N}) \quad \text{as } x \to \infty$$

for all $x \in \mathbb{R}^2$ and $N \in \mathbb{N}$. This follows from the derivative properties of the Fourier transform in Proposition 1.1.8.

Our goal is to apply Theorem 3.4.8 to T_{m_R} but recall that (3.5) only holds in the weak sense of Definition 2.2.3. So we need to reformulate the definition of T_{m_R} as a. By substituting in the explicit definition of the inverse Fourier Transform for any $f \in L^2(\mathbb{R}^2)$ we can rewrite $T_{m_R}f$ as

$$T_{m_R}(x) = \mathcal{F}^{-1}[m_R\hat{f}](x) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle} m_R(\xi)\hat{f}(\xi)d\xi. \tag{4.1}$$

Now we can view T_{m_R} as an L^2 inner product, which is what we need if we want to reconstruct m_R using the cone-adapted shearlet reconstruction formula from Theorem

3.4.8. However, before we do this we need to break m_R up into it's horizontal and vertical frequency parts.

Definition 4.0.7

Let Ξ_h and Ξ_v be smooth functions such that $\Xi_h(x) = 1$ for all $x \in \mathcal{C}_h$ and $\Xi_v(x) = 1$ for all $x \in \mathcal{C}_v$. Also let Ξ_h and Ξ_v have Schwartz like decay outside of $\mathcal{C}_h$ and $\mathcal{C}_v$ respectively. Lastly we require that

$$\Xi_h(x) + \Xi_v(x) = 1$$

for all $x \in \mathbb{R}^2$.

We can view Ξ_h and Ξ_v as smooth indicator functions for the horizontal and vertical frequency cones, respectively. Then we can split f into horizontal and vertical frequency parts. For any $g \in L^1(\mathbb{R}^2)$ we have that

$$\mathcal{F}^{-1}[g](\xi) = \mathcal{F}[g](-\xi),$$

then

$$\begin{aligned}\hat{f}(\xi) &= \mathcal{F}^{-1}[f](-\xi)\\ &= \mathcal{F}^{-1}[f(\Xi_h + \Xi_v)](-\xi)\\ &= \mathcal{F}^{-1}[f\Xi_h](-\xi) + \mathcal{F}^{-1}[f\Xi_v](-\xi)\\ &= \hat{f}_h(\xi) + \hat{f}_v(\xi),\end{aligned}$$

where

$$\hat{f}_h(\xi) = \mathcal{F}^{-1}[f\Xi_h](-\xi) \quad \text{and} \quad \hat{f}_v(\xi) = \mathcal{F}^{-1}[f\Xi_v](-\xi).$$

Then we have that

$$T_{m_R}(x) = \mathcal{F}^{-1}[m_R\hat{f}](x) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle} m_R(\xi)\hat{f}_h(\xi)d\xi + \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle} m_R(\xi)\hat{f}_h(\xi)d\xi. \quad (4.2)$$

We will use Theorem 3.4.8 to reconstruct m_R. Recall that the reconstruction formula (3.5) holds only for functions with frequency support contained in $\mathcal{C}_h$. The function m_R does not have frequency support contained in $\mathcal{C}_h$. This isn't a problem, since we are multiplying m_R by $\hat{f}_h$, which does have a fourier transform concentrated in $\mathcal{C}_h$. The parts of the reconstruction that come from outside of $\mathcal{C}_h$ will be dominated by the decay of $\hat{f}_h$. The same is true in $\mathcal{C}_v$. So we can safely apply (3.5) to m_R. First we need to pick a shearlet with which we define our transform.

Definition 4.0.8

Let $\psi \in L^2(\mathbb{R}^2)$ be our shearlet function then we assume ψ satisfies the following;

1. *We assume that ψ has compact support and we define*

$$U_\psi := \operatorname{supp} \psi.$$

Furthermore we assume that $U_\psi = \overline{S_1(0)} := \{(x, y) \in \mathbb{R}^2 : |x|, |y| \leq 1\}$.

2. *We assume that ψ is smooth on U_ψ and bounded with $||\psi||_\infty = 1$.*

3. *We assume that $\psi(x, y) = \psi_1(x)\psi_2(y)$ with*

$$\int_{-\infty}^{\infty} \psi_1(x)dx = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \psi_2(y)dy = 0. \tag{4.3}$$

4. *We also assume that ψ satisfies the admissibility condition*

$$\int_{\mathbb{R}^2} |\xi|^{-1}|\hat{\psi}(\xi)|d\xi < \infty.$$

5. *We also assume that there exists smooth, bounded and compactly supported functions Θ_1, Θ_2 on $\mathbb{R}$ such that*

$$\psi_1(x) = \frac{d^n}{dx^n}\Theta_1(x) \quad \text{and } \psi_2(x) = \frac{d^n}{dy^n}\Theta_2(x), \tag{4.4}$$

for some $n \in \mathbb{N}$. A specific n will be chosen later.

Remark 4.0.9

Definitions such as Definition 4.0.8 above are obviously only useful when they don't define an empty set. Luckily shearlet's of this form are easy to find. We can simply take ψ_1 and ψ_2 to be any derivative of the Gaussian Bump function.

Theorem 4.0.10

Let

$$C_\psi = \int_0^\infty \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1 = \int_{-\infty}^0 \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1$$

and define the integral functions

$$\Psi_m^h f(x) := \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{-2}^{2} \int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a, s, t)\psi_{a,s,t}(\xi)\hat{f}_h(\xi)e^{i\langle x, \xi\rangle} dadsdtd\xi,$$

$$\Phi_m^h f(x) := \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{i\langle x, \xi\rangle} \hat{f}_h(\xi)\langle m_R, \phi(\cdot - t)\rangle\phi(\xi - t)dtd\xi.$$

Here ϕ is the smooth function given by Proposition 3.4.10. Then there exist functions $T^h_{m_R}f$ and $T^v_{m_R}f$ with $\widehat{T^h_{m_R}f}$ and $\widehat{T^v_{m_R}f}$ rapidly decaying outside of $\mathcal{C}_h$ and $\mathcal{C}_v$ respectively, such that

$$T_{m_R}f(x) = T^h_{m_R}f(x) + T^v_{m_R}f(x)$$

for all $x \in \mathbb{R}^2$ and

$$2\pi C_\psi T^h_{m_R}f(x) = \Psi^h_m f(x) + \Phi^h_m f(x).$$

The function $T^v_{m_R}f$ can be reconstructed in an analogous way using the vertical frequency cone.

proof: Recall from (4.2) that

$$T_{m_R}f(x) = \underbrace{\frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle} m_R(\xi)\hat{f}_h(\xi)d\xi}_{=:T^h_{m_R}f(x)} + \underbrace{\frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle} m_R(\xi)\hat{f}_h(\xi)d\xi}_{=:T^v_{m_R}f(x)}.$$

The function $T^h_{m_R}f$ corresponds to the horizontal frequency part of $T_{m_R}f$ and $T^v_{m_R}f$ to the vertical frequency part of $T_{m_R}f$. In the rest of our argument we will focus on $T^h_{m_R}f$, but the entire argument will also hold for $T^v_{m_R}f$ if we redo the working in the vertical cone. Using (3.5) we find.

$$C_\psi m_R(\xi) = \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\psi_{a,s,t}(\xi)dadsdt + \int_{\mathbb{R}^2}\langle m_R, \phi(\cdot - t)\rangle\phi(\xi - t)dt.$$

It then follows that

$$\begin{aligned}
2\pi C_\psi T^h_{m_R}f(x) =& \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\psi_{a,s,t}(\xi)\hat{f}_h(\xi)e^{i\langle x,\xi\rangle}dadsdtd\xi \\
&+ \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}\hat{f}_h(\xi)\langle m_R, \phi(\cdot - t)\rangle\phi(\xi - t)dtd\xi \\
=&\Psi^h_m f(x) + \Phi^h_m f(x),
\end{aligned}$$

where $\Psi^h_m f$ is the shearlet part of $T^h_{m_R}f$ given by

$$\Psi^h_m f(x) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\psi_{a,s,t}(\xi)\hat{f}_h(\xi)e^{i\langle x,\xi\rangle}dadsdtd\xi$$

and $\Phi^h_m f$ is the auxillary part given by

$$\Phi^h_m f(x) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}\hat{f}_h(\xi)\langle m_R, \phi(\cdot - t)\rangle\phi(\xi - t)dtd\xi.$$

$\square$

Remark 4.0.11

Since $L^p(\mathbb{R}^2)$ is a vector space, if we can show that $T^h_{m_R}f,\ T^v_{m_R}f \in L^p(\mathbb{R}^2)$, then we have that $T_{m_R}f \in L^p(\mathbb{R}^2)$. Similarly if we can show that $\Psi^h_m f,\ \Phi^h_m f \in L^p(\mathbb{R}^2)$ then we have that $T^h_{m_R}f \in L^p(\mathbb{R}^2)$. The same arguments work for $T^v_{m_R}f$ and its shearlet and auxiliary parts. So in the rest of our working we will only be considering $T^h_{m_R}f$, since the conclusions we make for it will carry over to $T^v_{m_R}f$ by repeating our arguments in the vertical frequency cone.

We start by considering the auxiliary part of $T^h_{m_R}f$. It can be shown easily that $\Phi^h_m f$ is indeed contained in $L^p(\mathbb{R}^2)$ for all $p \in [2,\infty)$. This is done in Proposition 4.0.12 below.

Propostion 4.0.12

For all $p \in [2,\infty)$ we have that $\Phi^h_m f \in L^p(\mathbb{R}^2)$ and

$$||\Phi^h_m f||_{L^p} \leq CR^{-1/2}||f||_{L^2}$$

for some C depending on p.

proof: First we notice that

$$\begin{aligned}\langle m_R, \phi(\cdot - t)\rangle &= \int_{\mathbb{R}^2} m_R(y)\phi(y-t)dy \\ &= \int_{\mathbb{R}^2} m_R(y)\phi(-(t-y))dy \\ &= (m_R \star \phi^*)(t),\end{aligned}$$

where $\phi^*(y) = \phi(-y)$. Now we have

$$\begin{aligned}\Phi^h_m f(x) &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}\hat{f}_h(\xi)(m_R \star \phi^*)(t)\phi(\xi - t)dtd\xi \\ &= \int_{\mathbb{R}^2}(m_R \star \phi^*)(t)\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}\hat{f}_h(\xi)\phi(\xi-t)d\xi dt \quad \text{[Interchange integrals by Fubini]} \\ &= \int_{\mathbb{R}^2}(m_R \star \phi^*)(t)([f\Xi_h]^* \star \mathcal{F}^{-1}[\phi(\cdot - t)])(x)dt \\ &= (2\pi)^{-1}([f\Xi_h]^* \star \mathcal{F}^{-1}[\phi])(x)\int_{\mathbb{R}^2} e^{i\langle x,t\rangle}(m_R \star \phi^*)(t)dt \quad \text{[By Theorem 1.1.3.3]} \\ &= ([f\Xi_h]^* \star \mathcal{F}^{-1}[\phi])(x)\,\mathcal{F}^{-1}[m_R](x)\,\mathcal{F}^{-1}[\phi^*](x),\end{aligned}$$

where

$$[f\Xi_h]^*(x) = [f\Xi_h](-x).$$

Recall from Proposition 3.4.10 that

$$\int_{-2}^{2}\int_{0}^{1} |\hat{\psi}\left(ax_1, a^{1/2}(x_2 - sx_1)\right)|^2 a^{-3/2} da\ ds + |\hat{\phi}(x)|^2 = C_\psi < \infty.$$

Therefore $\mathcal{F}^{-1}\,\phi^*$ is bounded. By definition we know that $m_R \in L^1(\mathbb{R}^2)$ and therefore

$$||\,\mathcal{F}^{-1}[m_R]||_\infty \leq ||m_R||_{L^1} = R^{-1/2}||\chi(||\cdot||)||_{L^1}.$$

We also know that $||\Xi_h||_\infty = 1$ so $||[f\Xi_h]^*||_{L^2} \leq ||f||_{L^2}$. Now we have that

$$\begin{aligned}||\Phi_m^h f||_{L^p} &\leq R^{-1/2}||\chi(||\cdot||)||_{L^1}||\hat{\phi}||_\infty||([f\Xi_h] \star \mathcal{F}^{-1}[\phi])||_{L^p}\\ &\leq R^{-1/2}||\chi(||\cdot||)||_{L^1}||\hat{\phi}||_\infty||f\Xi_h||_{L^2}||\hat{\phi}||_{L^r}\\ &\leq R^{-1/2}||\chi(||\cdot||)||_{L^1}||\hat{\phi}||_\infty||f||_{L^2}||\hat{\phi}||_{L^r}.\end{aligned}$$

where the second last line follows from Young's convolution inequality if

$$r = \frac{2p}{2+p}.$$

However we need $r \geq 1$ which holds if $p \geq 2$. Since $\hat{\phi}$ is bounded and rapidly decaying $\hat{\phi} \in L^r(\mathbb{R}^2)$.Thus if $p \geq 2$ we have

$$||\Phi_m^h f||_{L^p} \leq CR^{-1/2}||f||_{L^2} < \infty.$$

where

$$C = ||\chi(||\cdot||)||_{L^1}||\hat{\phi}||_\infty||\hat{\phi}||_{L^r}$$

Therefore $\Phi_m^h f \in L^p(\mathbb{R}^2)$ for any $p \in [2, \infty)$. □

The proof of Proposition 4.0.12 can be repeated exactly the same in the veritical cone to show that $\Phi_m^v f \in L^p(\mathbb{R}^2)$ for all $p \in [2, \infty]$, with the same L^p bound. Thus the auxiliary parts of $T_{m_R} f$ are in any L^p space that we want as long as $p \geq 2$. Now all that we still need to do is find bounds on the L^p norms of the Shearlet parts. This will take a bit more work to accomplish and we also expect them to depend on R. First we find a preliminary bound for $||\Psi_m^h f||_{L^p}$.

Propostion 4.0.13

For any $f \in L^2(\mathbb{R}^2)$ and $p \geq 2$ we have

$$||\Psi_m^h f||_{L^p} \leq (2\pi)^{-1}||f||_{L^2} \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}|[\mathcal{SH}_\psi m_R](a,s,t)| \ \left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} dadsdt, \qquad (4.5)$$

where

$$r = \frac{2p}{2+p}.$$

proof: We start by computing;

$$\begin{aligned}
\Psi_m^h f(x) &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\psi_{a,s,t}(\xi)\hat{f}_h(\xi)e^{i\langle x,\xi\rangle}dadsdtd\xi \\
&= \int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\int_{\mathbb{R}^2}\psi_{a,s,t}(\xi)\hat{f}_h(\xi)e^{i\langle x,\xi\rangle}d\xi dadsdt \quad \text{[Interchange by Fubini]} \\
&= (2\pi)^{-1}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\left(\widehat{\psi_{a,s,t}}^* \star [f\Xi_h]\right)(x)\ dadsdt,
\end{aligned}$$

where $\widehat{\psi_{a,s,t}}^*(x) = \widehat{\psi_{a,s,t}}(x)$. We can then estimate the L^p norm of $\Psi_m^h f$ by;

$$\begin{aligned}
||\Psi_m^h f||_{L^p} &= (2\pi)^{-1}\left(\int_{\mathbb{R}^2}\left|\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}[\mathcal{SH}_\psi m_R](a,s,t)\left(\widehat{\psi_{a,s,t}}^* \star [f\Xi_h]\right)(x)dadsdt\right|^p dx\right)^{1/p} \\
&\leq (2\pi)^{-1}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}\left|[\mathcal{SH}_\psi m_R](a,s,t)\right|\left(\int_{\mathbb{R}^2}\left|(\widehat{\psi_{a,s,t}}^* \star [f\Xi_h])(x)\right|^p dx\right)^{1/p} dadsdt \\
&\leq (2\pi)^{-1}||f||_{L^2}\int_{\mathbb{R}^2}\int_{-2}^{2}\int_0^1 a^{-3}|[\mathcal{SH}_\psi m_R](a,s,t)| \ \left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} dadsdt,
\end{aligned}$$

where the first step followed from Minkowski's inequality and the last line follows from Young's convolution inequality if

$$r = \frac{2p}{p+2}.$$

and the fact that $||f\Xi_h||_{L^2} \leq ||f||_{L^2}$. Since we want $r \geq 1$ we choose $p \geq 2$. □

Before we can use (4.5) to find a bound for the L^p norm of Ψ_m^h we need to look at the shearlet coefficients a little closer. Proposition 4.0.14 below has less to do with shearlet transforms and more with coordinate geometry and measure theory.

Propostion 4.0.14

Let $U_\psi = S_1(0) = \{(x,y) \in \mathbb{R}^2 : ||x|,|y| \leq 1\}$ and fix shearlet parameters a, s and t we define the values

$$l_1 = a^{1/2}\sqrt{s^2+1}, \quad l_2 = \frac{a}{\sqrt{s^2+1}} \quad \textit{and} \quad l = \sqrt{(a+a^{1/2}s)^2+a}$$

Then the following statements hold;

1. *The region $U_{a,s,t} := [S_s A_a(U_\psi) + t]$ is a parallelogram with semi-major axis length l_1, semi-minor axis length l_2 and diameter l.*

2. $\lambda(U_{a,s,t}) = 2a^{3/2}$.

proof:

1. The fact that U_ψ is square with side-length 2 immediately implies that $A_a U_\psi$ is a rectangle height $a^{1/2}$ and width a. Now shearing $A_a U_\psi$ by applying S_s we turn $A_a U_\psi$ into a parallelogram, then shifting it by t simply changes the centre. In the calculations that follow we will assume that $t = 0$. The major-axis L_1 of this parallelogram is parallel to the image of $(0,1)$. The semi major-axis is the line from this image and the origin. Then the semi-major axis length, l_1 is then given by

$$l_1 = \left\| S_s A_a \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{1/2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} a^{1/2}s \\ a^{1/2} \end{pmatrix} \right\| = a^{1/2}\sqrt{s^2+1}.$$

Now the minor-axis, L_2 is the line perpendicular to L_1. The semi-minor axis is the line from the origin to the edge of $U_{a,s,0}$. To find the length , l_2 of this line we need to do some more work. We start by finding the angle, θ that L_2 makes with the x-axis. The slope of L_1 is s^{-1} so the slope of L_2 is then $-s$ which means that the angle between L_2 and the x-axis is given by

$$\theta = \arctan(-s).$$

Then we find that

$$l_2 = a\cos(\arctan(-s)) = \frac{1}{\sqrt{s^2+1}}.$$

We can also compute the diameter l of $U_{a,s,t}$ by computing the length of $S_s A_s(1,1)^T$. Then we find that

$$l = \left\| S_s A_a \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{1/2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} a + a^{1/2}s \\ a^{1/2} \end{pmatrix} \right\| = \sqrt{(a + a^{1/2}s)^2 + a}.$$

2. This is a consequence of the identity from measure theory that states $\lambda(AR) = |\det A|\lambda(R)$ for any invertible matrix A and measurable set R, and the fact that $\det S_S A_a = a^{3/2}$ and $\lambda(U_\psi) = 2$.

□

We can use Proposition 4.0.14 to derive a simple estimate for $|[\mathcal{SH}m_{R,h}](a,s,t)|$.

Propostion 4.0.15

For any $a \in (0,1)$ and $s \in (-2,2)$ we have

1. *If $||t|| > R^{-1} + 4$ then $|[\mathcal{SH}m_R](a,s,t)| = 0$, and*
2. *if $||t|| \le R^{-1} + 5$ then $|[\mathcal{SH}m_R](a,s,t)| \le 2a^{3/4}R^{1/2}||\chi||_\infty ||\psi||_{L^2}$.*

proof: First note that the Shearlet coefficients of $m_{R,h}$ are given by

$$\begin{aligned}[\mathcal{SH}m_R](a,s,t) &= \langle m_R, \psi_{a,s,t}\rangle \\ &= \int_{\mathbb{R}^2} m_R(x)\psi_{a,s,t}(x)dx \\ &= a^{-3/4}\int_{\mathbb{R}^2} m_R(x)\psi(A_a^{-1}S_s^{-1}(x-t))dx. \end{aligned} \tag{4.6}$$

Now note that the integrand in (4.6) is non zero if and only if the following inclusions hold;

$$x \in \mathcal{M} \quad \text{and} \quad A_a^{-1}S_s^{-1}(x-t) \in U_\psi.$$

The second inclusion is equivalent to

$$x \in S_sA_a(U_\psi) + t =: U_{a,s,t}.$$

Now if we define the set

$$B_{a,s,t} := \mathcal{M} \cap U_{a,s,t}.$$

Then we can restate our original statement as

$$m_R(x)\psi(A_a^{-1}S_s^{-1}(x-t)) \neq 0 \iff x \in B_{a,s,t}.$$

This allows us to reformulate (4.6) as

$$[\mathcal{SH}m_R](a,s,t) = a^{-3/4}\int_{B_{a,s,t}} m_R(x)\psi(A_a^{-1}S_s^{-1}(x-t))dx. \tag{4.7}$$

Reformulating the shearlet coefficient in this way immediately allows us to read off an important fact about $[\mathcal{SH}m_R](a,s,t)$. If $B_{a,s,t} = \emptyset$ then $[\mathcal{SH}m_R](a,s,t)$ is an integral over the empty set and as such will be zero. Thus we only need to consider the case

$$B_{a,s,t} \neq \emptyset.$$

1. Now we can prove part 1 of Proposition 4.0.15. For a fixed $a \in (0,1)$ and $s \in (-2,2)$ we know from Proposition 4.0.14 that the diameter l of $U_{a,s,t}$ is given by

$$l = \sqrt{(a + a^{1/2}s)^2 + a}.$$

Now if $||t|| \geq R^{-1} + l$ then $B_{a,s,t} = \emptyset$ since $U_{a,s,t}$ is too far a way. So for fixed $a \in (0,1)$ and $s \in (-2,2)$ we know that $|[\mathcal{SH}m_R](a,s,t)| = 0$ for $||t|| \geq R^{-1} + l$. Notice that for all $a \in (0,1)$ and $s \in (-2,2)$ we have

$$l = \sqrt{(a + a^{1/2}s)^2 + a} \leq \sqrt{(1 + 1^{1/2}2)^2 + 1} = \sqrt{10} \leq 4.$$

Therefore we can conclude that for all $a \in (0,1)$ and $s \in (-2,2)$ if $||t|| \geq R^{-1} + 4$ then $|[\mathcal{SH}m_R](a,s,t)| = 0$.

2. Now we can derive an estimate for the non-zero shearlet coefficients using Proposition 4.0.14.2. We compute

$$\begin{aligned}
|[\mathcal{SH}m](a,s,t)| &= \left| a^{-3/4} \int_{B_{a,s,t}} m_R(x)\psi(A_a^{-1}S_s^{-1}(x-t))dx \right| \\
&\leq a^{-3/4} \int_{\mathbb{R}^2} |\chi_{B_{a,s,t}}(x) m_R(x)\psi(A_a^{-1}S_s^{-1}(x-t))dx| \\
&\leq a^{-3/4}\lambda(B_{a,s,t})^{1/2}||m_R||_\infty ||\psi(A_a^{-1}S_s^{-1}(\cdot - t))||_{L^2} && \text{[Cauchy-Schwartz]} \\
&= \lambda(B_{a,s,t})^{1/2}||m_R||_\infty||\psi||_{L^2} \\
&\leq \lambda(U_{a,s,t})^{1/2}||m_R||_\infty||\psi||_{L^2} && [B_{a,s,t} \subseteq U_{a,s,t}] \\
&\leq 2a^{3/4}||m_R||_\infty||\psi||_{L^2}. && \text{[Proposition 4.0.14.2]}
\end{aligned}$$

Since

$$m_R(\xi) = R^{1/2}\chi(R||\xi|| - R)$$

we have that

$$||m_R||_\infty = R^{1/2}||\chi||_\infty.$$

So we find that shearlet coefficients are bounded by

$$|[\mathcal{SH}m_R](a,s,t)| \leq 2a^{3/4}R^{1/2}||\chi||_\infty||\psi||_{L^2}.$$

□

Using Proposition 4.0.15.1 we can rewrite (4.5) as

$$||\Psi_m^h||_{L^p} \leq (2\pi)^{-1}||f||_{L^2}\int_{||t||\leq R^{-1}+4}\int_{-2}^{2}\int_0^1 a^{-3}|[\mathcal{SH}_\psi m_R](a,s,t)| \left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} dadsdt, \tag{4.8}$$

We could substitute in Proposition 4.0.15.2 as well, but this will lead to problems with the convergence of the integral integral over a. So close to $a = 0$ we need to find a better estimate. To do this we will use the directionally vanishing moments of ψ from Definition 4.0.8.5.

Propostion 4.0.16

For all $s \in (-2, 2)$ and $t \in \mathbb{R}^2$ with $||t|| \leq R^{-1} + 4$, if $a \leq R^{-1}$ then

$$|[\mathcal{SH}m_R](a,s,t)| \leq a^n R^n ||\omega||_{L^2} \left|\left|\left(\frac{\partial}{\partial x}\right)^n ||\cdot||\right|\right|_\infty ||\chi'(||\cdot||)||_{L^2} \tag{4.9}$$

proof: First we note, since

$$\psi(x,y) = \psi_1(x)\psi_2(x)$$

and

$$A_a^{-1}S_s^1\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}a^{-1} & 0\\0 & a^{-1/2}\end{pmatrix}\begin{pmatrix}1 & -s\\0 & 1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}a^{-1}(x-sy)\\a^{-1/2}y\end{pmatrix}$$

we have that

$$\psi_{a,s,t} = a^{-3/4}\psi_1(a^{-1}(x-sy) - a^{-1}(t_1 - st_2)))\psi_2(a^{-1/2}y - a^{-1/2}t_2).$$

Recall that we assumed that there exists smooth and compactly supported functions Θ_1, Θ_2 such that

$$\psi_1(x) = \frac{d^n}{dx^n}\Theta_1(x) \quad \text{ and } \psi_2(x) = \frac{d^n}{dy^n}\Theta_2(x),$$

for some $n \in \mathbb{N}$. This allows us to write

$$\psi_{a,s,t}(v,y) = a^{-3/4}a^n\left[\left(\frac{d}{dx}\right)^n\Theta_1(a^{-1}(x-sy) - a^{-1}(t_1 - st_2))\right]\psi_2(a^{-1/2}(y-t_2)).$$

Now we can estimate $|[\mathcal{SH}m_R]|(a,s,t)$, but first we note that integration by parts implies that

$$\begin{aligned}&\int_{\mathbb{R}}\left[\left(\frac{d}{dx}\right)^n\Theta_1(a^{-1}(x-sy) - a^{-1}(t_1 - st_2))\right]m_R(x,y)dx\\&= \int_{\mathbb{R}}\Theta_1(a^{-1}(x-sy) - a^{-1}(t_1 - st_2))\left[\left(\frac{\partial}{\partial x}\right)^n m_R(x,y)\right]dx.\end{aligned} \tag{4.10}$$

The boundary terms went to zero since Θ_1 and m_R are Schwartz functions. The derivatives of m_R exist because it is a Schwartz function. Now define

$$\omega(x,y) = \Theta_1(x)\psi_2(y)$$

and

$$\omega_{a,s,t}(v,y) = \omega(a^{-1}(x-sy) - a^{-1}(t_1 - st_2), a^{-1/2}(y-t_2)) = \omega(A_a^{-1}S_s^{-1}((x,y)^T - t)).$$

This means that

$$||\omega_{a,s,t}||_{L^2} = a^{3/4}||\omega||_{L^2}.$$

Then we compute

$$\begin{aligned} \left(\frac{\partial}{\partial x}\right)^n m_R(x,y) &= R^{1/2}\left(\frac{\partial}{\partial x}\right)^n \chi(R||(x,y)|| - R) \\ &= R^{1/2+n}\chi'(R||(x,y)|| - R)\left[\left(\frac{\partial}{\partial x}\right)^n ||(x,y)||\right]. \end{aligned} \tag{4.11}$$

Note, for $n > 2$ the nth derivative of the Euclidean norm is bounded and since χ is a Schwartz function, so are its derivatives. Therefore $\chi'(||\cdot||)$ is a 2-dimensional Schwartz function. At the moment we can choose n however we need to since we haven't decided on a specfic value. When we choose a specific n later we will choose $n \geq 2$. Then we have that

$$\begin{aligned} \left|\left|\left(\frac{\partial}{\partial x}\right)^n m_R(x,y)\right|\right|_{L^2} &\leq R^{1/2+n}\left|\left|\left(\frac{\partial}{\partial x}\right)^n ||\cdot||\right|\right|_{\infty} ||\chi'(R||\cdot|| - R)||_{L^2} \\ &= R^n \left|\left|\left(\frac{\partial}{\partial x}\right)^n ||\cdot||\right|\right|_{\infty} ||\chi'(||\cdot||)||_{L^2} \end{aligned} \tag{4.12}$$

By (4.10), (4.12) and substituting in ω we can say that

$$\begin{aligned} &|[\mathcal{SH}m_R](a,s,t)| = |\langle m_R, \psi_{a,s,t}\rangle| \\ &= a^{-3/4+n}R^n\left|\int_{\mathbb{R}}\int_{\mathbb{R}} \omega_{a,s,t}(x,y)\left(\frac{\partial}{\partial x}\right)^n m_R(x,y)dxdy\right| \\ &\leq a^{-3/4+n}R^n||\omega_{a,s,t}||_{L^2}\left|\left|\left(\frac{\partial}{\partial x}\right)^n m_R(x,y)\right|\right|_{L^2} \quad \text{[Cauchy-Schwartz Inequality]} \\ &= a^n R^n ||\omega||_{L^2}\left|\left|\left(\frac{\partial}{\partial x}\right)^n ||\cdot||\right|\right|_{\infty} ||\chi'(||\cdot||)||_{L^2}. \end{aligned} \tag{4.13}$$

However we do not want our estimate to blow up as R goes to infinity, so we need the a^n term to dominate the R^n term. If we pick $a \leq R^{-1}$ then

$$a^n R^n \leq 1$$

then (4.13) is constant as R goes to infinity. So we accept (4.13) if $a \leq R^{-1}$. □

We can combine(4.8), Theorem (4.0.15.2) and (4.13) to find

$$\begin{aligned}
&(2\pi)||\Psi_m^h||_{L^p}\\
\leq& 2R^{1/2}||f||_{L^2}||\chi||_\infty||\psi||_{L^2}\int_{||t||\leq R^{-1}+4}\int_{-2}^{2}\int_{R^{-1}}^{1} a^{-9/4}\ \left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} dadsdt \qquad (4.14)\\
&+R^n||f||_{L^2}||\omega||_{L^2}\left|\left|\left(\frac{\partial}{\partial x}\right)^n||\cdot||\right|\right|_\infty ||\chi'(||\cdot||)||_{L^2}\int_{||t||\leq R^{-1}+4}\int_{-2}^{2}\int_{0}^{R^{-1}} a^{-3+n}\left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} dadsdt.
\end{aligned}$$

The last thing that we need to do is calculate the L^r norm of $\widehat{\psi_{a,s,t}}$. Before we can do so we need to mention a general theorem about the Fourier Transforms and coordinate transforms.

Propostion 4.0.17

Let $f \in L^1(\mathbb{R}^N)$ and let $A \in GL_N(\mathbb{R})$ (Real, invertible $N \times N$ martices) then for all $\xi \in \mathbb{R}^N$ we have

$$\mathcal{F}[f(A\cdot)](\xi) = |\det A|^{-1}\hat{f}((A^{-1})^T\xi).$$

Proposition 4.0.17 follows directly from the Transformation Theorem from Measure Theory, [Theorem 15.1 Chapter 15 page 154 of [14]]. Using this result we can now calculate the L^r norm of $\widehat{\psi_{a,s,t}}$.

Propostion 4.0.18

For fixed $a \in (0,1), s \in (-2,2)$ and $t \in \mathbb{R}^2$ we have that

$$\left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} = a^{\frac{3r-6}{4r}}||\hat{\psi}||_{L^r}. \qquad (4.15)$$

proof: Since L^r norms are conserved under translation we assume $t = 0$. Then recall that

$$\psi_{a,s,0}(x) = a^{-3/4}\psi(A_a^{-1}S_s^{-1}x).$$

We can then compute

$$\begin{aligned}
\widehat{\psi_{a,s,t}}(\xi) &= a^{-3/4}\,\mathcal{F}[\psi(A_a^{-1}S_s^{-1}\cdot)](\xi)\\
&= a^{-3/4}|\det(S_sA_a)|\hat{\psi}(A_a^TS_s^T\xi)\\
&= a^{3/4}\hat{\psi}(A_a^TS_s^T\xi).
\end{aligned}$$

It then follows that for any $r \in [1, \infty]$

$$\left|\left|\widehat{\psi_{a,s,t}}\right|\right|_{L^r} = a^{3/4} a^{-\frac{3}{2r}} ||\hat{\psi}||_{L^r} = a^{\frac{3r-6}{4r}} ||\hat{\psi}||_{L^r}.$$

□

Theorem 4.0.19

For any $f \in L^2(\mathbb{R}^2)$ and $p \in [2, \infty)$ we have

$$||\Psi_m^h f||_{L^p} \leq C||f||_{L^2}(R^{-1}+4)^2 \left(R^{\frac{2r+3}{2r}} + R^{\frac{7r+3}{2r}} - R^{\frac{1}{2}}\right), \tag{4.16}$$

for some $C > 0$ and

$$r = \frac{2p}{2+p}.$$

proof: Plugging (4.15) into (4.14) we find that

$$\begin{aligned}
&(2\pi)||\Psi_m^h f||_{L^p} \\
\leq& 2R^{1/2}||f||_{L^2}||\chi||_\infty ||\psi||_{L^2}||\hat{\psi}||_{L^r} \int_{||t|| \leq R^{-1}+4} \int_{-2}^{2} \int_{R^{-1}}^{1} a^{-\frac{3}{2}\left(1+\frac{1}{r}\right)} \, dadsdt \\
&+ R^n ||f||_{L^2}||\omega||_{L^2} \left|\left|\left(\frac{\partial}{\partial x}\right)^n || \cdot ||\right|\right|_\infty ||\chi'(|| \cdot ||)||_{L^2}||\hat{\psi}||_{L^r} \int_{||t|| \leq R^{-1}+4} \int_{-2}^{2} \int_{0}^{R^{-1}} a^{-\frac{3}{2}\left(3+\frac{1}{r}\right)+n} dadsdt \\
=& 8\pi(R^{-1}+4)^2 R^{1/2}||f||_{L^2}||\chi||_\infty ||\psi||_{L^2}||\hat{\psi}||_{L^r} \int_{R^{-1}}^{1} a^{-\frac{3}{2}\left(1+\frac{1}{r}\right)} \, da \\
&+ 4\pi(R^{-1}+4)^2 R^n ||f||_{L^2}||\omega||_{L^2} \left|\left|\left(\frac{\partial}{\partial x}\right)^n || \cdot ||\right|\right|_\infty ||\chi'(|| \cdot ||)||_{L^2}||\hat{\psi}||_{L^r} \int_{0}^{R^{-1}} a^{-\frac{3}{2}\left(3+\frac{1}{r}\right)+n} da \\
=& 8\pi(R^{-1}+4)^2 R^{1/2}||f||_{L^2}||\chi||_\infty ||\psi||_{L^2}||\hat{\psi}||_{L^r} \left(R^{\frac{r+3}{2r}} - 1\right)\left(\frac{r+3}{2r}\right) \\
&+ 4\pi(R^{-1}+4)^2 R^n ||f||_{L^2}||\omega||_{L^2} \left|\left|\left(\frac{\partial}{\partial x}\right)^n || \cdot ||\right|\right|_\infty ||\chi'(|| \cdot ||)||_{L^2}||\hat{\psi}||_{L^r} \int_{0}^{R^{-1}} a^{-\frac{3}{2}\left(3+\frac{1}{r}\right)+n} da.
\end{aligned}$$

The last integral is not guaranteed to be convergent yet, but this isn't a problem. Note that we haven't chosen a specific value for n yet. This is because we introduced it specifically with this integral in mind. To get convergence we need

$$-\frac{3}{2}\left(3+\frac{1}{r}\right) + n > -1 \tag{4.17}$$

which means

$$n > \frac{7r+3}{2r}.$$

Now recall that $p \in [2, \infty)$ and

$$r = \frac{2p}{2+p}$$

so $r \in [1,2)$. If we pick $n > 5$ then (4.17) is satisfied for all $r \in [1,2)$, since $n \in \mathbb{N}$ we pick $n = 6$. Thus assuming that $n \geq 2$ in the proof of Propostion 4.0.16 was reasonable. Finally we can say that

$$\begin{aligned}||\Psi^h_m f||_{L^p} \leq & 4(R^{-1}+4)^2 R^{1/2}||f||_{L^2}||\chi||_\infty||\psi||_{L^2}||\hat{\psi}||_{L^r}\left(R^{\frac{r+3}{2r}}-1\right)\left(\frac{r+3}{2r}\right) \\ & +2(R^{-1}+4)^2 R^6||f||_{L^2}||\omega||_{L^2}\left|\left|\left(\frac{\partial}{\partial x}\right)^n||\cdot||\right|\right|_\infty ||\chi'(||\cdot||)||_{L^2}||\hat{\psi}||_{L^r}\left(\frac{2r}{5r-3}\right)R^{\frac{3-5r}{2r}}.\end{aligned} \tag{4.18}$$

However we need to check that this bound is non-negative to ensure that it doesn't give a contradiction. The only terms that could be negative are

$$\left(R^{\frac{r+3}{2r}}-1\right)$$

and

$$\left(\frac{2r}{5r-3}\right).$$

But since $r \in [1,2)$ and $R > 1$ we have that both of these terms are positive, and therefore the estimate is positive. Since all functional norms in (4.18) are constant in R it then follows that

$$||\Psi^h_m f||_{L^p} \leq C||f||_{L^2}(R^{-1}+4)^2\left(R^{\frac{2r+3}{2r}}+R^{\frac{7r+3}{2r}}-R^{\frac{1}{2}}\right),$$

for some $C > 0$. □

The proofs of Proposition 4.0.12 and Theorem 4.0.19 can be repeated in the vertical frequency cone to conclude that there exist $C_1, C_2 > 0$ so that

$$\begin{aligned}&||\Phi^v_m f||_{L^p} \leq C_1 R^{-1/2}||f||_{L^2}, \\ &||\Psi^v_m f||_{L^p} \leq C_2||f||_{L^2}(R^{-1}+4)^2\left(R^{\frac{2r+3}{2r}}+R^{\frac{7r+3}{2r}}-R^{\frac{1}{2}}\right).\end{aligned}$$

Where r has the usual dependence on p. We can now finally conclude that T_{m_R} also defines a bounded linear operator from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ for any $p \in [2,\infty]$.

Theorem 4.0.20

For every $R > 1$ the convolution operator generated by m_R, given by

$$T_{m_R} f = \mathcal{F}^{-1}[m_R \hat{f}], \qquad f \in L^2(\mathbb{R}^2),$$

is a bounded linear operator from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ for any $p \in [2, \infty]$, with

$$||T_{m_R} f||_{L^p} \leq C||f||_{L^2} \left[R^{-1/2}(R^{-1}+4)^2 \left(R^{\frac{2r+3}{2r}} + R^{\frac{7r+3}{2r}} - R^{\frac{1}{2}}\right)\right],$$

where $C > 0$ is a constant and where

$$r = \frac{2p}{2+p}.$$

proof: Since $m_R \in L^2(\mathbb{R}^2)$, the case $p = \infty$ is dealt with in Proposition 4.0.3, and since $m_R \in L^\infty(\mathbb{R}^2)$ the case $p = 2$ is dealt with in Proposition 4.0.2.

So suppose that $p \in (2, \infty)$. Then by Theorem 4.0.10 we know that we can write

$$T_{m_R} f(x) = C_\psi^{-1}(\Psi_m^h f(x) + \Psi_m^v f(x) + \Phi_m^h f(x) + \Phi_m^v f(x)).$$

Then by Proposition 4.0.12, Theorem 4.0.19 and their analogues in the verical frequency cone we have

$$\begin{aligned}
||T_{m_R} f||_{L^p} &= C_\psi^{-1}||\Psi_m^h f + \Psi_m^v f + \Phi_m^h f + \Phi_m^v f||_{L^p} \\
&\leq C_\psi^{-1}(||\Psi_m^h f||_{L^p} + ||\Psi_m^v f||_{L^p} + ||\Phi_m^h f||_{L^p} + ||\Phi_m^v f||_{L^p}) \quad \text{[Triangle Inequality]} \\
&\leq C_1 R^{-1/2}||f||_{L^2} + C_2||f||_{L^2}(R^{-1}+4)^2 \left(R^{\frac{2r+3}{2r}} + R^{\frac{7r+3}{2r}} - R^{\frac{1}{2}}\right) \\
&= C||f||_{L^2} \left[R^{-1/2}(R^{-1}+4)^2 \left(R^{\frac{2r+3}{2r}} + R^{\frac{7r+3}{2r}} - R^{\frac{1}{2}}\right)\right],
\end{aligned}$$

where $C > 0$ is a constant and

$$r = \frac{2p}{2+p}.$$

Therefore T_{m_R} is a bounded linear operator from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$. □

Before we conclude it is also worth considering the behaviour of the L^p norms of the shearlet and auxiliary parts of $T_{m_R} f$ as R goes to infinity. Recall from Proposition 4.0.12 that

$$||\Phi_m^h f||_{L^p} \leq CR^{-1/2}||\hat{\phi}||_{L^r}.$$

Since the only term that isn't constant in R is a negative power of R we can see that $||\Phi_m^h f||_{L^p}$ goes to zero as R goes to infinity. Similarly from Theorem 4.0.19 we have that

$$||\Psi_m^h||_{L^p} \leq C||f||_{L^2}(R^{-1}+4)^2\left(R^{\frac{2r+3}{2r}}+R^{\frac{7r+3}{2r}}-R^{\frac{1}{2}}\right).$$

Now taking a leading order expansion for large R we find that

$$||\Psi_m^h||_{L^p} \leq C'R^{\frac{7r+3}{2r}}.$$

The right hand side diverges as R goes to infinity due to the positive power of R. So we cannot use our estimates to claim that the family of convolution operators
$T := \{T_{m_R}|R>0\}$ converges to a well defined operator as R goes to infinity. If we want to show that the delta distribution along the unit circle generates a bounded convolution operator we will have to take a different approach using more techniques from functional analysis. It is also worth trying different transforms that are tailored towards this purpose.

Conclusion

In this dissertation we have reviewed techniques of function synthesis and analysis by taking an in depth look at three common transforms and their associated reconstruction formulas. The three transforms we covered were the Fourier transform, the wavelet transform and the shearlet transform. After reviewing these transforms we applied the shearlet transform to the study of L^2 to L^p behaviour of a particular family of convolution operators.

The first chapter isa summary of an honours level course in Fourier analysis. The Fourier transform was introduced first because it provides an easy introduction to integral transforms and function reconstruction. Rather than focussing on the proofs of propositions and theorem we focussed instead on definitions and the results themselves. We did this because the Fourier transform was fundamental to the rest of this dissertation. Almost every proof from Chapter 2 onwards utilised the Fourier transform in some way, usually by applying Plancherel's theorem (Theorem 1.1.4), Parseval's theorem (Theorem 1.1.9) or the Fourier Convolution theorem (Proposition 1.1.13).

After introducing the Fourier transform we discussed its lack of localisation and how that affects its ability as a tool for analysing local behaviour of functions. This justified the introduction of another, more localised, transform called the wavelet transform. Chapter 2 was essentially a review of the first two chapters of [5]. We then saw firstly how the wavelet transform is defined and discussed the underlying intuition. Then we covered various reconstruction formulas and introduced the of a "weak" reconstruction formula Proposition 2.2.2 and Definition 2.2.3. This was an important idea since all of the reconstruction formulas held only in this “weak” sense. Chapter 2 was concluded with a section on how the wavelet transform can be used to analyse global and local Hölder regularity of functions on $\mathbb{R}$.

In Chapter 3 we saw that in higher dimensions the wavelet transform lacks the ability to detect curvature. In the way that the wavelet transform is a response to the Fourier transform's lack of localisation, the shearlet transform is a response to the wavelet transform's lack of sensitivity to directionality. After reviewing the continuous shearlet transform we see that it had a directional bias, which lead us to discussing the cone adapted shearlet transform. Both forms of the shearlet transform have associated reconstruction formulas that we discussed in detail. As with the wavelet transform, these reconstruction formulas hold in the weak sense of Definition 2.2.3.

In Chapter 4 we apply the shearlet transform to a problem in functional analysis. Specifically we analysed the L^2 to L^p behaviour of a family of convolution operators defined by a family of 2-dimensional multiplier functions concentrated on the unit circle. Our main strategy was to reformulate the convolution operator by applying the Fourier transform and reconstructing the multipliers using the cone adapted shearlet transform. We were then able to show that this family of convolution operators is a family of bounded linear operators from $L^2(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ for any $p \in [2, \infty]$.

While we were able to arrive at a result in Chapter 4, this did not come without its fair share of difficulty. We found that while the shearlet transform works very well in the numerical setting, it can make things quite complicated in the analytic setting. The main problem we faced was that the shearlet transform is strongly dependent on the geometry of the problem we are applying it to. This means that it is easy to get stuck on solving geometry problems that have little to do with the problem we are actually trying to solve. Furthermore, the orientation and scaling of the shearlet are quite strongly linked, as seen in Proposition 4.0.14. This can make dealing with the dilation and shearing parameters separately quite difficult. So we suspect that for future work it would be worth investigating a transform that independently scales and rotates the analysing functions. An example of such a transform uses a rotation matrix rather than a shearing matrix to change the orientation of the analysing functions. This was briefly mentioned at the start of Chapter 3. There are two transforms that use this idea, namely the curvelet transform and the ridglet transform. See [2] for more on curvelets and [1] for more on ridgelets.

We have seen throughout this dissertation the usefulness of transforms and their associated inverse transforms. However, we have also seen that each transform has its own strengths and weaknesses. Each transform is specialised to analyse specific properties of functions. Which transform we choose to apply should depend on our goal. We have reviewed three powerful transforms and can now add them to our mathematical tool belt.

Bibliography

[1] Emmanuel J Candès and David L Donoho, *Ridgelets: A key to higher-dimensional intermittency?*, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **357** (1999), no. 1760, 2495–2509.

[2] Emmanuel J Candes and David L Donoho, *Curvelets: A surprisingly effective nonadaptive representation for objects with edges*, Stanford Univ Ca Dept of Statistics, 2000.

[3] David Cruz-Uribe and Alberto Fiorenza, *Convergence in measure of approximate identities in variable Lebesgue spaces*, Anal. Appl. (Singap.) **13** (2015), no. 4, 413–418. MR3340668

[4] Stephan Dahlke, Gitta Kutyniok, Peter Maass, Chen Sagiv, Hans-Georg Stark, and Gerd Teschke, *The uncertainty principle associated with the continuous shearlet transform*, Int. J. Wavelets Multiresolut. Inf. Process. **6** (2008), no. 2, 157–181. MR2450406

[5] Ingrid Daubechies, *Ten lectures on wavelets*, Philadelphia, Pa. : Society for Industrial and Applied Mathematics (English).

[6] Philipp Grohs, *Continuous shearlet tight frames*, J. Fourier Anal. Appl. **17** (2011), no. 3, 506–518. MR2803946

[7] ———, *Shearlets and microlocal analysis*, Shearlets, 2012, pp. 39–67. MR2896275

[8] Matthias Holschneider and Philippe Tchamitchian, *Régularite locale de la fonction "non-différentiable" de Riemann*, Les ondelettes en 1989 (Orsay, 1989), 1990, pp. 102–124, 209–210. MR1083586

[9] Ram P. Kanwal, *Generalized functions: theory and applications*, 3rd ed., Boston:Birkhauser (English).

[10] Gitta Kutyniok and Demetrio Labate, *Introduction to shearlets*, Shearlets, 2012, pp. 1–38.

[11] ———, *Shearlets: Multiscale analysis for multivariate data*, Springer Science & Business Media, 2012.

[12] Gitta Kutyniok, Jakob Lemvig, and Wang-Q Lim, *Shearlets and optimally sparse approximations*, Shearlets, 2012, pp. 145–197.

[13] Stephane Mallat et al., *A wavelet tour of signal processing: The sparce way*, AP Professional, Third Edition, London (2009).

[14] Rene L. Schilling, *Measures, integrals and martingales*, Second Edition, Cambridge, United Kingdom ; New York, NY : Cambridge University Press (English).

[15] Melissa Tacy, *Fourier analysis and distribution theory lecture notes.*

[16] Anders Vretblad, *Fourier analysis and its applications*, New York : Springer (English).