FFI RAPPORT

CHARACTERIZATION OF LOCAL AND GLOBAL REGULARITY BY THE CONTINUOUS WAVELET TRANSFORM, APPLIED TO REAL AND SYNTHETIC DATA

TORSÅS Morten

FFI/RAPPORT-2007/00117

CHARACTERIZATION OF LOCAL AND GLOBAL REGULARITY BY THE CONTINUOUS WAVELET TRANSFORM, APPLIED TO REAL AND SYNTHETIC DATA

TORSÅS Morten

FFI/RAPPORT-2007/00117

FORSVARETS FORSKNINGSINSTITUTT
Norwegian Defence Research Establishment
P O Box 25, NO-2027 Kjeller, Norway

FORSVARETS FORSKNINGSINSTITUTT(FFI)
Norwegian Defence Research Establishment

P O BOX 25
NO-2027 KJELLER, NORWAY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE
(when data entered)

REPORT DOCUMENTATION PAGE

1) PUBL/REPORT NUMBER	2) SECURITY CLASSIFICATION	3) NUMBER OF PAGES
FFI/RAPPORT-2007/00117	UNCLASSIFIED	96
1a) PROJECT REFERENCE	2a) DECLASSIFICATION/DOWNGRADING SCHEDULE	
FFI-IV/1062/914	-	

4) TITLE

CHARACTERIZATION OF LOCAL AND GLOBAL REGULARITY BY THE CONTINUOUS WAVELET TRANSFORM, APPLIED TO REAL AND SYNTHETIC DATA

5) NAMES OF AUTHOR(S) IN FULL (surname first)

TORSÅS Morten

6) DISTRIBUTION STATEMENT

Approved for public release. Distribution unlimited. (Offentlig tilgjengelig)

7) INDEXING TERMS

IN ENGLISH	IN NORWEGIAN
a) Wavelet	a) Wavelet
b) Continuous Wavelet Transform (CWT)	b) Kontinuerlig Wavelettransform
c) Regularity	c) Regularitet
d) Edge Detection	d) Kantdeteksjon
e)	e)

THESAURUS REFERENCE: Tesaurus

8) ABSTRACT

This FFI Report is a reprint, with some minor bug-fixes, of my Cand. Scient. Thesis with the same title at The University of Oslo, Mathematics Department, spring 2006.
Some important Theorems regarding the connection between the decay of the Wavelet Transform across scales and the regularity of functions, locally and globally, will be presented and proved, and an algorithm for computing the local regularity of functions based on the WTMM (Wavelet Transform Modulus Maxima) methods derived from these Theorems will be described and used on a Laser Altimetry data set.

9) DATE	AUTHORIZED BY	POSITION
2007-01-10	This page only Elling Tveit	Director of Research

ISBN 978-82-464-1071-5

FFI-B-22-1982

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE
(when data entered)

CONTENTS

page

1	THESIS OVERVIEW	7
1.1	Overview	7
2	REGULARITY	8
2.1	Introduction	8
2.2	CWT	9
2.3	Hölder Regular Functions	14
2.4	Oscillating Singularities	22
2.5	Global Hölder Regularity	23
2.6	Local Hölder Regularity	24
2.7	Pointwise Differentiability	26
2.8	CWT Local Maxima	27
3	THE PROOFS	30
3.1	Hölder Regular Functions	30
3.2	Global Hölder Regularity	32
3.3	Local Hölder Regularity	33
3.4	CWT Local Maxima	36
4	WTMM ALGORITHM	59
4.1	Introduction	59
4.2	Data Description	59
4.3	WTMM Algorithm	60
4.4	Synthetic Data	68
4.5	The Sandvika Data Set	69
4.6	Conclusions	76
5	THESIS SUMMARY	76

APPENDIX

A	PRELIMINARIES	78
A.1	Introduction	78
A.2	Integration Theory	78
A.3	General Theory	82
A.4	Vector Spaces	82
A.5	Distribution theory	88
A.6	Fourier Transforms	90
A.7	Function Spaces	92
	References	92

CHARACTERIZATION OF LOCAL AND GLOBAL REGULARITY BY THE CONTINUOUS WAVELET TRANSFORM, APPLIED TO REAL AND SYNTHETIC DATA

1 THESIS OVERVIEW

1.1 Overview

The ‘irregular’ points or areas of a signal (1 to N dimensions) carry most of the information in the data. The objects in an image are outlined by their borders, i.e. the region (usually N-1 dimensional, i.e. a line in a 2D data set, a point in a 1D signal, etc), where the data value changes abruptly. Getting a description of where the edges or borders of objects are and some characterization of what type of edge it is, makes us able to separate real objects from noise and smooth areas with slow-moving changes.

The theory of characterizing function regularity by the decay of the wavelet transform has become standard wavelet theory since I started this thesis. The number of papers and books on this theme is enormous, and to mention all or lots of them here is beyond scope. The bibliography includes some, mostly from the pre 1995 period.

The purpose of this thesis is to present and prove the most important 1-D theorems regarding the connection between the continuous wavelet transform (CWT) and local and global Hölder regularity. The starting point was the paper by Mallat and Hwang (39) in the early 1990s. The work on this thesis has had some long and irregular breaks, due to un-mathematical events, but the search for related papers, thesis and books has never stopped completely. Many authors write about these topics, but it seems that the original paper covered most of what is todays knowledge on the theme. The book by Mallat from 1999 (37) and by Holschneider from 1995 (14) is the most interesting additional contributions that we have found. These will be presented without proofs. The proofs in the paper by Mallat and Hwang are not completed, so the main contribution by this thesis is the completion of these proofs in a consistent form, following the sketched proofs in that paper and in Holschneider, Tchamitchian (15) and Jaffard (18), (19). A description of the differences and similarities between the theorems in Mallat and Hwang and in these two other books is also added. In addition, a theorem from Jaffard, Meyer and Ryan (29), an algorithm for computing pointwise Hölder exponents, is included.

Secondly, the implementation of the WTMM (Wavelet Transform Modulus Maxima) algorithm in Matlab is done from scratch, except for the *cwt* function from Matlab’s Wavelet Toolbox. We are aware of other implementations of similar algorithms, including the WaveLab toolbox (2), but learning and using such large packages is both time-consuming in it self, and also makes it hard to understand exactly what is happening. So, making everything our self makes the result less perfect, but improves the understanding. This algorithm, like ‘all’ other such algorithms made by others are based, more or less stringent, on the theorems in this thesis. What is common to them all, is that they find the *modulus maxima lines* in the CWT plane, and estimate the decay of the wavelet transform along these lines. We then get an estimate of the regularity at the points that have a maxima line pointing at them.

We also want to include the definitions and preliminaries needed to prove the theorems and explain the theory.

Some work has been done to extract the ground from a ‘scene’ with buildings and vegetation from laser altimetry data in (57), (58), (59), using wavelet methods.

In Chapter 2, we will present all the mathematical results regarding Hölder regularity from Mallat and Hwang (39), Holschneider (14) and Mallat (37) and some related results, including some general theory and lemmas related to the topics. Most of the material here is gathered from these three sources.

In Chapter 3 we will prove the theorems from Mallat and Hwang (39), based on the more or less completed proofs there, and in a notation that is consistent with the rest of this thesis.

In Chapter 4, we want to implement an WTMM (Wavelet Transform Modulus Maxima) algorithm based on following the wavelet transform maxima across scales in all the 1D lines in each direction of some laser altimetry datasets, to identify and characterize the various singularities we discover, and then put all these lines together in a location-preserving matrix with the *characterized singularities*.

In Appendix A, the relevant mathematical preliminaries not included in the text is added.

2 REGULARITY

2.1 Introduction

In this chapter we want to present theorems from the book of Mathias Holschneider (14), the book of Stephane Mallat (37) and from the paper by Stephane Mallat and Wen Liang Hwang (38), regarding characterization of singularities by the asymptotic decay of the wavelet transform. Theorems from this last paper will be proven completely in Chapter 3. The similarities and differences between the results in Mallat and Hwang (38) from 1992 and the more recent Mallat (37) from 1999 will be commented, but the results in Holschneider (14) will only be presented without much comment. The most fundamental theorems will be named according to the naming and numbering in the book/paper they are found. The mathematical background, the definitions and the notations not included in the text are found in Appendix A.

Some of the theorems in the paper and the two books are if-and-only-if and some of them have slightly different conditions in the different directions. The theorems are separated, one part showing the properties of the wavelet transform of functions with pre-described Hölder regularity and another part showing consequences of the wavelet transform properties, i.e. a characterization of the regularity, the Hölder exponent, by the asymptotic decay of the wavelet transform. The naming convention of a)’s and b)’s for each direction is used.

In Section 2.2, the ‘Continuous Wavelet Transform’ (CWT) section, we define the continuous wavelet transform and some properties thereof, including theorems regarding necessary and sufficient conditions on functions to be *admissible*, i.e. *wavelets*.

In Section 2.3, the ‘Hölder Regular Functions’ section, we will analyse properties of functions with known Hölder regularity globally and locally, specifically the behavior of the CWT of such functions.

In Section 2.4, the 'Oscillating Singularity' section, oscillating singularities are defined and discussed.

In Section 2.5, the 'Global Hölder regularity' section, sufficient conditions on the CWT for the function to be Hölder regular globally or on an interval are outlined.

In Section 2.6, the 'Local Hölder regularity' section, similar conditions for local, i.e. at a point $x_0 \in \mathbb{R}$, are outlined.

Section 2.7, the 'Pointwise Differentiability' section, is a presentation of the implication these results has on the differentiability of functions.

Section 2.8, the 'Wavelet transform local maxima' section, shows rather strong results regarding the maxima lines of $|\mathcal{W}_\psi f(s,x)|$ in the time-frequency plane.

2.2 CWT

The continuous wavelet transform is defined by a convolution, which is a sliding of one function over another, specific the sliding of the wavelet over the function to be analysed or transformed. The Fourier transform is the classical 'frequency transform' which gives us the contribution of each frequency to the total signal. It is 'totally un-localized' in that a small change to a small part of the signal gives contribution to the whole transform, hiding the *localization* in the transformed signal. This can partly be fixed by windowing the frequency atom, i.e. by multiplying with a localized function. In the *Windowed Fourier Transform* we use a fixed window, and thus get partly localized information. In the wavelet transform, we have, in effect, a window that is dynamically scalable, giving us a much better localization property.

Definition 2.1 (Fourier Transform). *We define the* Fourier-transform $\hat{f}$ *of* f *as* $\hat{f}(\xi) = \int_{\mathbb{R}} f(t)e^{-it\xi}\, dt$.

The Fourier Transform is sometimes defined with a $\frac{1}{2\pi}$ or $\frac{1}{\sqrt{2\pi}}$ factor added for symmetry with the inverse transform or for isometric purposes, and these other definitions might be in use, by mistake, in some of the calculations in this thesis. Warning given.

A wavelet is a localized function with waves or oscillations, with average zero. It's a pretty 'weak' definition in that there are a lot of wavelets, giving us enormous freedom in selecting wavelets that suits our problem to be solved. The flexibility also offers challenges to the understanding of exactly what the analysis of a function shows us. There are many slightly different definitions of wavelets and of wavelet transforms, each emphasizing different perspectives of the theory. In the discrete theory, the Hilbert space $L^2(\mathbb{R})$ is often used with the many Hilbert space results as tools and with wavelet transforms defined by the $L^2(\mathbb{R})$ innerproduct. Similarly with periodic functions on $L^2[a,b]$or $L^2([0,2\pi))$. We will be using the standard definition for continuous wavelet transform which is the convolution-perspective.

Definition 2.2 (Admissible Wavelets and Wavelet Transform).

- $C_\psi = 2\pi \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|}\, d\xi$.
- $C_\psi^+ = \int_0^\infty \frac{|\hat{\psi}(\xi)|^2}{\xi}\, d\xi$.
- $C_\psi^- = \int_{-\infty}^0 \frac{|\hat{\psi}(\xi)|^2}{|\xi|}\, d\xi$.
- *If $C_\psi < \infty$, then we say that ψ is* admissible*, or that ψ satisfies* the admissibility condition.
- *ψ is a* wavelet *if $\psi \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $C_\psi^+ = C_\psi^- < \infty$.*
- *The Continuous Wavelet Transform (CWT):*

$$\mathcal{W}_\psi f(s,x) = (f * \psi_s)(x) = \int_{\mathbb{R}} f(u) \frac{1}{s} \psi \left(\frac{x-u}{s} \right)\, du.$$

For any real function f, we have the *Hermittian Symmetry* given by $\hat{f}(-\xi) = \overline{\hat{f}(\xi)}$, so for real wavelets, the condition $C_\psi^+ = C_\psi^-$ is automatically satisfied.

The CWT has a weak inverse, given by the following lemma:

Lemma 2.3. *Given $f \in L^1(\mathbb{R})$,*

$$f(x) = \frac{1}{C_\psi} \int_0^\infty \int_{\mathbb{R}} \mathcal{W}_\psi f(s,x) \overline{\psi_s(u-x)}\, du\, \frac{ds}{s},$$

where the equality is in a weak sense.

Proof. This is outlined on pages 24 and 25 in Daubechies (9) and Proposition 2.4.1 there. □

The admissibility condition on $L^1(\mathbb{R})$-functions implies that the function has average zero, which justifies the ‘wave’-part of the notion of ‘wavelets’. They are also time-localized or ‘small’, meaning that they are $L^1(\mathbb{R})$ and some are also compactly supported, which is the motivation for the ‘-lets’ part of the name. The mostly used compactly supported wavelets are the compactly supported spline- and Daubechies-wavelets, including the Haar-wavelet. Other examples of non compact but highly localized wavelet are the Gaussian family including the Mexican Hat wavelet. This family consists of derivatives of a Gaussian function.

Lemma 2.4. *If ψ is a wavelet then $\hat{\psi}(0) = \int \psi(x)dx = 0$.*

Proof. Suppose $\hat{\psi}(0) = \delta \neq 0$. Since $\hat{\psi}(\xi)$ is continuous by Lemma A.60, there exists $\epsilon > 0$ such that $|\hat{\psi}(\xi)| \geq \delta/2$ i $[0, \epsilon\rangle$. Consequently

$$C_\psi \geq \int_0^\epsilon \frac{\delta/2}{\xi} d\xi = \infty.$$

□

Lemma 2.5. *Let $k \in \mathbb{N}$, $\phi \in C^k(\mathbb{R})$ and suppose that $\phi^{(k)} \in L^2(\mathbb{R})$ is not identically zero. Then $\psi(x) = \phi^{(k)}(x)$ is admissible.*

Proof. Since $|\hat{\psi}(\xi)| = |\xi|^k|\hat{\phi}(\xi)|$ by induction on A.5 in Lemma A.60, page 90

$$\begin{aligned}
C_\psi &= 2\pi \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi \\
&= 2\pi \int_{\mathbb{R}} \frac{|\xi|^{2k}|\hat{\phi}(\xi)|^2}{|\xi|} \\
&= 2\pi \int_{-1}^{1} |\xi|^{2k-1}|\hat{\phi}(\xi)|^2 d\xi + 2\pi \int_{|\xi|>1} \frac{|\xi|^{2k}|\hat{\phi}(\xi)|^2}{|\xi|} d\xi \\
&\leq 2\pi \left(||\phi||^2_{L^\infty([-1,1])} + ||\phi^{(k)}||^2_{L^2(\mathbb{R})} \right) < \infty.
\end{aligned}$$

□

Lemma 2.6. *The set $\{f \in L^2(\mathbb{R}) : f$ is admissible$\}$ is dense in $L^2(\mathbb{R})$.*

Proof. $f \in L^2(\mathbb{R}) \Rightarrow \hat{f} \in L^2(\mathbb{R})$ by Lemma A.61. Let χ_A be the characteristic function of a set A and define $\hat{f}_\epsilon = \hat{f}(\xi)\chi_{\{\xi:|\xi|>\epsilon\}}(\xi)$. For every ϵ, f_ϵ is admissible. Since $||f||_{L^2(\mathbb{R})} = ||\hat{f}||_{L^2(\mathbb{R})}$,

$$||f - f_\epsilon||^2_{L^2(\mathbb{R})} = \int_{-\epsilon}^{\epsilon} |\hat{f}(\xi)|^2 d\xi \to 0 \text{ when } \epsilon \to 0, \tag{2.1}$$

so every $L^2(\mathbb{R})$-function is the limit of a sequence of admissible functions. □

This calls for some remarks. We see that any function in $L^2(\mathbb{R})$ with any average can be approximated by a function of zero average that satisfies the admissibility condition. Usually when we talk about wavelets we think of the Daubechies family, the Meyer family or some other well localized functions, usually localized somewhere around zero, and with a few approximately symmetric or anti-symmetric bumps. But from Lemma 2.6 we see that admissible functions can have about *any* shape or structure. This is of course not so for $L^1(\mathbb{R})$. As an illustration, let us consider an example:

Example. *Let $\psi_n : \mathbb{R} \to \mathbb{R}$ be defined by:*

$$\psi_n(x) = \begin{cases} 0 & \text{for } x \in (-\infty, 0) \cup [n+1, \infty) \\ 1 & \text{for } x \in [0, 1) \\ -1/n & \text{for } x \in [1, n+1), \end{cases} ,$$

and let $\chi_{(a,b)}(x) \in L^2(\mathbb{R})$ be the characteristic function of (a, b). We have:

$$\widehat{\chi_{(-1,1)}}(\xi) = 2\frac{\sin(x)}{x} = 2\,\mathrm{sinc}(\xi) \text{ and}$$

$$\widehat{\psi_s(\cdot - b)}(\xi) = e^{-ib\xi}\hat{\psi}(s\xi).$$

Then

$$\begin{aligned}\hat{\psi}_n(\xi) &= e^{-i\xi/2}\operatorname{sinc}(\xi/2) - e^{-i(\frac{n+2}{2})\xi}\operatorname{sinc}(n\xi/2)\\ &= e^{-i\xi}(e^{i\xi/2}\operatorname{sinc}(\xi/2) - e^{-in\xi/2}\operatorname{sinc}(n\xi/2),\end{aligned}$$

and

$$\frac{|\hat{\psi}_n(\xi)|^2}{|\xi|} < \lim_{\xi\to 0}\frac{|\hat{\psi}_n(\xi)|^2}{|\xi|} < 2n.$$

So

$$\begin{aligned}C_{\psi_n} &= 2\pi\int_{\mathbb{R}}\frac{|\hat{\psi}_n(\xi)|^2}{|\xi|}\,d\xi &(2.2)\\ &= 2\pi\int_{\mathbb{R}}\frac{|e^{i\xi/2}\operatorname{sinc}(\xi/2) - e^{-in\xi/2}\operatorname{sinc}(n\xi/2)|^2}{|\xi|}\,d\xi &(2.3)\\ &\le 2\pi\int_{|\xi|\le 1}2n\,d\xi + \int_{|\xi|>1}\frac{4}{|\xi|^3}\,d\xi &(2.4)\\ &< \infty. &(2.5)\end{aligned}$$

so ψ_n *is admissible* $\forall n\in\mathbb{N}$*, and*

- $\int_{\mathbb{R}} f_n(x)\,dx = 0,\ \forall n\in\mathbb{N}$,
- $f_n \to \chi_{(a,b)},\ n\to\infty$ *and*
- $||\chi_{(0,1)} - f_n||^2_{L^2(\mathbb{R})} = 1/n^2 \to 0,\ n\to\infty.$

So we have a set of admissible functions that converges to the characteristic function of $[0,1)$ *in* $L^2(\mathbb{R})$*. That shows that any simple function, which is a finite linear combination of characteristic functions, is in the closure of the set of admissible functions.*

Lemma 2.7. *Suppose* $0\neq\psi\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$, $\hat{\psi}(0)=0$ *and*

$$\int_{\mathbb{R}}|x|^{\beta}|\psi(x)|dx < \infty$$

for some $\beta > 1/2$*. Then* ψ *is admissible.*

Proof. This is Lemma 1.1.4 in Louis, Maas, Riedler (31). □

Corollary 2.8. *Suppose* $0\neq\psi\in L^2(\mathbb{R})$ *has compact support. Then*

$$\hat{\psi}(0) = 0 \Leftrightarrow \psi \text{ is admissible.} \tag{2.6}$$

Proof.

- $\Leftarrow$: If $\psi \in L^2(\mathbb{R})$ has compact support, then $\psi \in L^1(\mathbb{R})$ and the result follows from Lemma 2.4.
- $\Rightarrow$: The compact support of ψ gives $\int_{\mathbb{R}} |x|^{\beta}|\psi(x)|dx < \infty$ for all $\beta > 0$, so this follows from Lemma 2.7.

□

The following lemma is simply a splitting of a function f in one smooth ($C^n(I)$) part and a 'rest', the irregular part. It's simply a tool to be used in later sections.

Lemma 2.9. *Let $I = (a, b) \subset \mathbb{R}$ be an interval, $s_0 > 0$, $f \in L^1(I)$ and $\psi \in C^n(I)$, for $n \in \mathbb{N}$. Then*

$$\begin{aligned} f(x) &= \frac{1}{C_\psi}\int_0^{s_0}\int_{\mathbb{R}} \mathcal{W}_\psi f(s,u)\overline{\psi}_s(u-x)\,du\,\frac{ds}{s} && (2.7)\\ &+ \frac{1}{C_\psi}\int_{s_0}^{\infty}\int_{\mathbb{R}} \mathcal{W}_\psi f(s,u)\overline{\psi}_s(u-x)\,du\,\frac{ds}{s} && (2.8)\\ &= f_{small}(x) + f_{large}(x), && (2.9)\end{aligned}$$

where $f_{large}(x) \in C^n(I)$.

Proof. We have $\psi_s(u-x-h) - \psi_s(u-x) = \frac{h}{s^2}\psi'(\frac{u-x-\tau}{s})$ for some $\tau \in (0, h)$ by the *Mean Value Theorem*, so

$$\begin{aligned} 0 &\le |f_{large}(x+h) - f_{large}(x)| && (2.10)\\ &= \frac{1}{C_\psi}\int_{s_0}^{\infty}\int_{\mathbb{R}} \mathcal{W}_\psi f(s,x)(\overline{\psi_s}(u-x-h) - \overline{\psi_s}(u-x))du\frac{ds}{s}\\ &= \frac{1}{C_\psi}\int_{s_0}^{\infty}\int_{\mathbb{R}} \mathcal{W}_\psi f(s,x)\frac{h}{s^2}\overline{\psi'}(\frac{u-x-\tau}{s})du\frac{ds}{s}\\ &\le h\frac{||f||_{L^1(I)}||\psi||_{L^\infty(I)}||\psi'||_{L^1(I)}}{C_\psi}\int_{s_0}^{\infty}\frac{1}{s^2}\,ds\\ &= h\frac{||f||_{L^1(I)}||\psi||_{L^\infty(I)}||\psi'||_{L^1(I)}}{s_0 C_\psi}\\ &= Ch, && (2.11)\end{aligned}$$

and the result follows by induction on n. □

Example (Commonly used Wavelets). *As we have seen, there are an infinite number of wavelets, but there are a number of classes or families of wavelets that are more used and studied than others because of their nice properties. We will not be defining or analysing these families here, only mention and plot (Figure 2.1) some of the most famous and popular ones, and mention in particular two of the mostly used, namely the* Haar wavelet *and the* Mexican Hat *wavelet, whom we will be using later in this thesis.*

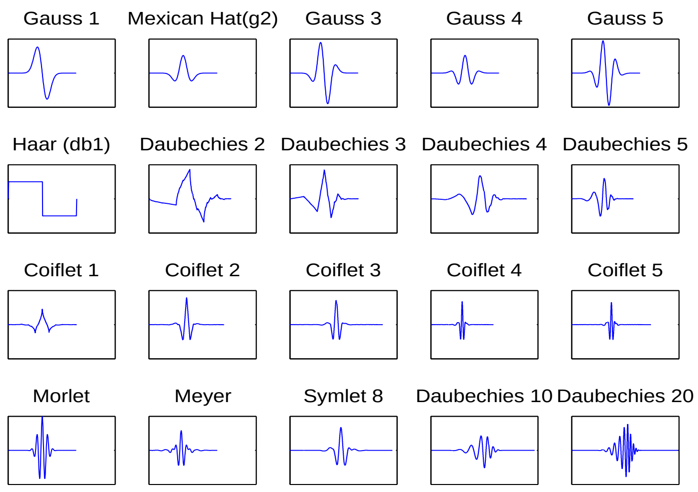

Figure 2.1 Different wavelet functions

2.3 Hölder Regular Functions

The smoothness of a function at a point or in some interval is a description of how fast it is changing. Is it continuous? Is it differentiable? Is its derivative continuous etc. The mathematical description of this is contained in the definition of Hölder regularity. Lipschitz regularity gives a similar description and is used by some authors, sometimes with a slightly different definition. The *Lipschitz condition* on a function is usually for the function to be Hölder 1.

First we include some notations.

Definition 2.10 ('Big O' (O) and 'Small o' (o)). *Let f and g be functions.*

- $f(x) = \mathbf{O}(g(x)),\ x \to x_0 \Leftrightarrow f(x)/g(x) < C < \infty$ *for* $|x - x_0| < \delta$.
- $f(x) = \mathbf{o}(g(x)),\ x \to x_0 \Leftrightarrow f(x)/g(x) \to 0$ *for* $|x - x_0| \to 0$.

'Big O' and 'Small o' are notations used to characterize upper bounds of a function by easily manageable functions when $x \to x_0$. For instance, if the modulus of a function f is bounded by a constant times $1/|x|$ when $x \to 0$, we write $f(x) = \mathbf{O}(1/x),\ x \to 0$.

Definition 2.11 (Fast Decay). *A function ψ has* fast decay *if*

$$|\psi(x)| = \mathbf{O}((1 + |x|^m)^{-1}),\ \forall m \in \mathbb{N},\ x \to \infty.$$

This definition shows us that a function has fast decay if it converges to zero when $x \to \infty$ even if it is multiplied by any polynomial.

Definition 2.12 (Smoothing function). *A* smoothing function *is a real function* $\Theta(x)$ *such that*

$$\Theta(x) = \mathbf{O}\left(\frac{1}{1+x^2}\right),$$

and

$$\hat{\Theta}(0) \neq 0.$$

The typical smoothing function used is a 'bump'-function, essentially positive, like the Gaussian or the characteristic function. The convolution of a smoothing function and a function is typically a smoothed version of the original function, explaining the name.

Definition 2.13 (Hölder regularity). *Let* $n \in \mathbb{N}$ *,* $n \leq \alpha < n+1$*. A function* $f(x)$ *is* Hölder α *at* x_0 *if there exist two constants* A *and* $h_0 > 0$*, and a polynomial* $P_n(x)$*, (typically the* Taylor Polynomial *if the function is* n *times differentiable,) of order* n *such that for* $h < h_0$*,*

$$|f(x_0+h) - P_n(h)| \leq A|h|^\alpha. \tag{2.12}$$

The supremum of all the values α *such that* f *is Hölder* α *at* x_0 *is called the* Hölder regularity *of* f *at* x_0*. A function is* singular *at* x_0 *if it is not Hölder* 1 *at* x_0*. A function* $f(x)$ *is* uniformly Hölder α *over an interval* (a,b) *if there exists a constant* A *such that for all* $x_0 \in (a,b)$ *there exists a polynomial of order* n*,* $P_n(x)$ *such that* $|f(x_0+h) - P_n(h)| \leq A|h|^\alpha$ *for any* $x_0 + h \in \langle a,b\rangle$*. If* f *is a tempered distribution of finite order ,* α *is a non-integer real number and* $[a,b] \subset \mathbb{R}$*, then the distribution* $f(x)$ *is* uniformly Hölder α *on* (a,b) *if its primitive is uniformly Hölder* $\alpha+1$ *on* (a,b)*. A distribution* f *has an* isolated singularity Hölder α *at* x_0 *if* $f(x)$ *is uniformly Hölder* α *over an interval* (a,b)*, with* $x_0 \in (a,b)$*, and* f *is uniformly Hölder* 1 *over any subinterval of* (a,b) *that does not include* x_0*.*

A more general definition of Hölder-like properties are included in the following function spaces, which are used in Holschneider (14):

Definition 2.14 (Λ^α**,** λ^α**,** $\Lambda^{\alpha,\beta}_{\log}$**,** $\lambda^{\alpha,\beta}_{\log}$**).** *Let* $n \in \mathbb{N}$ *and* P_n *be the polynomial of degree at most* n *that best approximates the function* f *in a neighborhood of* x_0 *(the Taylor Polynomial if f is* n *times differentiable) and let*

$$f(x_0 + x) = P_n(x) + f_{loc}(x).$$

- $f \in \Lambda^\alpha(x_0),\ n < \alpha \leq (n+1) \Leftrightarrow f_{loc}(x) = \mathbf{O}(x^\alpha),\ (x \to 0).$
- $f \in \lambda^\alpha(x_0),\ n \leq \alpha < (n+1) \Leftrightarrow f_{loc}(x) = \mathbf{o}(x^\alpha),\ (x \to 0).$
- $f \in \Lambda^{\alpha,\beta}_{\log}(x_0),\ n < \alpha \leq (n+1) \Leftrightarrow f_{loc}(x) = \mathbf{O}(x^\alpha \log^\beta x),\ (x \to 0).$
- $f \in \lambda^{\alpha,\beta}_{\log}(x_0),\ n \leq \alpha < (n+1) \Leftrightarrow f_{loc}(x) = \mathbf{o}(x^\alpha \log^\beta x),\ (x \to 0).$
- $\Lambda^\alpha(\mathbb{R})$, $\lambda^\alpha(\mathbb{R})$, $\Lambda^{\alpha,\beta}_{\log}(\mathbb{R})$ *and* $\lambda^{\alpha,\beta}_{\log}(\mathbb{R})$ *are the spaces where the above estimates hold uniformly in* x*.*

Lemma 2.15. *Let* $\alpha < \alpha',\ \beta < \beta'$ *and* $\gamma > 0$. *Then*
$\Lambda^{\alpha,\beta}_{\log}(\mathbb{R}) \subset \Lambda^{\alpha,\beta'}_{\log}(\mathbb{R}) \subset \Lambda^{\alpha}(\mathbb{R}) \subset \lambda^{\alpha,\beta}_{\log}(\mathbb{R}) \subset \lambda^{\alpha,\beta'}_{\log}(\mathbb{R}) \subset \lambda^{\alpha}(\mathbb{R}) \subset \Lambda^{\alpha',\gamma}_{\log}(\mathbb{R})$.

Lemma 2.16. *We have the following results, concerning Hölder regularity:*

1. f *Hölder* α*,* $\beta \le \alpha$*,* $\Rightarrow f$ *Hölder* β*.*
2. f *bounded,* $\alpha \le 0 \Rightarrow f$ *Hölder* α*.*
3. f *continuous,* $\alpha < 1 \Rightarrow f$ *Hölder* α*.*
4. f *Hölder* $\alpha \Rightarrow F(x) = \int_a^x f(u)du$ *Hölder* $(\alpha + 1)$*.*
5. f *Hölder* $\alpha \notin \mathbb{Z} \Rightarrow f'$ *Hölder* $(\alpha - 1)$*.*

Proof. 1. $\alpha - \beta > 0,\ |h| < 1 \Rightarrow |h|^{\alpha-\beta} < 1$. So
$|f(x+h) - P_n(h)| \le A|h|^{\alpha} = A|h|^{\beta}\ |h|^{\alpha-\beta} \le A|h|^{\beta}$.

2. $|h| < 1,\ \alpha \le 0 \Rightarrow |h|^{\alpha} \ge 1$. So
$|f(x+h) - P_n(h)| = |f(x+h) - f(x)| \le |f(x+h)| + |f(x)| \le 2A \le 2A|h|^{\alpha}$.

3. $|h| < 1,\ \alpha < 1 \Rightarrow |h|^{\alpha} > |h|$. So
$|f(x+h) - f(x)| \le A|h| \le A|h|^{\alpha}$.

4. Select $\tau \in (-h, h)$ such that $\frac{F(x+h)-F(x)}{h} = f'(x+\tau)$. (It exists by the *Mean Value Theorem*.) Then

$$\begin{aligned}
|F(x+h) - P_n(h)| &= |F(x+h) - F(x) - f(x)h \\
&\quad - \sum_{k=1}^{n} \frac{1}{(k+1)!} f^{(k)}(x)h^{k+1}| \\
&= |h|\ |(\frac{F(x+h)-F(x)}{h} \\
&\quad - \sum_{k=0}^{n} \frac{1}{(k+1)!} f^{(k)}(x)h^{k})| \\
&= |h|(|f(x+\tau) - \sum_{k=0}^{n} \frac{1}{(k+1)!} f^{(k)}(x)h^{k})| \\
&= |h|(|f(x+\tau) - \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x)h^{k} \\
&\quad + \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x)h^{k} - \sum_{k=0}^{n} \frac{1}{(k+1)!} f^{(k)}(x)h^{k}|) \\
&\le |h|\ C|h|^{\alpha} + |h|\ |\sum_{k=0}^{n} \frac{k}{(k+1)!} f^{(k)}(x)h^{k})| \\
&\le |h|\ C|h|^{\alpha} + |h|\ C_1|h|^{\alpha} \\
&= |h|\ C_2|h|^{\alpha} \\
&= C_2|h|^{(\alpha+1)},
\end{aligned}$$

for h small enought, since the polynomials are smooth.

5.

$$\begin{aligned} C|h|^{\alpha} &\geq |f(x+h) - P_n(h)| \\ &= |f(x+h) - f(x) - \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(x) h^k \\ &= |h| \, (|\frac{f(x+h) - f(x)}{h} - \sum_{k=0}^{n-1} \frac{1}{(k+1)!} f^{(k+1)}(x) h^k|) \\ &= |h| \, |f'(x+\tau) - P_2(h)| \end{aligned}$$

Then

$|f'(x+\tau) - P_2(h)| \leq C|h|^{(\alpha-1)}$,

by the same argument. The $\alpha \notin \mathbb{N}$ reservation comes from the case of oscillating singularities.

□

To measure regularity of functions, the smoothness, or the ‘narrowness’ in the frequency domain, of the analysing wavelet is not that important. But to measure Hölder regularity higher than 1, vanishing moments of the wavelet is crucial.

Definition 2.17 (Vanishing Moments, $M_n(\Omega)$). *A function ψ has $n \in \mathbb{N}$* vanishing moments *on $\Omega \subset \mathbb{R}$ if*

$$\int_{\Omega} x^k \psi(x) \, dx = 0 \text{ for } k = 0, \cdots, n.$$

$$M_n(\Omega) = \{f \in L^1(\Omega) : f \text{ has } n \text{ vanishing moments.}\}$$

By the Hölder property (2.12), we approximate f with a polynomial $P_n(x)$ in a neighborhood of x,

$$f(x_0 + h) = P_n(h) + g(h), \text{ with } |g(h) \leq A|h|^{\alpha}. \tag{2.13}$$

If the analysing wavelet has n vanishing moments, the polynomial part $P_n(x)$ of $f(x)$ gives no contribution to $\mathcal{W}_\psi f(s,x)$. So to find the exact Hölder regularity α, when $m-1 < \alpha < m$ for some large $m \in \mathbb{N}$, simply use a wavelet with at least m vanishing moments, and the smooth, polynomial part of f will not affect $\mathcal{W}_\psi f(s,x)$.

Theorem 2.18 (Mallat (37) (Theorem 6.2)). *Given a wavelet ψ with fast decay.*

$$\psi \in M_n(\mathbb{R}) \Leftrightarrow \exists \theta \text{ with fast decay, such that } \psi(x) = (-1)^n \frac{d^n \theta(x)}{dx^n}.$$

Further,

$$m > n, \ \psi \notin M_m(\mathbb{R}) \Leftrightarrow \int_{\mathbb{R}} \theta(x) \, dx \neq 0.$$

Proof. This is Theorem 6.2 in Mallat (37) and is proved there. □

Corollary 2.19. *If* $\psi \in M_n(\mathbb{R})$ *, then*

$$\mathcal{W}_\psi f(s,x) = s^n \frac{d^n}{dx^n}(f * \theta_s^{inv})(x),$$

where $\theta_s^{inv}(x) = \frac{1}{s}\theta(\frac{-t}{s})$.

Proof. This is included in Theorem 6.2 in Mallat (37) and is proved there. □

Corollary 2.20. *If* $\psi \in M_n(\mathbb{R})$ *then*

$$\hat{\psi}(\xi) = (i\xi)^n \hat{\theta}(\xi).$$

Proof. This is a simple consequence of the previous lemma and of Lemma A.60. □

It is, however, important to know that there exists wavelets with *all* the moments vanishing, as described in this next lemma. This is the case for the *Meyer Wavelet*, introduced by Yves Meyer: Wavelets and Operators (45).

Lemma 2.21 (Louis et. al (31)(Lemma 1.4.5)). *There exists admissible* $\psi \in \mathcal{S}(\mathbb{R})$ *such that*

$$\int_{\mathbb{R}} x^k \psi(x)\, dx = 0, \forall k \in \mathbb{N}.$$

Proof. Choose $0 \neq \phi \in \mathcal{C}_0^\infty(\mathbb{R})$ such that $\phi(\xi) = 0$ in a neighbourhood of $\xi = 0$. Then $\phi^{(k)}(\xi)|_{\xi=0} = 0,\ \forall k \in \mathbb{N}$. Define $\psi(x)$ as the inverse Fourier Transform of $\phi(\xi)$. Then

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

so ψ is admissible. Let $g(x) = x^k\psi(x)$. Then

$$\begin{aligned}
0 = \phi^{(k)}(\xi)|_{\xi=0} &= \hat{\psi}^{(k)}(\xi)|_{\xi=0} \\
&= (-1)^k \hat{g}(\xi)|_{\xi=0} \\
&= (-1)^k \int_{\mathbb{R}} g(x) e^{-ix\xi}\, dx|_{\xi=0} \\
&= (-1)^k \int_{\mathbb{R}} x^k \psi(x) e^{-ix\xi}\, dx|_{\xi=0} \\
&= (-1)^k \int_{\mathbb{R}} x^k \psi(x)\, dx.
\end{aligned}$$

□

Now we have established some notations and properties of wavelets, and are ready to examine some consequences of functions being Hölder regular. The first two theorems are regarding uniform Hölder regularity and the next two are about pointwise Hölder regularity.

Theorem 2.22 (Mallat, Hwang (38) (Theorem 3.3 a)). *Let* $0 < \alpha < n \in \mathbb{N}$. *Let* $[a,b] \subset \mathbb{R}$ *be an interval and* $(b-a) > 2\epsilon > 0$. *Suppose that* $\psi \in M_n(\mathbb{R})$ *is a wavelet and* $||x^\alpha \psi||_{L^1(\mathbb{R})} < \infty$. *If a function* $f(x) \in L^2(\mathbb{R})$ *is uniformly Hölder* α *over any interval* $(a+\epsilon, b-\epsilon)$, *then*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha),\ x \in (a+\epsilon, b-\epsilon),\ s > 0. \tag{2.14}$$

Proof. This is proved in Chapter 3. □

Theorem 2.23 (Mallat (37) (Theorem 6.3 a)). *Let* $\psi \in M_n(\mathbb{R})$ *have* n *derivatives having fast decay. If* $f \in L^2(\mathbb{R})$ *is uniformly Hölder* $\alpha \leq n$ *over* $[a,b] \subset \mathbb{R}$ *then*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha),\ \forall (s,x) \in \mathbb{R}^+ \times [a,b]. \tag{2.15}$$

Proof. This is half of Theorem 6.3 in Mallat (37) and is proved there. □

We see that the differences between Theorem 2.22 and the more recent by Mallat in 2.23 is that with the stronger condition that ψ have n derivatives having fast decay, instead of $||x^\alpha \psi||_{L^1(\mathbb{R})} < \infty$, we get the result on the whole interval, $[a,b]$, and also that the result is valid for integer $\alpha = n \in \mathbb{N}$. Now to the pointwise cases:

Theorem 2.24 (Mallat, Hwang (38) (Theorem 3.4 a)). *Let* $\alpha \leq n \in \mathbb{N}$. *Suppose* $\psi \in C^n(\mathbb{R}) \cap M_n(\mathbb{R})$ *is a wavelet, such that* $||x^\alpha \psi||_{L^1(\mathbb{R})} < \infty$. *If a function* $f(x) \in L^2(\mathbb{R})$ *is Hölder* α *at* x_0, *then for all points* x *in a neighborhood of* x_0 *and any scale* s,

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha + |x-x_0|^\alpha). \tag{2.16}$$

Proof. This is proved in Chapter 3. □

Theorem 2.25 (Mallat (37) (Theorem 6.4 a)). *Let* $\psi \in M_n(\mathbb{R})$ *have* n *derivatives having fast decay. If* $f \in L^2(\mathbb{R})$ *is Hölder* $\alpha \leq n$ *at* x_0 *then*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha + |x-x_0|^\alpha),\ \forall (s,x) \in \mathbb{R}^+ \times \mathbb{R}.$$

Proof. This is half of Theorem 6.4 in Mallat (37) and is proved there. □

Again, the condition that ψ have n derivatives having fast decay in Theorem 2.25 replaces the condition $||x^\alpha \psi||_{L^1(\mathbb{R})} < \infty$ in Theorem 2.22, with the same conclusion, making the first one the strongest.

As a corollary of this, we get the following 'algorithm' for computing the Hölder exponent, with notation $\alpha(f, x_0)$:

Corollary 2.26 (Mallat (37) (Theorem 10.2)). *Suppose* f *uniformly Hölder* β *and bounded for some* β, $0 < \beta < 1$. *Then, for every* $x_0 \in \mathbb{R}$, *the* Hölder exponent $\alpha(f, x_0)$ *is given by*

$$\alpha(f, x_0) = \lim_{s \to 0 x \to x_0} \inf \frac{\log |\mathcal{W}_\psi f(s,x)|}{\log(s + |x - x_0|)} \tag{2.17}$$

Proof. This is Theorem 10.2 in Mallat (37) and is proved there. □

Definition 2.27 (Submultiplicative Function). *A function* $r : \mathbb{R}_+ \to \mathbb{R}_+$ *is called* submultiplicative *if:*

- $r(x + \epsilon) \geq r(x), \; \forall x, \epsilon > 0.$
- $\exists C > 0$ *such that* $r(t\,x) \leq C\; r(t)\; r(x), \; \forall t, x \in \mathbb{R}_+.$

A function r *is submultiplicative over* $\mathbb{R}$ *if* $r(x)$ *and* $r(-x), \; x \geq 0$ *are submultiplicative.*

A submultiplicative function is an increasing function, with a restriction on the asymptotic increase. For instance, the functions $r(x) = C\; x^{\alpha}$ are submultiplicative functions for all $\alpha, \; C \; > \; 0$. A function $r(x)$ being submultiplicative means that the logarithm, $p(x) = \log(r(x))$ is *subadditive.* This notation is used by Holschneider (14) as a generalization of the x^{α}-perspective of Mallat (37), and Mallat and Hwang (38) in Theorem 2.22 and Theorem 2.24 and in theorems we will be studying in later sections.

Theorem 2.28 (Holschneider (14) (Theorem 2.0.5)). *Let* r *be a submultiplicative function and let* $\psi \in M_n(\mathbb{R})$ *be a wavelet such that* $||r\psi||_{L^1(\mathbb{R})} < \infty$. *Then*

- $f(x_0 + x) = P_n(x) + \mathbf{O}(r(|x|)) \Rightarrow \mathcal{W}_{\psi} f(s, x_0 + x) = \mathbf{O}(r(x) + r(s)).$
- $f(x_0 + x) = P_n(x) + \mathbf{o}(r(|x|)) \Rightarrow \mathcal{W}_{\psi} f(s, x_0 + x) = \mathbf{o}(r(x) + r(s)).$

Proof. This is Theorem 2.0.5 in Holschneider (14) and is proved there. □

Corollary 2.29 (Holschneider (14) (Theorem 2.0.3)). *Let* $\psi \in M_n(\mathbb{R})$ *be a wavelet and let* $|f(x)| \leq c(1 + |x|^{\alpha})$.

- *If* $||x^{\alpha}\psi||_{L^1(\mathbb{R})} < \infty$, *then*
  - $f \in \Lambda^{\alpha}(x_0) \Rightarrow \mathcal{W}_{\psi} f(s, x_0 + x) = \mathbf{O}(s^{\alpha} + |x|^{\alpha}), \; (s \to 0).$
  - $f \in \lambda^{\alpha}(x_0) \Rightarrow \mathcal{W}_{\psi} f(s, x_0 + x) = \mathbf{o}(s^{\alpha} + |x|^{\alpha}), \; (s \to 0).$
- *If* $||x^{\alpha} \log^{\beta}(x)\psi(x)||_{L^1(\mathbb{R})} < \infty$, *then*
  - $f \in \Lambda^{\alpha,\beta}_{\log} \Rightarrow \mathcal{W}_f(s, x_0 + x) = \mathbf{O}(s^{\alpha} \log^{\beta} s + |x|^{\alpha} \log^{\beta} |x|), \; (s \to 0).$
  - $f \in \lambda^{\alpha,\beta}_{\log} \Rightarrow \mathcal{W}_f(s, x_0 + x) = \mathbf{o}(s^{\alpha} \log^{\beta} s + |x|^{\alpha} \log^{\beta} |x|), \; (s \to 0).$

Proof. This corollary is Theorem 2.0.3 in Holschneider (14) and is proved there. □

A slightly different class of functions from the Hölder-classes of functions, or specific the Hölder $\alpha = 1$ class is the *Class of Zygmund*:

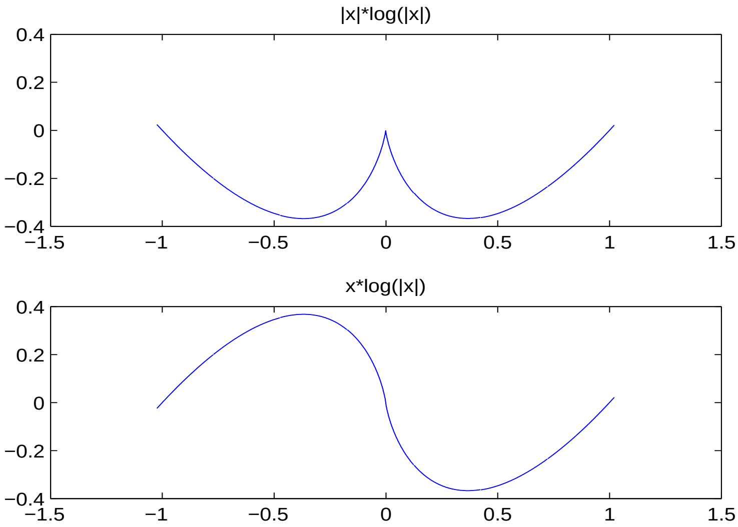

Figure 2.2   Functions in the Zygmund Class does not have cusps

Definition 2.30 (The Class of Zygmund, $\Lambda^*(\mathbb{R})$, $\lambda^*(\mathbb{R})$).

$\Lambda^*(\mathbb{R}) = \{f \in C(\mathbb{R}) : |f(x + x_0) + f(x - x_0) - 2f(x_0)| = \mathbf{O}(x), x \to 0, \forall x_0 \in \mathbb{R}\}.$

$\lambda^*(\mathbb{R}) = \{f \in C(\mathbb{R}) :\ |f(x + x_0) + f(x - x_0) - 2f(x_0)| = \mathbf{o}(x),\ x \to 0,\ \forall x_0 \in \mathbb{R}\}.$

The functions in $\Lambda^*(\mathbb{R})$ do not have cusps. The example illustrating this in Holschneider (14) is the function

$f(x) = x\ \log(|x|) \in \Lambda^*(\mathbb{R}),$

and

$f(x) = |x|\ \log(|x|) \notin \Lambda^*(\mathbb{R}),$

plotted in Figure 2.2.

Theorem 2.31 (Holschneider (14) (Theorem 2.2.2)). *Suppose $\psi \in S_0(\mathbb{R})$ (Schwarz class, page 92) is a wavelet. Then,*

- $f \in \Lambda^*(\mathbb{R}) \Rightarrow \mathcal{W}_\psi f(s, x) = \mathbf{O}(s),\ (s \to 0).$
- $f \in \lambda^*(\mathbb{R}) \Rightarrow \mathcal{W}_\psi f(s, x) = \mathbf{o}(s),\ (s \to 0).$

Proof. This is Theorem 2.2.2 in Holschneider (14) and is proved there. $\square$

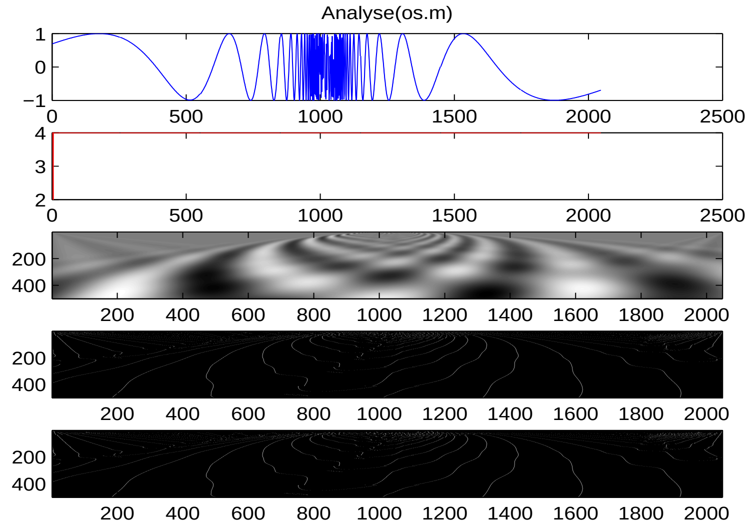

Figure 2.3 A function with oscillating singularity on top. The other plots are a regularity analysis, the magnitude of the wavelet transform, and two versions of the maxima lines og this CWT. This is a type of plots that will be used later in this thesis.

2.4 Oscillating Singularities

Oscillating singularities occur when an otherwise (in a neighbourhood) smooth function is singular at a point due to oscillations that tends to infinity towards that point. That means that in a neighborhood that does not contain the singular point, the function is smooth, but it becomes arbitrary fast changing close to the point. We define oscillating singularities by properties of the wavelet transform, according to the themes in this thesis:

Definition 2.32 (Oscillating Singularities). *A function $f(x)$ has an* Oscillating singularity *at x_0 if there exists $\alpha > 0$ such that f is not Hölder α at x_0 but the primitive $F(x) = \int_a^x f(u)\, du$ is Hölder $(\alpha + 1)$ at x_0.*

For instance, the function

$$f(x) = \sin(\frac{1}{x})$$

has an isolated singularity $\alpha = 0$, at $x = 0$ and is C^∞ elsewhere.

For functions with oscillating singularities we must consider the wavelet transform *outside* The Cone of Influence, as in the second term $\frac{|x-x_0|^\alpha}{|\log(|x-x_0|)|}$ in (2.21) of Theorem 2.39. The reason for this is that we don't have any maxima lines that converges to the singular point, but we have maxima lines converging to $x = 1/(n\pi)$ with

$$|\mathcal{W}_\psi f(s,x)| \leq A_n s, (\alpha = 1),$$

where $A_n = \mathbf{O}(n^2)$ which means that the A_ns grow to infinity when we get closer to 0.

The maxima of the maxima lines, meaning the points $(s, X(s))$ on one of the maxima lines where we have a local maximum of $|\mathcal{W}_\psi f(s,x)|$ in a small 2D neighborhood are located along, for the $\sin(1/x)$-case, a parabola that is outside the Cone of Influence. A more thorough discussion of this is found in Section 5.3 on pages 70-78 in Mallat and Hwang (39).

2.5 Global Hölder Regularity

With the Fourier transform, we are able to characterize global Hölder regularity by the decay of the Fourier transform as shown in Lemma 2.33. We are, however, not able to tell whether the function is locally more regular.

Lemma 2.33 (Mallat (37) (Theorem 6.1)). *A bounded function $f(x)$ is uniformly Hölder α over $\mathbb{R}$ if it satisfies $\int_{\mathbb{R}} |\hat{f}(\xi)|(1+|\xi|)^\alpha \, d\xi < +\infty$.*

Proof. This is Theorem 6.1 in Mallat (37) and is proved there. □

With the wavelet transform, however, we will in the next chapter show that we can find the Hölder regularity of a given function at a particular point x_0. This section will show some results regarding characterization of Hölder regularity on regions or neighborhoods in $\mathbb{R}$ (or all of $\mathbb{R}$) by the asymptotic behavior of the wavelet transform when the scale goes to zero.

Theorem 2.34 (Mallat, Hwang (38) (Theorem 3.3 b)). *Suppose that $0 < \alpha \leq n \in \mathbb{N}$, $\alpha \notin \mathbb{N}$, $\psi \in C(\mathbb{R}) \cap M_n(\mathbb{R})$ is a wavelet and $||\psi'||_{L^1(\mathbb{R})} < \infty$. Let $[a,b] \subset \mathbb{R}$ be an interval. If*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha). \tag{2.18}$$

for any $x \in (a+\epsilon, b-\epsilon)$ $((b-a) > 2\epsilon > 0)$ and any scale $s > 0$, then $f(x)$ is uniformly Hölder α over any such interval $(a+\epsilon, b-\epsilon)$.

Proof. This is half of Theorem 3.3 in Mallat, Hwang (38) and is proved in Chapter 3. □

Theorem 2.35 (Mallat (37) (Theorem 6.3 b)). *Let $\psi \in M_n(\mathbb{R})$ have n derivatives having fast decay. If $f \in L^2(\mathbb{R})$, $[a,b] \subset \mathbb{R}$, $n < \alpha \notin \mathbb{N}$ and $\mathcal{W}_\psi f(s,x)$ satisfies*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha),\ \forall (s,x) \in \mathbb{R}^+ \times [a,b] \tag{2.19}$$

then f is uniformly Hölder α on $[a+\epsilon, b-\epsilon]$, $\forall \epsilon > 0$.

Proof. This is the second part of Theorem 6.3 in Mallat (37) and is proved there. □

As for the theorems in the previous section, the main difference between Theorem 2.34 and Theorem 2.35 is that in the first $\psi \in C(\mathbb{R})$ and $\psi' \in L^1(\mathbb{R})$ whilst in the latter, ψ is supposed to be $C^n(\mathbb{R})$ with n derivatives having fast decay. The important difference between these two theorems and their counterparts, Theorem 2.22 and Theorem 2.23, is the non-integer demand on α ($\alpha \notin \mathbb{N}$) .

Theorem 2.36 (Holschneider (14) (Theorem 2.1.1)). *Suppose* $0 < \alpha < n \in \mathbb{N}$, $\psi \in C^{n+1}(\mathbb{R})$ *is compactly supported and* $|\mathcal{W}_\psi f(s,x)|$ *is rapidly decreasing for large* s. *Then, when* $a \to 0$*:*

- $|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha) \Rightarrow f \in \Lambda^\alpha(\mathbb{R})$
- $|\mathcal{W}_\psi f(s,x)| = \mathbf{o}(s^\alpha) \Rightarrow f \in \lambda^\alpha(\mathbb{R})$
- $|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha \log^\beta(s)) \Rightarrow f \in \Lambda^{\alpha,\beta}_{log}(\mathbb{R})$
- $|\mathcal{W}_\psi f(s,x)| = \mathbf{o}(s^\alpha \log^\beta(s)) \Rightarrow f \in \lambda^{\alpha,\beta}_{log}(\mathbb{R})$

Proof. This is Theorem 2.1.1 in Holschneider (14) and is proved there. □

Theorem 2.37 (Holschneider (14) (Theorem 2.2.3)). *Let* $\psi \in C^2(\mathbb{R})$ *be compactly supported and* $\mathcal{W}_\psi f(s,x) = 0$ *for* $s > 1$.

- $\mathcal{W}_\psi f(s,x) = \mathbf{O}(s) \Rightarrow f \in \Lambda^*(\mathbb{R})$.
- $\mathcal{W}_\psi f(s,x) = \mathbf{o}(s) \Rightarrow f \in \lambda^*(\mathbb{R})$.

Proof. This is Theorem 2.2.3 in Holschneider (14) and is proved there. □

Corollary 2.38 (Holschneider (14) (Corollary of Theorem 2.2.3)). *Let* $\alpha < 1$. *Then,*

$\Lambda^1(\mathbb{R}) \subset \Lambda^*(\mathbb{R}) \subset \Lambda^\alpha(\mathbb{R})$.

Proof. This is a corollary of Theorem 2.2.3 in Holschneider (14) and is proved there. □

2.6 Local Hölder Regularity

We now turn to the more important part of this chapter. We will present some theorems that gives us a tool for characterizing the pointwise regularity of functions based on the decay of the wavelet transform when the scale goes to zero.

Theorem 2.39 (Mallat, Hwang (38) (Theorem 3.4 b)). *Let* $f \in L^2(\mathbb{R})$ *and* $\psi(x) \in C^n(a,b) \cap M_n(\mathbb{R})$ *be a wavelet with compact support. Let* $0 < \alpha < n,\ \alpha \notin \mathbb{N}$. *If the two following conditions hold:*

- *There exists* $\epsilon > 0$ *such that for all points* x *in a neighborhood of* x_0 *and any scale* s,

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\epsilon). \tag{2.20}$$

- *For all points* x *in a neighborhood of* x_0 *and any scale* s

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}\left(s^\alpha + \frac{|x-x_0|^\alpha}{|\log(|x-x_0|)|}\right). \tag{2.21}$$

then $f(x)$ is Hölder α at x_0

Proof. This is half of Theorem 3.4 in Mallat, Hwang (38) and is proved in Chapter 3. □

Theorem 2.40 (Mallat (37) (Theorem 6.4 b)). *Let $\psi \in M_n(\mathbb{R})$ have n derivatives having fast decay. If $n > \alpha \notin \mathbb{N}$ and there exists $\alpha' < \alpha$ such that*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha + s^{(\alpha-\alpha')}|x - x_0|^{\alpha'}),\ \forall (s,x) \in \mathbb{R}^+ \times \mathbb{R}, \tag{2.22}$$

then f is Hölder α at x_0.

Proof. This is half of Theorem 6.4 in Mallat (37) and is proved there. □

The first condition in Theorem 2.39 shows that f is uniformly Hölder ϵ (typically $\epsilon < \alpha$) in a region containing x_0. Theorem 2.40 has a slightly different condition:

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha + s^{(\alpha-\alpha')}|x - x_0|^{\alpha'})$$

which is supposed to be valid globally and makes it almost an if-and-only-if theorem, together with Theorem 2.25. The other difference between these two theorems is the compactly supported ψ in the first as opposed to the fast decay of the n derivatives of ψ in the latter. The second term in (2.21) represents a restriction on $\mathcal{W}_\psi f(s,x)$ also *outside* The Cone of Influence.

Theorem 2.41 (Holschneider (14) (Theorem 2.3.2)). *Let r be a submultiplicative, even function, $n \in \mathbb{N}$ and let $\psi \in C^{(n+1)}$ be compactly supported wavelet. Suppose*

1. *$\int_0^1 r(x)x^{-(n+1)}\, dx < \infty$ and $r(x) = \mathbf{O}(x^n),\ (x \to 0)$.*
2. *$\int_1^\infty r(x)x^{-(n+2)}\, dx < \infty$ and $r(x) = \mathbf{O}(x^{-(n+1)}),\ (x \to \infty)$.*
3. *$\exists \gamma > 0$ such that $|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\gamma)$ for $s < 1$ uniformly in x.*

Then, for $(s \to 0, x)$,

- $\mathcal{W}_\psi f(s, x_0 + x) = \mathbf{O}(r(s) + \frac{r(x)}{\log(r(x))}) \Rightarrow |f(x + x_0) - P_n(x)| = \mathbf{O}(r(x))$.
- $\mathcal{W}_\psi f(s, x_0 + x) = \mathbf{o}(r(s) + \frac{r(x)}{\log(r(x))}) \Rightarrow |f(x + x_0) - P_n(x)| = \mathbf{o}(r(x))$.

Proof. This is Theorem 2.3.2 in Holschneider (14) and is proved there. □

Corollary 2.42 (Holschneider (14) (Corollary 2.3.3)). *If*

- $|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\gamma)$ *for some* $\gamma > 0$.
- $|\mathcal{W}_\psi f(s, x_0 + x)| = \mathbf{O}(s^\alpha + x^\alpha),\ (s \to 0, x)$,

then

$|f(x+x_0) - f(x_0)| = \mathbf{O}(x^\alpha \log(x)),\ (x \to 0).$

Proof. This is Corollary 2.3.3 in Holschneider (14) and is proved there. □

Theorem 2.43 (Holschneider (14) (Theorem 2.3.1)). *Suppose*

1. $\exists \gamma > 0$ *such that* $|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\gamma)$ *for* $s < 1$ *uniformly in* x.
2. $|\mathcal{W}_\psi f(s, x_0 + x)| = \mathbf{O}(s^\alpha) + \mathbf{O}(\frac{x^\alpha}{\log x}),\ (s \to 0, x)$,
3. $|\mathcal{W}_\psi f(s,x)|$ *is rapidly decreasing for large* s.

Then

$$|f(x+x_0) - P_n(x)| = \mathbf{O}(x^\alpha)$$

for $n < \alpha < n+1$.

Proof. This is Theorem 2.3.1 in Holschneider (14) and is proved there. □

2.7 Pointwise Differentiability

This section is included to show how the results concerning Hölder regularity affects the differentiability of functions, and only Holschneider (14) has these results explicitly as theorems.

Theorem 2.44 (Holschneider (14) (Theorem 2.1.2)). *Let* r *be a submultiplicative, even function that satisfies, for* $n \in \mathbb{N}$,

- $\int_0^1 r(x) x^{-(n+1)}\ dx < \infty$ *and* $r(x) = \mathbf{O}(x^n),\ (x \to 0)$.
- $\int_1^\infty r(x) x^{-(n+2)}\ dx < \infty$ *and* $r(x) = \mathbf{O}(x^{-(n+1)}),\ (x \to \infty)$.

Suppose $\psi \in C^{(n+1)}$ *has compact support and that* $\mathcal{W}_\psi f(s,x) = 0$ *for* $s > 1$.

- *If* $|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(r(s)),\ (s \to 0)$ *then*
  - $f \in C^n(\mathbb{R})$,
  - $|\frac{d^n f(x+u)}{dx^n} - \frac{d^n f(x)}{dx^n}| = \mathbf{O}(r(x)/x^n), (x \to 0)$.
- *If* $|\mathcal{W}_\psi f(s,x)| = \mathbf{o}(r(s)),\ (s \to 0)$ *then*
  - $f \in C^n(\mathbb{R})$,
  - $|\frac{d^n f(x+u)}{dx^n} - \frac{d^n f(x)}{dx^n}| = \mathbf{o}(r(x)/x^n), (x \to 0)$.

Proof. This is Theorem 2.1.2 in Holschneider (14) and is proved there. □

Theorem 2.45 (Holschneider (14) (Theorem 2.4.1)). *Suppose the wavelet $\psi \in C^{(n+1)}$ has compact support and a non-negative monotonic function $r(x) = r(|x|)$ satisfies the ''Condition of Dini":*

$$\int_0^1 \frac{r(x)}{x^{(n+2)}}\, dx < \infty.$$

If

1. $\mathcal{W}_\psi f(s,x) = 0$ *for* $s > 1$,
2. $\mathcal{W}_\psi f(s,x) = \mathbf{O}(s^\gamma)$ *for some* $\gamma > 0$ *and*
3. $\mathcal{W}_\psi f(s, x + x_0) = \mathbf{O}(r(s) + r(x))$,

then the nth differential quotient of f ($\Delta^n(x) = \Delta^{(n-1)}(\frac{f(x-x_0)-f(x_0)}{x})$) exists at x_0. Furthermore, the condition on r is optimal.

Proof. This is Theorem 2.4.1 in Holschneider (14) and is proved there. □

Theorem 2.46 (Holschneider (14) (Theorem 2.4.2)). *Suppose f is a periodic function or measure, $\partial_t f(x_0)$ exists and the wavelet $\psi \in L^1(\mathbb{R}) \cap M_{n-1}(\mathbb{R})$ and $(x^n\psi) \in L^1(\mathbb{R})$. If*

$$\int_{-\infty}^{\infty} x^n \psi(x)\, dx = n! \Leftrightarrow (i\partial)^n \hat{\psi}(0) = 2\pi n!,$$

then

$$\lim_{s\to 0} \frac{\mathcal{W}_\psi f(s,x_0)}{s^n} = \partial_x^n f(x_0).$$

Proof. This is Theorem 2.4.2 in Holschneider (14) and is proved there. □

2.8 CWT Local Maxima

The continuous wavelet transform is a function of 2 variables, the 'time' and the 'frequency' variable. We are interested in the set of maxima of the one-dimensional functions we get when we fix the 'frequency' variable. The maxima of the functions $g_s(x) = |\mathcal{W}_\psi f(s,x)|$ are on the *ridges* of the $|\mathcal{W}_\psi f(s,x)|$ surface, whereas the *zero-crossings* of the functions $g_s(x) = \mathcal{W}_\psi f(s,x)$ are the delimiting lines between the different such ridges, or the 'valleys' between them. Properties of $\mathcal{W}_\psi f(s,x)$ at the zero-crossings or at the ridges has been studied extensively in (1), (36),(55) and in many other papers.

Definition 2.47 (Maxima, Local Maxima, Modulus Maxima, Local Modulus Maxima). *We say that a wavelet transform $\mathcal{W}_\psi f(s,x)$ has a* maximum *(plural;* maxima *or one of the more precise;* Local Maxima, Modulus Maxima *or* Local Modulus Maxima*) at $(s_0, x_0) \in \mathbb{R}^2$ if the function $g(x) = |\mathcal{W}_\psi f(s_0, x)|$ has a local maximum at x_0, strictly on one of the sides (left or right).*

Definition 2.48 (The set of maxima of the wavelet transform). *Let* $g(x) = |\mathcal{W}_\psi f(s_0, x)|$. *We define the* set of maxima *of the wavelet transform* $\mathcal{W}_\psi f(s, x)$ *by*

$$\mathrm{Max}(\mathcal{W}_\psi f(s,x)) = \{(s_0, x_0) \in \mathbb{R}^2 \mid g(x) \textit{ has a local maximum at } x_0\}$$

This next Proposition is used in the proof of Theorem 2.50 by induction on n.

Proposition 2.49 (Mallat, Hwang (38) (Proposition A.1)). *Let* $n \in \mathbb{N}$ *and* ψ *be a wavelet that can be written* $\psi(x) = \frac{d^n \phi(x)}{dx^n}$, *where* $\phi(x)$ *is a continuous function with compact support. Let* $f(x)$ *be a function and suppose that for any* $\epsilon > 0$, *there exists a constant* K_ϵ, *such that at all scales* s,

$$\int_{a+\epsilon}^{b-\epsilon} |f * \phi_s(x)| \, dx \leq K_\epsilon. \tag{2.23}$$

If $|\mathcal{W}_\psi f(s,x)|$ *has no local maxima for* $x \in (a,b)$ *and* $s < s_0$, *for some* $s_0 > 0$, *then for any* $\epsilon > 0$, *there exists a constant* $A_{\epsilon,n}$ *such that for any* $x \in (a+\epsilon, b-\epsilon)$ *and* $s < s_0$

$$|\mathcal{W}_\psi f(s,x)| \leq A_{\epsilon,n} s^n. \tag{2.24}$$

Proof. This is Proposition A.1 in the appendix of Mallat and Hwang (38) and is proved in Chapter 3. □

Theorem 2.50 (Mallat, Hwang (38) (Theorem 5.2)). *Let* $n \in \mathbb{N}$ *and* $\psi \in C^n(a,b) \cap M_n(a,b)$ *be a wavelet with compact support. If there exists a scale* $s_0 > 0$ *such that for all scales* $s < s_0$ *and* $x \in (a,b)$, $|\mathcal{W}_\psi f(s,x)|$ *has no local maxima, then for any* $\epsilon > 0$ *and* $\alpha < n$, f *is uniformly Hölder* α *on* $(a+\epsilon, b-\epsilon)$.
If $\psi = \frac{d^n}{dx^n}\Theta(x)$ *where* $\Theta(x)$ *is a* smoothing function, *then* f *is uniformly Hölder* n *on any such interval* $(a+\epsilon, b-\epsilon)$.

Proof. This is Theorem 5.2 in Mallat and Hwang (38) and is proved in Chapter 3. □

Corollary 2.51 (Mallat, Hwang (38) (Corollary of Theorem 5.2)). *The closure of the set of points where* f *is not Hölder* n *is included in the closure of the wavelet transform maxima of* f.

Proof. This is a simple consequence of Theorem 2.50. □

We see that given the conditions in Theorem 2.50, all the irregular points (not Hölder $\alpha = n$) can be located by following the maxima lines when the scale goes to zero.

Theorem 2.52 (Mallat (37) (Theorem 6.5)). *Let* $n \in \mathbb{N}$, $f \in L^1[a,b]$ *and* ψ *be a compactly supported wavelet such that*

$$\psi(x) = (-1)^n \frac{d^n}{dx^n}\theta(x) \in C^n(\mathbb{R})$$

where $\int_{\mathbb{R}} \theta(x)\, dx \neq 0$. *If there exists* $s_0 > 0$ *such that* $|\mathcal{W}_\psi f(s,x)|$ *has no local maxima for* $x \in [a,b]$ *and* $s < s_0$, *then* f *is uniformly Hölder* n *on* $[a+\epsilon, b-\epsilon]$ *for any* $\epsilon > 0$.

Proof. This is Theorem 6.5 in Mallat (37) and is proved there. □

Here we see that the main differences between Theorem 2.50 and the more recent Theorem 2.52 is the $\psi \in C^n(\mathbb{R})$ condition in the first gives an additional result for $\alpha < n$ as opposed to the condition that ψ is the n-th derivative of a smoothing function $\psi(x) = (-1)^n \frac{d^n}{dx^n}\theta(x)$. This last condition gives us the result for $\alpha = n$ in both theorems.

This next theorem shows us that with strict conditions on the wavelet used, the wavelet transform modulus maxima lines are indeed *lines*, i.e. they are connected and they continue all the way down to the finest scale.

Proposition 2.53 (Mallat (37) (Proposition 6.1)). *Let θ be a Gaussian and $\psi(x) = (-1)^n\theta^{(n)}(x)$. For any $f \in L^2(\mathbb{R})$, the modulus maxima of $\mathcal{W}_\psi f(s,x)$ belong to connected curves that are never interrupted when the scale decreases.*

Proof. This is Proposition 6.1 in Mallat (37) and is proved there. □

Theorem 2.54 (Mallat, Hwang (38) (Theorem 5.3)). *Let $\psi \in C^n([a,b])$ be compactly supported, $\Theta(x)$ a smoothing function such that $\psi(x) = \frac{d^n}{dx^n}\Theta(x)$. Let f be a tempered distribution whose wavelet transform is well defined over (a,b) and let $x_0 \in (a,b)$. We suppose that there exists a scale $s_0 > 0$ and a constant C such that for $x \in (a,b)$ and $s < s_0$, $\mathrm{Max}(\mathcal{W}_\psi f(s,x)) \subset \mathrm{Cone}(x_0, C)$.*

- *$x_1 \in (a,b)$, $x_1 \neq x_0 \Rightarrow f$ is uniformly Hölder n in a neighborhood of x_1.*
- *$\alpha < n$ non-integer. f is Hölder α at x_0 if and only if there exists a constant A such that at each local maxima $(s,x) \in \mathrm{Cone}(x_0, C)$, we have*

$$|\mathcal{W}_\psi f(s,x)| \leq As^\alpha. \tag{2.25}$$

Proof. This is Theorem 5.3 in Mallat, Hwang (38) and is proved in Chapter 3. □

So, if all the modulus maxima for $x \in (a,b)$ and $s < s_0$ are within the Cone of Influence at x_0, we know that the function is Hölder n at $x_1 \neq x_0$ and we get a characterization of the regularity at x_0 as in Theorem 2.24 and Theorem 2.39 by considering the decay of $\mathcal{W}_\psi f(s,x)$ *only inside* the Cone of Influence.

Theorem 2.55 (Mallat, Hwang (38) (Theorem 5.4)). *Let $\psi \in C^n(\mathbb{R})$ be a wavelet such that $\mathrm{supp}(\psi) \subset [-K,K]$ and $\psi(x) = \frac{d^n}{dx^n}\Theta(x)$ where Θ is a strictly positive function on $(-K,K)$. Let $x_0 \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$. Suppose:*

- *There exists an interval (a,b), with $x_0 \in (a,b)$, and a scale $s_0 > 0$ such that the wavelet transform $\mathcal{W}_\psi f(s,x)$ has constant sign for $s < s_0$ and $x \in (a,b)$.*
- *There exists constants B, $\epsilon > 0$ such that for all points $x \in (a,b)$ and any scale s,*

$$|\mathcal{W}_\psi f(s,x)| \leq Bs^\epsilon. \tag{2.26}$$

- *Let $x = X(s)$ be a curve such that*

  $$(s, X(s)) \in \mathrm{Cone}(x_0, K),\ \forall s < s_0,$$

  with $K < C$ (i.e the curve $\{(s, X(s)\}$ is in a cone strictly smaller than the Cone of Influence). Then there exists a constant A such that for any scale $s < s_0$, the wavelet transform satisfies

  $$|\mathcal{W}_\psi f(s, X(s))| \leq As^\gamma \text{ with } 0 \leq \gamma \leq n, \tag{2.27}$$

Then f is Hölder α at x_0, for any $\alpha < \gamma$.

Proof. This is Theorem 5.4 in Mallat, Hwang (38) and is proved in Chapter 3. □

So if Θ is strictly positive on the interior of its support, $\psi(x) = \frac{d^n}{dx^n}\Theta(x)$, $|\mathcal{W}_\psi f(s, x)| \leq Bs^\epsilon$ for some $\epsilon > 0$ on $\mathbb{R} \times (a, b)$ and $\mathcal{W}_\psi f(s, x)$ has constant sign on $\{s < s_0\} \times (a, b)$, where $x_0 \in (a, b)$, then we can estimate α at x_0 by the decay of $|\mathcal{W}_\psi f(s, x)|$ along any curve strictly inside the Cone of Influence.

We have now presented several theorems, some of them quite similar but with slightly different conditions and conclusions, regarding necessary or sufficient conditions on the decay of the wavelet transform across scales and the regularity of a function. The overall conclusion is that the wavelet transform, with rather weak conditions on the wavelets used, enables us to characterize pointwise behavior of functions.

3 THE PROOFS

In this chapter we will prove the theorems of Mallat and Hwang (38). All the theorems have sketched proofs in that paper, or refer to other sources that have more or less complete proofs, so the work done here is to collect all of them in one consistent form, and to fill in the gaps. Especially Lemma 3.5 on page 36 has a long proof (22 pages) even though it is rather elementary. All the special cases that needs to be considered is what makes it that long.

3.1 Hölder Regular Functions

Theorem 3.1 (Mallat, Hwang (Theorem 3.3 a)). *Let $0 < \alpha \leq n \in \mathbb{N}$. Let $[a, b] \subset \mathbb{R}$ be an interval and $(b - a) > 2\epsilon > 0$. Suppose that $\psi \in M_n(\mathbb{R})$ is a wavelet and $||x^\alpha \psi||_{L^1(\mathbb{R})} < \infty$. If a function $f(x) \in L^2(\mathbb{R})$ is uniformly Hölder α over any interval $(a + \epsilon, b - \epsilon)$, then*

$$|\mathcal{W}_\psi f(s, x)| = \mathbf{O}(s^\alpha),\ x \in (a + \epsilon, b - \epsilon),\ s > 0. \tag{3.1}$$

Proof. Let $v = \frac{x-u}{s}$. By the definition of the wavelet transform (Definition 2.2),

$$\begin{aligned}
|\mathcal{W}_\psi f(s,x)| &= |\int_{\mathbb{R}} f(u)\frac{1}{s}\psi\left(\frac{x-u}{s}\right)\,du| \\
&= |\int_{\mathbb{R}} f(u)\frac{1}{s}\psi\left(\frac{x-u}{s}\right)\,du| \\
&\quad -|\int_{\mathbb{R}} f(x)\frac{1}{s}\psi\left(\frac{x-u}{s}\right)\,du| \\
&= |\int_{\mathbb{R}} (f(u)-f(x))\frac{1}{s}\psi\left(\frac{x-u}{s}\right)\,du| \\
&\leq C\int_{\mathbb{R}} |u-x|^\alpha\frac{1}{s}|\psi\left(\frac{x-u}{s}\right)|\,du \\
&= C\int_{\mathbb{R}} |sv|^\alpha\frac{1}{s}|\psi(v)|s\,dv \\
&= C\int_{\mathbb{R}} |v^\alpha\psi(v)|\,dv\; s^\alpha \\
&= C\,||x^\alpha\psi||_{L^1(\mathbb{R})}s^\alpha \\
&= As^\alpha.
\end{aligned}$$

$\square$

Theorem 3.2 (Mallat, Hwang (Theorem 3.4 a)). *Let* $\alpha \leq n \in \mathbb{N}$. *Suppose* $\psi \in C^n(\mathbb{R}) \cap M_n(\mathbb{R})$ *is a wavelet, such that* $||x^\alpha\psi||_{L^1(\mathbb{R})} < \infty$. *If a function* $f(x)$ *is Hölder* α *at* x_0, *then for all points* x *in a neighborhood of* x_0 *and any scale* s,

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\alpha + |x-x_0|^\alpha). \tag{3.2}$$

Proof. By an overall translation, we may assume that $x_0 = 0$.

$$\begin{aligned}
|\mathcal{W}_\psi f(s,x)| &= |\int_{\mathbb{R}} f(u)\frac{1}{s}\psi\left(\frac{x-u}{s}\right)\,du| \\
&= |\int_{\mathbb{R}} (f(u)-f(0))\frac{1}{s}\psi\left(\frac{x-u}{s}\right)\,du| \\
&\leq \int_{\mathbb{R}} C\,|x|^\alpha\frac{1}{s}|\psi\left(\frac{x-u}{s}\right)|\,du \\
&= C\int_{\mathbb{R}} |x-sv|^\alpha\,|\psi(v)|\,dv \\
&\leq C(C_1 s^\alpha||x^\alpha\psi||_{L^1(\mathbb{R})} + C_2|x|^\alpha||\psi||_{L^1(\mathbb{R})}) \\
&= A(s^\alpha + |x|^\alpha)
\end{aligned}$$

$\square$

3.2 Global Hölder Regularity

Theorem 3.3 (Mallat,Hwang (3.3 b)). *Suppose that* $0 < \alpha < n \in \mathbb{N}$, $\alpha \notin \mathbb{N}$, $\psi \in C(\mathbb{R}) \cap M_n(\mathbb{R})$ *is a wavelet and* $||\psi'||_{L^1(\mathbb{R})} < \infty$. *Let* $[a, b] \subset \mathbb{R}$ *be an interval. If*

$$|\mathcal{W}_\psi f(s, x)| = \mathbf{O}(s^\alpha). \tag{3.3}$$

for any $x \in (a + \epsilon, b - \epsilon)$ $((b - a) > 2\epsilon > 0)$ *and any scale* $s > 0$, *then* $f(x)$ *is uniformly Hölder* α *over any such interval* $(a + \epsilon, b - \epsilon)$.

Proof. Let $0 < s_0 < \infty$. By the inversion formula (Lemma 2.3),

$$\begin{aligned} f(x) &= \frac{1}{C_\psi} \int_0^\infty \int_{\mathbb{R}} \mathcal{W}_\psi f(s, u) \overline{\psi}_s(u - x)\, du\, \frac{ds}{s} \\ &= \frac{1}{C_\psi} \int_0^{s_0} \int_{\mathbb{R}} \mathcal{W}_\psi f(s, u) \overline{\psi}_s(u - x)\, du\, \frac{ds}{s} \\ &\quad + \frac{1}{C_\psi} \int_{s_0}^\infty \int_{\mathbb{R}} \mathcal{W}_\psi f(s, u) \overline{\psi}_s(u - x)\, du\, \frac{ds}{s} \\ &= f_{small}(x) + f_{large}(x), \end{aligned}$$

as in Lemma 2.9.

- The function f_{large} is a smooth function by Lemma 2.9 and consequently Hölder α for $0 < \alpha < 1$.

- Let $s < s_0$ and $0 \leq h \leq s_0$.

$$\begin{aligned}
f(x) &= |f_{small}(x+h) - f_{small}(x)| \\
&= \left| \int_0^{s_0} \int_{\mathbb{R}} \{\psi_s(x+h-u) - \psi_s(x-u)\} \mathcal{W}_\psi f(s,u)\, du \frac{ds}{s} \right| \\
&\leq \left| \int_0^{h} \int_{\mathbb{R}} \psi_s(x+h-u) \mathcal{W}_\psi f(s,u)\, du \frac{ds}{s} \right| \\
&\quad + \left| \int_0^{h} \int_{\mathbb{R}} \psi_s(x-u) \mathcal{W}_\psi f(s,u)\, du \frac{ds}{s} \right| \\
&\quad\quad + \left| \int_h^{s_0} \int_{\mathbb{R}} \{\psi_s(x+h-u) - \psi_s(x-u)\} \mathcal{W}_\psi f(s,u)\, du \frac{ds}{s} \right| \\
&\leq \int_0^{h} \int_{\mathbb{R}} |\psi_s(x+h-u)| A_\epsilon s^\alpha \, du \, \frac{ds}{s} \\
&\quad + \int_0^{h} \int_{\mathbb{R}} |\psi_s(x-u)| A_\epsilon s^\alpha \, du \, \frac{ds}{s} \\
&\quad\quad + \int_h^{s_0} \int_{\mathbb{R}} \frac{h}{s^2} \left| \psi' \left(\frac{x+\tau-u}{s} \right) \right| |\mathcal{W}_\psi f(s,u)| \, du \, \frac{ds}{s} \\
&\leq A_\epsilon ||\psi||_{L^1(\mathbb{R})} \int_0^h s^\alpha \frac{ds}{s} \\
&\quad + A_\epsilon ||\psi||_{L^1(\mathbb{R})} \int_0^h s^\alpha \frac{ds}{s} \\
&\quad\quad + \int_h^{s_0} \int_{\mathbb{R}} \frac{h}{s^2} |\psi' \left(\frac{x+\tau-u}{s} \right) | A_\epsilon s^\alpha \, du \, \frac{ds}{s} \\
&\leq 2A_\epsilon \int_0^h s^\alpha \frac{ds}{s} \\
&\quad + A_\epsilon h ||\psi'||_{L^1(\mathbb{R})} \int_h^{s_0} s^{\alpha-1} \frac{ds}{s} \\
&\leq Ch^\alpha,
\end{aligned}$$

  since $\psi_s\left(x+h-u\right) - \psi_s\left(x-u\right) = \frac{h}{s^2}\psi'\left(\frac{x+\tau-u}{s}\right)$ for a $\tau \in (0,h)$ by the *Mean Value Theorem*.

We have proven that

$$|f(x+h) - f(x)| \leq Ch^\alpha, \tag{3.4}$$

i.e. the function is uniformly Hölder α.

□

3.3 Local Hölder Regularity

Theorem 3.4 (Mallat,Hwang (Theorem 3.4 b)). *Let $\psi(x) \in C^n(a,b) \cap M_n(\mathbb{R})$ be a wavelet with compact support. Let $0 < \alpha < n,\ \alpha \notin \mathbb{N}$. A function $f(x)$ is Hölder α at x_0, if the two* *following conditions hold.*

- *There exists* $\epsilon > 0$ *such that for all points* x *in a neighborhood of* x_0 *and any scale* s*,*

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}(s^\epsilon). \tag{3.5}$$

- *For all points* x *in a neighborhood of* x_0 *and any scale* s

$$|\mathcal{W}_\psi f(s,x)| = \mathbf{O}\left(s^\alpha + \frac{|x-x_0|^\alpha}{|\log(|x-x_0|)|}\right). \tag{3.6}$$

then $f(x)$ *is Hölder* α *at* x_0

Proof. By an overall translation and dilation, we may assume that $x_0 = 0$ and $\mathrm{supp}(\psi) \subset [-1/2, 1/2]$. Let $0 < h < s_0 < 1$. We only consider the case $h > 0$. $h < 0$ can be treated analogously. The inversion formula gives,

$$\begin{aligned} f(x) &= \frac{1}{C_\psi}\int_0^\infty \int_{\mathbb{R}} \mathcal{W}_\psi f(s,u)\overline{\psi}_s(u-x)\,du\,\frac{ds}{s} \\ &= \frac{1}{C_\psi}\int_0^{s_0} \int_{\mathbb{R}} \mathcal{W}_\psi f(s,u)\overline{\psi}_s(u-x)\,du\,\frac{ds}{s} \\ &\quad + \frac{1}{C_\psi}\int_{s_0}^\infty \int_{\mathbb{R}} \mathcal{W}_\psi f(s,u)\overline{\psi}_s(u-x)\,du\,\frac{ds}{s} \\ &= f_{small}(x) + f_{large}(x). \end{aligned}$$

Again, f_{large} is smooth, and consequently Hölder α by Lemma 2.16. Define

$$\eta(h) = h^{\alpha/\epsilon}. \tag{3.7}$$

We typically have $\alpha > \epsilon$, otherwise $f(x)$ is uniformly Hölder $\alpha < \epsilon$ in the neighborhood of x_0 by (3.5) and Theorem 3.3 and the result would be trivial. For $0 < h < 1$, we have $0 < \eta(h) < h$. We then have

$$\Delta(h) = f_{small}(h) - f_{small}(0) \tag{3.8}$$

$$= \int_0^{s_0}\int_{\mathbb{R}} \{\psi_s(h-x) - \psi_s(-x)\}\mathcal{W}_\psi f(s,x)\,dx\,\frac{ds}{s} \tag{3.9}$$

$$= \int_0^{\eta(h)}\int_{\mathbb{R}} \psi_s(h-x)\mathcal{W}_\psi f(s,x)\,dx\,\frac{ds}{s} \tag{3.10}$$

$$+ \int_{\eta(h)}^{h}\int_{\mathbb{R}} \psi_s(h-x)\mathcal{W}_\psi f(s,x)\,dx\,\frac{ds}{s} \tag{3.11}$$

$$- \int_0^{h}\int_{\mathbb{R}} \psi_s(h-x)\mathcal{W}_\psi f(s,x)\,dx\,\frac{ds}{s} \tag{3.12}$$

$$+ \int_h^{s_0}\int_{\mathbb{R}} \{\psi_s(h-x) - \psi_s(-x)\}\mathcal{W}_\psi f(s,x)\,dx\,\frac{ds}{s} \tag{3.13}$$

We will estimate (3.10) -(3.13) separately.

(3.10):

$$
\begin{aligned}
|(3.10)| &\leq \int_0^{\eta(h)} \int_{\mathbb{R}} |\psi_s(h-x)|\ |\mathcal{W}_\psi f(s,x)|\ dx\ \frac{ds}{s} \\
&\leq \int_0^{\eta(h)} \int_{\mathbb{R}} |\psi_s(h-x)|\ As^\epsilon\ dx\ \frac{ds}{s} \\
&\leq \int_0^{\eta(h)} ||\psi_s||_1 As^\epsilon\ \frac{ds}{s} \\
&= ||\psi_s||_1 A \int_0^{h^{\alpha/\epsilon}} s^{\epsilon-1}\ ds \\
&= ||\psi||_1 A\,[s^\epsilon]_0^{h^{\alpha/\epsilon}} \\
&= ||\psi||_1 Ah^\alpha \\
&= C_1 h^\alpha.
\end{aligned}
$$

(3.11):

$$
\begin{aligned}
|(3.11)| &\leq \int_{\eta(h)}^{h} \int_{\mathbb{R}} |\psi_s(h-x)|\ |\mathcal{W}_\psi f(s,x)|\ dx\ \frac{ds}{s} \\
&\leq \int_{\eta(h)}^{h} \int_{\mathbb{R}} |\psi_s(h-x)| B\left(s^\alpha + \frac{|x|^\alpha}{|\ln(|x|)|}\right)\ dx \frac{ds}{s} \\
&= B\int_{\eta(h)}^{h} \int_{\mathbb{R}} |\psi_s(h-x)| s^\alpha\ dx \frac{ds}{s} \\
&\quad +\ B\int_{\eta(h)}^{h} \int_{\mathbb{R}} |\psi_s(h-x)| \frac{|x|}{|\ln(|x|)|}\ dx\ \frac{ds}{s} \\
&\leq B||\psi||_1 \int_{\eta(h)}^{h} s^\alpha \frac{ds}{s} \\
&\quad +\ B\frac{|h|^\alpha}{|\ln(h)|}||\psi||_1 \int_{\eta(h)}^{h} \frac{ds}{s} \\
&= B||\psi||_1\,[s^\alpha]_{h^{\alpha/\epsilon}}^{h} \\
&\quad +\ B\frac{|h|^\alpha}{|\ln(h)|}||\psi||_1\,[\ln(s)]_{h^{\alpha/\epsilon}}^{h} \\
&= B|\psi||_1(h^\alpha - h^{2\alpha/\epsilon}) \\
&\quad +\ B||\psi||_1 \frac{|h|^\alpha}{|\ln(h)|}(\ln(h) - \ln(h^{\alpha/\epsilon})) \\
&= B||\psi||_1\left(h^\alpha - h^{2\alpha/\epsilon} + \frac{h^\alpha(1-\alpha/\epsilon)\ln(h)}{|\ln(h)}\right) \\
&\leq B||\psi||_1(1-\alpha/\epsilon+1)h^\alpha \\
&= C_2 h^\alpha.
\end{aligned}
$$

(3.12):

$$
\begin{aligned}
|(3.12)| &= |\int_0^h \int_{\mathbb{R}} \psi_s(h-x)\mathcal{W}_\psi f(s,x)\ dx\ \frac{ds}{s}| \\
&\leq \int_0^h \int_{\mathbb{R}} |\psi_s(h-x)|\ |\mathcal{W}_\psi f(s,x)|\ dx\ \frac{ds}{s} \\
&\leq A_\epsilon ||\psi||_1 \int_0^h s^\alpha \frac{ds}{s} \\
&\leq C_3 h^\alpha,
\end{aligned}
$$

since this integral runs inside the Cone of Influence and by Theorem 3.3.

(3.13): As in the proof of Theorem 3.3, by the Mean Value Theorem there exists a $\tau \in (0,h)$, such that

$$
\begin{aligned}
|(3.13)| &= |\int_h^{s_0} \int_{\mathbb{R}} \{\psi_s(h-x) - \psi_s(-x)\}\mathcal{W}_\psi f(s,x)\ dx\ \frac{ds}{s}| \\
&\leq \int_h^{s_0} \int_{\mathbb{R}} |\frac{h}{s^2}\psi'(\frac{\tau - x}{s})|\ |\mathcal{W}_\psi f(s,x)| dx\ \frac{ds}{s} \\
&= C_4 h^\alpha
\end{aligned}
$$

Thus $|\Delta(h)| \leq Ch^\alpha$ and f is Hölder α at x_0

$\square$

3.4 CWT Local Maxima

The following lemma is a rather simple curve-analysing lemma, but with a long proof, given all the different cases that needs to be considered. Is is used in the proof of Proposition 3.6.

Lemma 3.5. *Let $[c,d] \subset \mathbb{R}$, $0 < \beta < \frac{(d-c)}{4}$ and $K > 0$. Let $g \in C^2([c,d])$ be a function which satisfies*

$$\int_c^d |g(x)|\ dx < K. \tag{3.14}$$

- *If $\left|\frac{dg(x)}{dx}\right|$ has no local maxima on (c,d) and $x \in [c+\beta, d-\beta]$, then*

$$|g(x)| \leq 2\frac{K}{\beta} = B_{K,\beta}, \tag{3.15}$$

and

$$\left|\frac{d\ g(x)}{dx}\right| \leq 12\frac{K}{\beta^2} = C_{K,\beta}. \tag{3.16}$$

- *If* $\left|\frac{d^2g(x)}{dx^2}\right|$ *has no local maxima on* (c,d) *and* $x \in [c+\beta, d-\beta]$*, then*

$$\left|\frac{d^2g(x)}{dx^2}\right| \leq 120\frac{K}{\beta^3} = D_{K,\beta}. \tag{3.17}$$

Proof. The proof is elementary, but we have to consider several different cases. The overall structure for the proof of 3.15 is as follows:

- $g'(x)$ has constant sign:
  1. $g'(x) > 0$ and $g(x) > 0$:
  2. $g'(x) > 0$ and $g(x) < 0$:
  3. $g'(x) > 0$ and $g(x)$ changes sign:
  4. $g'(x) < 0$ and $g(x) > 0$:
  5. $g'(x) < 0$ and $g(x) < 0$:
  6. $g'(x) < 0$ and $g(x)$ changes sign:
- $g'(x)$ is monotonic and changes sign:
  1. $g'(x)$ decreases:
     - $g(x)$ is negative:
     - $g(x)$ is positive:
     - $g(x)$ changes sign:
       * $g(c+\beta) < 0$ and $g(d-\beta) < 0$:
       * $g(c+\beta) < 0$ and $g(d-\beta) > 0$:
       * $g(c+\beta) > 0$ and $g(d-\beta) < 0$:
  2. $g'(x)$ increases:
     - $g(x)$ is positive:
     - $g(x)$ is negative:
     - $g(x)$ changes sign:
       * $g(c+\beta) > 0$ and $g(d-\beta) > 0$:
       * $g(c+\beta) > 0$ and $g(d-\beta) < 0$:
       * $g(c+\beta) < 0$ and $g(d-\beta) > 0$:

For a given function g, $|g'(x)|$, $|g'(x)|$ and $|g''(x)|$ are trivially bounded by continuity of the functions and compactness of the interval. We will prove that the bounds are independent of the function g. We will prove (3.15), (3.16) and (3.17) separately.

(3.15): Since $|g'(x)|$ has no local maxima, either $g'(x)$ has constant sign or it is monotonic, and it looks like one of the graphs in Figure 3.1 on page 38. Then $g(x)$ looks like one of the graphs in Figure 3.2 on page 38.

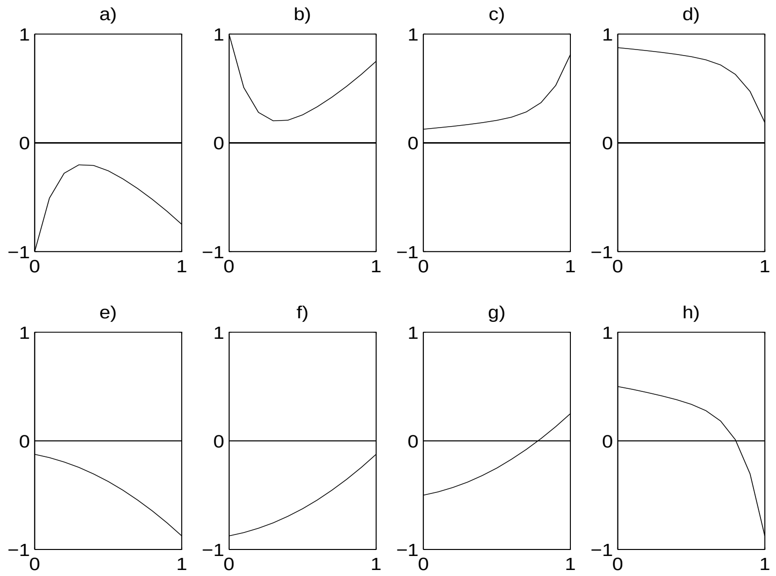

Figure 3.1 *The possible graphs of $g'(x)$ when $|g'(x)|$ has no maxima*

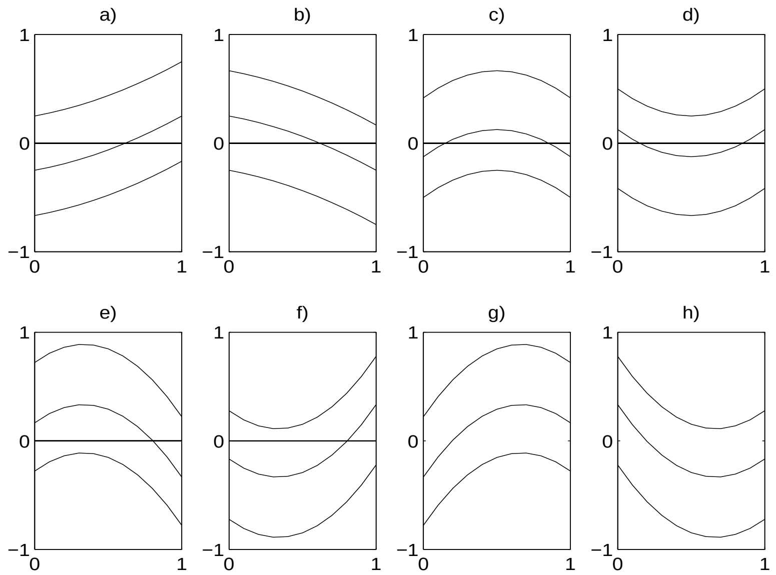

Figure 3.2 *The possible graphs og $g(x)$ when $|g'(x)|$ has no maxima.*

- $g'(x)$ has constant sign, as the graphs a)-f) in Figure 3.1: $g(x)$ is monotonic and

$$|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|). \tag{3.18}$$

From (3.14) we have

$$\int_c^{c+\beta} |g(x)|\, dx \leq K \text{ and } \int_{d-\beta}^{d} |g(x)|\, dx \leq K. \tag{3.19}$$

This implies

$$|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|) \leq \frac{K}{\beta}. \tag{3.20}$$

To prove (3.20), we must distinguish several cases:

1. $g'(x) > 0$ and $g(x) > 0$ as in a) i), Figure 3.2: The second integral in (3.19) implies that

$$\beta|g(d-\beta)| \leq \int_{d-\beta}^{d} |g(x)|\, dx \leq K,$$

for $x \in [d-\beta, d]$, i.e.
$|g(d-\beta)| \leq \dfrac{K}{\beta}$.
We also know that $|g(c+\beta)| \leq |g(d-\beta)|$.

2. $g'(x) > 0$ and $g(x) < 0$ as in a) iii), Figure 3.2: The first integral in (3.19) implies that
$|g(c+\beta)| \leq \dfrac{K}{\beta}$.
We also know that $|g(d-\beta)| \leq |g(c+\beta)|$.

3. $g'(x) > 0$ and $g(x)$ changes sign as in a) ii), Figure 3.2: The first integral in (3.19) implies that
$|g(c+\beta)| \leq \dfrac{K}{\beta}$.
The second integral in (3.19) implies that
$|g(d-\beta)| \leq \dfrac{K}{\beta}$.

4. $g'(x) < 0$ and $g(x) > 0$ as in b) i), Figure 3.2: The first integral in (3.19) implies that
$|g(c+\beta)| \leq \dfrac{K}{\beta}$.
We also know that $|g(d-\beta)| \leq |g(d-\beta)|$.

5. $g'(x) < 0$ and $g(x) < 0$ as in b) iii), Figure 3.2: The second integral in (3.19) implies that
$|g(d-\beta)| \leq \dfrac{K}{\beta}$.
We also know that $|g(c+\beta)| \leq |g(d-\beta)|$.

6. $g'(x) < 0$ and $g(x)$ changes sign as in b) ii), Figure 3.2: The first integral in (3.19) implies that

   $$|g(c+\beta)| \le \frac{K}{\beta}.$$

   The second integral in (3.19) implies that

   $$|g(d-\beta)| \le \frac{K}{\beta}. \tag{3.21}$$

From the marks 1 to 6 we get

$$|g(x)| \le \max(|g(c+\beta)|, |g(d-\beta)|) \le \frac{K}{\beta}, \tag{3.22}$$

for $x \in [c+\beta, d-\beta]$, proving (3.15) when $g'(x)$ has constant sign.

- $g'(x)$ is monotonic and changes sign as g) and h) in Figure 3.1: The curvature of $g(x)$ does not change sign.

  1. $g'(x)$ decreases. Then $g(x)$ is concave, as c), e) and g) in Figure 3.2.

     – $g(x)$ is negative as c) iii), e) iii) and g) iii) in Figure 3.2: Since $g(x)$ is negative and concave,

     $$|g(x)| \le \max(|g(c+\beta)|, |g(d-\beta)|).$$

     Since $g'(x)$ is monotonically decreasing, either it is positive on $[c, c+\beta]$ or it is negative on $[c+\beta, d]$. Since $g(x)$ remains negative, $\beta \le \frac{d-c}{4}$ and

     $$\int_c^{c+\beta} |g(x)|\, dx \le K, \text{ and } \int_{c+\beta}^{d} |g(x)|\, dx \le K,$$

     we get

     $$|g(c+\beta)| \le \max\left(\frac{K}{\beta}, \frac{K}{d-(c+\beta)}\right) = \frac{K}{\beta}. \tag{3.23}$$

     Similarly

     $$\int_c^{d-\beta} |g(x)|\, dx \le K, \text{ and } \int_{d-\beta}^{d} |g(x)|\, dx \le K,$$

     and we get

     $$|g(d-\beta)| \le \max(\frac{K}{((d-\beta)-c)}, \frac{K}{\beta}) = \frac{K}{\beta}. \tag{3.24}$$

     (3.23) and (3.24) together gives

     $$|g(x)| \le \max(|g(c+\beta)|, |g(d-\beta)|) \le \frac{K}{\beta}, \tag{3.25}$$

     when $g(x)$ is negative, concave and $g''(x)$ changes sign.

     – $g(x)$ is positive as c) i), e) i) and g) i) in Figure 3.2: There exists $e \in (c+\beta, d-\beta)$ such that $g(x) \le g(e), \forall x \in [c+\beta, d-\beta]$. Since $g(x)$ is concave, we get

     $$K \ge \int_{c+\beta}^{d-\beta} g(x)\, dx \ge \frac{g(e)((d-\beta)-(c+\beta))}{2}.$$

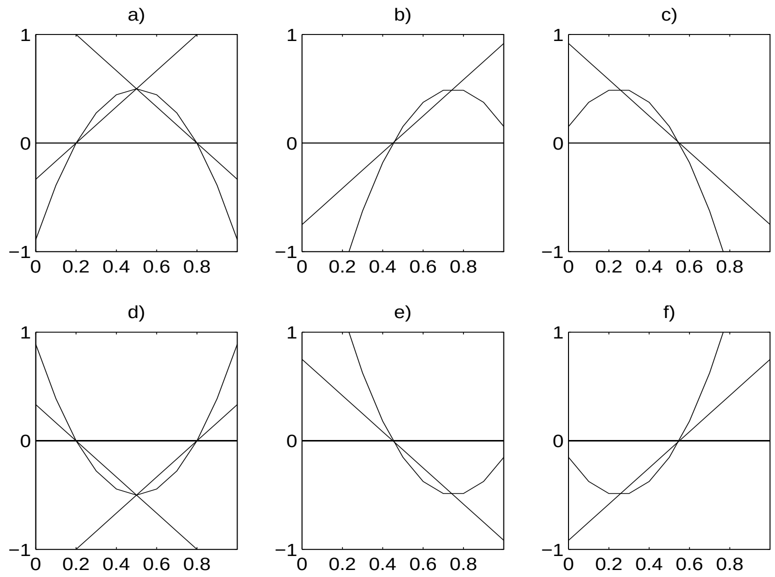

Figure 3.3 The concave and convex functions with $l_1(x)$ and $l_2(x)$

Since $\beta < \frac{d-c}{4}$, we obtain

$$\begin{aligned} |g(x)| &\leq g(e) \\ &\leq \frac{2K}{(d-\beta)-(c+\beta)} \\ &\leq \frac{2K}{4\beta - 2\beta} \\ &= \frac{K}{\beta}, \end{aligned} \tag{3.26}$$

when $g(x)$ is positive, concave and $g''(x)$ changes sign.

– $g(x)$ changes sign as c) ii), e) ii) and g) ii) in Figure 3.2:

∗ $g(c+\beta) < 0$ and $g(d-\beta) < 0$ as c) ii) in Figure 3.2: Since $g(x)$ is concave, it has two zero-crossings at the locations z_0 and z_1. For $x \in (c+\beta, z_0) \cup (z_1, d-\beta)$, $g(x)$ is negative, and

$$|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|). \tag{3.27}$$

For $x \in [c, c+\beta]$ and $x \in [d-\beta, d]$ the function is monotonic, and from the first part of the proof, we have

$$|g(c+\beta)| \leq \frac{K}{\beta} \text{ and } |g(d-\beta)| \leq \frac{K}{\beta}. \tag{3.28}$$

For $x \in [z_0, z_1]$, $g(x) \geq 0$ and there exists $e \in [z_0, z_1]$ such that $g(x) \leq g(e)$ for $x \in [z_0, z_1]$. Since $g(x)$ is concave over $[z_0, z_1]$ we have

$$K \geq \int_{z_0}^{z_1} g(x)\, dx \geq \frac{g(e)(z_1 - z_0)}{2},$$

which gives

$$g(e) \leq \frac{2K}{(z_1 - z_0)}. \tag{3.29}$$

To prove that $g(e)$ is bounded, we start by supposing that $g(e) \geq \frac{K}{\beta}$. Otherwise it is trivially true. Let $l_0(x)$ be the line that crosses $g(x)$ at the points $(z_0, 0)$ and $(e, g(e))$ as a) in Figure 3.3. Then by (3.28),

$$|l_0(c+\beta)| \leq |g(c+\beta)| \leq \frac{K}{\beta}, \tag{3.30}$$

and by assumption

$$l_0(e) = g(e) \geq \frac{K}{\beta}. \tag{3.31}$$

From Figure 3.3 we have

$$\frac{(e-z_0)}{|l_0(e)|} = \frac{(z_0-(c+\beta))}{|l_0(c+\beta)|}. \tag{3.32}$$

(3.30), (3.31) and (3.32) together gives

$$\begin{aligned} \frac{(e-z_0)}{\frac{K}{\beta}} &\geq \frac{(e-z_0)}{|l_0(e)|} \\ &= \frac{(z_0-(c+\beta))}{|l_0(c+\beta)|} \\ &\geq \frac{(z_0-(c+\beta))}{\frac{K}{\beta}}, \end{aligned} \tag{3.33}$$

which gives

$$e - z_0 \geq z_0 - (c+\beta). \tag{3.34}$$

Let $l_1(x)$ be the line that crosses $g(x)$ at $(e, g(e)$ and $(z_1, 0)$. Then by (3.28)

$$|l_1(d-\beta)| \leq |g(d-\beta)| \leq \frac{K}{\beta}, \tag{3.35}$$

and by assumption

$$|l_1(e)| = |g(e)| \geq \frac{K}{\beta}. \tag{3.36}$$

From Figure 3.3 we have

$$\frac{(z_1-e)}{|l_1(e)|} = \frac{((d-\beta)-z_1)}{|l_1(d-\beta)|}. \tag{3.37}$$

(3.35), (3.36) and (3.37) together gives

$$\frac{(z_1-e)}{\frac{K}{\beta}} \geq \frac{(z_1-e)}{|l_1(e)|} = \frac{((d-\beta)-z_1)}{|l_1(d-\beta)|} \geq \frac{((d-\beta)-z_1)}{\frac{K}{\beta}},$$

which gives

$$z_1 - e \geq (d-\beta) - z_1. \tag{3.38}$$

Adding (3.34) and (3.38) gives

$$\begin{aligned} z_1 - z_0 &\geq \frac{(d-\beta)-(c+\beta)}{2} \\ &\geq 2\frac{d-c}{4} - \beta \\ &\geq \beta. \end{aligned} \tag{3.39}$$

By inserting (3.40) into (3.29) we get

$$g(e) \leq \frac{2K}{(z_1-z_0)} \leq \frac{2K}{\beta}. \tag{3.40}$$

We then obtain, from (3.30) and (3.40),

$$|g(x)| \leq |g(e)| \leq \frac{2K}{\beta}, \tag{3.41}$$

when $g(x)$ is as c) ii) in Figure 3.2.

∗ $g(c+\beta) < 0$ and $g(d-\beta) > 0$ as g) ii) in Figure 3.2: Since $g(x)$ is concave, it has one zero-crossing at the location z_0. For $x \in (c+\beta, z_0)$, $g(x)$ is negative, and

$$|g(x)| \leq |g(c+\beta)|. \tag{3.42}$$

For $x \in [c, c+\beta]$ the function is monotonic, and from the first part of the proof, we have

$$|g(c+\beta)| \leq \frac{K}{\beta}. \tag{3.43}$$

For $x \in [z_0, d-\beta]$, $g(x) \geq 0$ and there exists $e \in [z_0, d-\beta]$ such that $g(x) \leq g(e)$ for $x \in [z_0, d-\beta]$. Since $g(x)$ is concave over $[z_0, d-\beta]$ we have

$$K \geq \int_{z_0}^{d-\beta} g(x)\, dx \geq \frac{g(e)((d-\beta)-z_0)}{2},$$

which gives

$$g(e) \leq \frac{2K}{((d-\beta)-z_0)}. \tag{3.44}$$

To prove that $g(e)$ is bounded, we start by supposing that $g(e) \geq \frac{K}{\beta}$. Otherwise it is trivially true. Let $l_0(x)$ be the line that crosses $g(x)$ at the points $(z_0, 0)$ and $(e, g(e))$ as b) in Figure 3.3. Then by (3.43),

$$|l_0(c+\beta)| \leq |g(c+\beta)| \leq \frac{K}{\beta}. \tag{3.45}$$

By assumption

$$l_0(e) = g(e) \geq \frac{K}{\beta}. \tag{3.46}$$

From b) in Figure 3.3 we have

$$\frac{(e-z_0)}{|l_0(e)|} = \frac{(z_0-(c+\beta))}{|l_0(c+\beta)|}. \tag{3.47}$$

(3.45), (3.46) and (3.47) together gives

$$\begin{aligned} \frac{(e-z_0)}{\frac{K}{\beta}} &\geq \frac{(e-z_0)}{|l_0(e)|} \\ &= \frac{(z_0-(c+\beta))}{|l_0(c+\beta)|} \\ &\geq \frac{(z_0-(c+\beta))}{\frac{K}{\beta}}, \end{aligned} \tag{3.48}$$

which gives

$$e - z_0 \geq z_0 - (c+\beta). \tag{3.49}$$

We also have

$$(d-\beta) - e \geq 0. \tag{3.50}$$

Adding (3.49) and (3.50) gives

$$(d-\beta) - z_0 \geq z_0 - (c+\beta). \tag{3.51}$$

Since $(d-c) > 4\beta$ we have

$$((d-\beta)-z_0)+(z_0-(c+\beta)) > 2\beta. \tag{3.52}$$

(3.51) and (3.52) implies that

$$((d-\beta)-z_0) \geq \beta. \tag{3.53}$$

By inserting (3.53) into (3.44) we get

$$g(e) \leq \frac{2K}{((d-\beta)-z_0)} \leq \frac{2K}{\beta}. \tag{3.54}$$

We then obtain

$$|g(x)| \leq g(e) \leq \frac{2K}{\beta}, \tag{3.55}$$

when $g(x)$ is as g) ii) in Figure 3.2.

* $g(c+\beta) > 0$ and $g(d-\beta) < 0$ as e) ii) in Figure 3.2: Since $g(x)$ is concave, it has one zero-crossing at the location z_1. For $x \in (z_1, d-\beta)$, $g(x)$ is negative, and

$$|g(x)| \leq |g(d-\beta)|. \tag{3.56}$$

For $x \in [d-\beta, d]$ the function is monotonic, and from the first part of the proof , we have

$$|g(d-\beta)| \leq \frac{K}{\beta}. \tag{3.57}$$

For $x \in [(c+\beta), z_1]$, $g(x) \geq 0$ and there exists $e \in [c+\beta, z_1]$ such that $g(x) \leq g(e)$ for $x \in [(c+\beta), z_1]$. Since $g(x)$ is concave over $[(c+\beta), z_1]$ we have

$$K \geq \int_{c+\beta}^{z_1} g(x)\, dx \geq \frac{g(e)(z_1-(c+\beta))}{2},$$

which gives

$$g(e) \leq \frac{2K}{(z_1-(c+\beta))}. \tag{3.58}$$

To prove that $g(e)$ is bounded, we start by supposing that $g(e) \geq \frac{K}{\beta}$. Otherwise it is trivially true. We have that

$$e-(c+\beta) \geq 0. \tag{3.59}$$

Let $l_1(x)$ be the line that crosses $g(x)$ at $(e, g(e))$ and $(z_1, 0)$ as c) in Figure 3.3. Then by (3.57)

$$|l_1(d-\beta)| \leq |g(d-\beta)| \leq \frac{K}{\beta}, \tag{3.60}$$

and by assumption

$$l_1(e) = g(e) \geq \frac{K}{\beta}. \tag{3.61}$$

From c) in Figure 3.3,

$$\frac{(z_1-e)}{|l_1(e)|} = \frac{((d-\beta)-z_1)}{|l_1(d-\beta)|}. \tag{3.62}$$

(3.60), (3.61) and (3.62) together gives

$$\begin{aligned}\frac{(z_1-e)}{\frac{K}{\beta}} &\geq \frac{(z_1-e)}{|l_1(e)|} \\ &= \frac{((d-\beta)-z_1)}{|l_1(d-\beta)|} \\ &\geq \frac{((d-\beta)-z_1)}{\frac{K}{\beta}},\end{aligned} \tag{3.63}$$

which gives

$$z_1 - e \geq (d-\beta) - z_1. \tag{3.64}$$

Adding (3.59) and (3.64) gives

$$\begin{aligned} z_1 - z_0 &\geq \frac{(d-\beta)-(c+\beta)}{2} \\ &\geq 2\frac{d-c}{4} - \beta \\ &\geq \beta. \end{aligned} \tag{3.65}$$

By inserting (3.66) into (3.58) we get

$$g(e) \leq \frac{2K}{(z_1-z_0)} \leq \frac{2K}{\beta}. \tag{3.66}$$

We then obtain

$$|g(x)| \leq g(e) \leq \frac{2K}{\beta}, \tag{3.67}$$

when $g(x)$ is as e) ii) in Figure 3.2.

From (3.25),(3.27), (3.41), (3.55) and (3.67) we get

$$\begin{aligned} |g(x)| &\leq \max(|g(c+\beta)|, g(e), |g(d-\beta)|) \\ &\leq 2\frac{K}{\beta}, \end{aligned} \tag{3.68}$$

when $g'(x)$ is decreasing and changes sign.

2. $g'(x)$ increases. Then $g(x)$ is convex as d), f) and h) in Figure 3.2.
   - $g(x)$ is positive as d) i), f) i) and h) i) in Figure 3.2:

     $|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|)$.

     Since $g'(x)$ is monotonically increasing, either it is negative on $[c, c+\beta]$ or it is positive on $[c+\beta, d]$. Since $g(x)$ remains positive, $\beta \leq \frac{d-c}{4}$ and

     $$\int_c^{c+\beta} |g(x)|\, dx \leq K, \text{ and } \int_{c+\beta}^{d} |g(x)|\, dx \leq K,$$

     we get

     $$\begin{aligned} |g(c+\beta)| &\leq \max\left(\frac{K}{\beta}, \frac{K}{d-(c+\beta)}\right) \\ &= \frac{K}{\beta}. \end{aligned} \tag{3.69}$$

     Similarly

     $$\int_c^{d-\beta} |g(x)|\, dx \leq K, \text{ and } \int_{d-\beta}^{d} |g(x)|\, dx \leq K,$$

and we get

$$|g(d-\beta)| \leq \max\left(\frac{K}{((d-\beta)-c)}, \frac{K}{\beta}\right) = \frac{K}{\beta}. \tag{3.70}$$

(3.70) and (3.71) together gives

$$|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|) \leq \frac{K}{\beta}. \tag{3.71}$$

- $g(x)$ is negative as d) iii), f) iii) and h) iii) in Figure 3.2: There exists $e \in (c+\beta, d-\beta)$ such that $|g(x)| \leq |g(e)|, \forall x \in [c+\beta, d-\beta]$. Since $g(x)$ is convex, we get

  $$K \geq \int_{c+\beta}^{d-\beta} g(x)\, dx \geq \frac{g(e)((d-\beta)-(c+\beta))}{2}.$$

  Since $\beta < \frac{d-c}{4}$, we obtain

  $$|g(x)| \leq |g(e)| \leq \frac{2K}{(d-\beta)-(c+\beta)} = \frac{2K}{2\beta} = \frac{K}{\beta}. \tag{3.72}$$

- $g(x)$ changes sign as d) ii), f) ii) and h) ii) in Figure 3.2:
  * $g(c+\beta) > 0$ and $g(d-\beta) > 0$ as d) ii) in Figure 3.2: Since $g(x)$ is convex, it has two zero-crossings at the locations z_0 and z_1. For $x \in (c+\beta, z_0) \cup (z_1, d-\beta)$, $g(x)$ is positive, and

    $$|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|). \tag{3.73}$$

    For $x \in [c, c+\beta]$ and $x \in [d-\beta, d]$ the function is monotonic, and from the first part of the proof, we have

    $$|g(c+\beta)| \leq \frac{K}{\beta} \text{ and } |g(d-\beta)| \leq \frac{K}{\beta}. \tag{3.74}$$

    For $x \in [z_0, z_1]$, $g(x) \leq 0$ and there exists $e \in [z_0, z_1]$ such that $|g(x)| \leq |g(e)|$ for $x \in [z_0, z_1]$. Since $g(x)$ is convex over $[z_0, z_1]$ we have

    $$K \geq \int_{z_0}^{z_1} |g(x)|\, dx \geq \frac{|g(e)|(z_1-z_0)}{2},$$

    which gives

    $$|g(e)| \leq \frac{2K}{(z_1-z_0)}. \tag{3.75}$$

    To prove that $g(e)$ is bounded, we start by supposing that $g(e) \geq \frac{K}{\beta}$. Otherwise it is trivially true. Let $l_0(x)$ be the line that crosses $g(x)$ at the points $(z_0, 0)$ and $(e, g(e))$. Then, by (3.74),

    $$|l_0(c+\beta)| \leq |g(c+\beta)| \leq \frac{K}{\beta}, \tag{3.76}$$

and, by assumption,

$$l_0(e) = g(e) \geq \frac{K}{\beta}. \tag{3.77}$$

Form d) in Figure 3.3 we have

$$\frac{(e - z_0)}{|l_0(e)|} = \frac{(z_0 - (c + \beta))}{|l_0(c + \beta)|}. \tag{3.78}$$

(3.76), (3.77) and (3.78) together gives

$$\begin{aligned}\frac{(e - z_0)}{\frac{K}{\beta}} &\geq \frac{(e - z_0)}{|l_0(e)|} \\ &= \frac{(z_0 - (c + \beta))}{|l_0(c + \beta)|} \\ &\geq \frac{(z_0 - (c + \beta))}{\frac{K}{\beta}},\end{aligned} \tag{3.79}$$

which gives

$$e - z_0 \geq z_0 - (c + \beta). \tag{3.80}$$

Let $l_1(x)$ be the line that crosses $g(x)$ at $(e, g(e))$ and $(z_1, 0)$ as . Then, by (3.74),

$$|l_1(d - \beta)| \leq |g(d - \beta)| \leq \frac{K}{\beta}, \tag{3.81}$$

and, by assumption,

$$l_1(e) = g(e) \geq \frac{K}{\beta}. \tag{3.82}$$

From d) in Figure 3.3 we have

$$\frac{(z_1 - e)}{|l_1(e)|} = \frac{((d - \beta) - z_1)}{|l_1(d - \beta)|}. \tag{3.83}$$

(3.81), (3.82) and (3.83) together gives

$$\begin{aligned}\frac{(z_1 - e)}{\frac{K}{\beta}} &\geq \frac{(z_1 - e)}{|l_1(e)|} \\ &= \frac{((d - \beta) - z_1)}{|l_1(d - \beta)} \\ &\geq \frac{((d - \beta) - z_1)}{\frac{K}{\beta}},\end{aligned} \tag{3.84}$$

which gives

$$z_1 - e \geq (d - \beta) - z_1. \tag{3.85}$$

Adding (3.80) and (3.85) gives

$$\begin{aligned}z_1 - z_0 &\geq \frac{(d - \beta) - (c + \beta)}{2} \\ &\geq 2\frac{d - c}{4} - \beta \\ &\geq \beta.\end{aligned} \tag{3.86}$$

By inserting (3.87) into (3.75) we get

$$g(e) \leq \frac{2K}{(z_1 - z_0)} \leq \frac{2K}{\beta}. \tag{3.87}$$

We then obtain

$$|g(x)| \leq |g(e)| \leq \frac{2K}{\beta}, \tag{3.88}$$

when $g(x)$ is as d) ii) in Figure 3.2.

* $g(c+\beta) > 0$ and $g(d-\beta) < 0$ as h) ii) in Figure 3.2: Since $g(x)$ is convex, it has one zero-crossing at the location z_0. For $x \in (c+\beta, z_0) \cup (z_1, d-\beta)$, $g(x)$ is negative, and

$$|g(x)| \leq \max(|g(c+\beta)|, |g(d-\beta)|). \tag{3.89}$$

For $x \in [c, c+\beta]$ the function is monotonic, and from the first part of the proof, we have

$$|g(c+\beta)| \leq \frac{K}{\beta}. \tag{3.90}$$

For $x \in [z_0, z_1]$, $g(x) \leq 0$ and there exists $e \in [z_0, d-\beta]$ such that $|g(x)| \leq |g(e)|$ for $x \in [z_0, d-\beta]$. Since $g(x)$ is convex over $[z_0, d-\beta]$ we have

$$K \geq \int_{z_0}^{d-\beta} |g(x)|\, dx \geq \frac{|g(e)|((d-\beta)-z_0)}{2},$$

which gives

$$|g(e)| \leq \frac{2K}{((d-\beta)-z_0)}. \tag{3.91}$$

To prove that $g(e)$ is bounded, we start by supposing that $|g(e)| \geq \frac{K}{\beta}$. Otherwise it is trivially true. Let $l_0(x)$ be the line that crosses $g(x)$ at the points $(z_0, 0)$ and $(e, g(e))$ as e) in Figure 3.3. Then, from (3.90),

$$|l_0(c+\beta)| \leq |g(c+\beta)| \leq \frac{K}{\beta}, \tag{3.92}$$

and, by assumption

$$|l_0(e)| = |g(e)| \geq \frac{K}{\beta} \tag{3.93}$$

and from e) in Figure 3.3,

$$\frac{(e-z_0)}{|l_0(e)|} = \frac{(z_0-(c+\beta))}{|l_0(c+\beta)|}. \tag{3.94}$$

(3.92), (3.93) and (3.94) together gives

$$\begin{aligned} \frac{(e-z_0)}{\frac{K}{\beta}} &\geq \frac{(e-z_0)}{|l_0(e)|} \\ &= \frac{(z_0-(c+\beta))}{|l_0(c+\beta)|} \\ &\geq \frac{(z_0-(c+\beta))}{\frac{K}{\beta}}, \end{aligned} \tag{3.95}$$

which gives

$$e - z_0 \geq z_0 - (c+\beta). \tag{3.96}$$

We also have

$$(d-\beta) - e \geq 0. \tag{3.97}$$

Adding (3.96) and (3.97) gives

$$(d-\beta)-z_0 \geq \frac{(d-\beta)-(c+\beta)}{2} \geq \beta. \tag{3.98}$$

By inserting (3.99) into (3.91) we get

$$|g(e)| \leq \frac{2K}{((d-\beta)-z_0)} \leq \frac{2K}{\beta}. \tag{3.99}$$

We then obtain

$$|g(x)| \leq |g(e)| \leq \frac{2K}{\beta}, \tag{3.100}$$

when $g(x)$ is as h) ii) in Figure 3.2.

∗ $g(c+\beta) < 0$ and $g(d-\beta) > 0$ as f) ii) in Figure 3.2: Since $g(x)$ is convex, it has one zero-crossing at the location z_1. For $x \in (z_1, d-\beta)$, $g(x)$ is positive, and

$$|g(x)| \leq |g(d-\beta)|. \tag{3.101}$$

For $x \in [d-\beta, d]$ the function is monotonic, and from the first part of the proof, we have

$$|g(d-\beta)| \leq \frac{K}{\beta}. \tag{3.102}$$

For $x \in [(c+\beta), z_1]$, $g(x) \leq 0$ and there exists $e \in [(c+\beta), z_1]$ such that $|g(x)| \leq |g(e)|$ for $x \in [(c+\beta), z_1]$. Since $g(x)$ is convex over $[(c+\beta), z_1]$ we have

$$K \geq \int_{c+\beta}^{z_1} |g(x)|\, dx \geq \frac{|g(e)|(z_1-(c+\beta))}{2},$$

which gives

$$g(e) \leq \frac{2K}{(z_1-(c+\beta))}. \tag{3.103}$$

To prove that $g(e)$ is bounded, we start by supposing that $|g(e)| \geq \frac{K}{\beta}$. Otherwise it is trivially true.
We have

$$e-(c+\beta) \geq 0. \tag{3.104}$$

Let $l_1(x)$ be the line that crosses $g(x)$ at $(e, g(e))$ and $(z_1, 0)$ as f) in 3.3. Then, by (3.102),

$$|l_1(d-\beta)| \leq |g(d-\beta)| \leq \frac{K}{\beta}, \tag{3.105}$$

and, by assumption,

$$|l_1(e)| = |g(e)| \geq \frac{K}{\beta}. \tag{3.106}$$

From f) in Figure 3.3 we have

$$\frac{(z_1-e)}{|l_1(e)|} = \frac{((d-\beta)-z_1)}{|l_1(d-\beta)|}. \tag{3.107}$$

(3.105), (3.106) and (3.107) together gives

$$
\begin{aligned}
\frac{(z_1-e)}{\frac{K}{\beta}}\frac{(z_1-e)}{|l_1(e)|} & \\
&= \frac{((d-\beta)-z_1)}{|l_1(d-\beta)|} \qquad (3.108)\\
&\geq \frac{((d-\beta)-z_1)}{\frac{K}{\beta}},
\end{aligned}
$$

which gives

$$z_1 - e \geq (d-\beta) - z_1. \qquad (3.109)$$

Adding (3.104) and (3.109) gives

$$
\begin{aligned}
z_1 - (c+\beta) &\geq \frac{(d-\beta)-(c+\beta)}{2} \qquad (3.110)\\
&\geq \frac{d-c}{4}\\
&\geq \beta.
\end{aligned}
$$

By inserting (3.111) into (3.103) we get

$$|g(e)| \leq \frac{2K}{(z_1-(c+\beta))} \leq \frac{2K}{\beta}. \qquad (3.111)$$

From (3.102(and (3.111), we obtain

$$|g(x)| \leq |g(e)| \leq \frac{2K}{\beta}, \qquad (3.112)$$

when $g(x)$ is as h) ii) in Figure 3.2.

This finishes the proof og (3.15), i.e.

$$|g(x)| \leq \frac{2K}{\beta}. \qquad (3.113)$$

(3.16): Since $|g'(x)|$ has no maxima on the interval $[c+\beta/2, d-\beta/2]$, we know that $|g'(x)| \leq \max(|g'(c+\beta)|, |g'(d-\beta)|)$ for $x \in [c+\beta, d-\beta]$.

- Suppose $|g'(c+\beta)| \geq |g'(d-\beta)|$: Then $|g'(x)|$ is decreasing on $[c+\beta/2, c+\beta]$ and $g'(x)$ does not change sign over this interval. Hence,

$$
\begin{aligned}
|g'(c+\beta)| &\leq \frac{2}{\beta}\left|\int_{c+\beta/2}^{c+\beta} g'(x)\, dx\right|\\
&= \frac{2}{\beta}|g(c+\beta) - g(c+\beta/2)|\\
&\leq \frac{2}{\beta}\left(2\frac{K}{\beta} + 2\frac{K}{\beta/2}\right)\\
&= \frac{2}{\beta}\left(\frac{2K+4K}{\beta}\right)\\
&= \frac{12K}{\beta^2}.
\end{aligned}
$$

- Suppose $|g'(c+\beta)| \leq |g'(d-\beta)|$: Then $|g'(x)|$ is increasing on $[d-\beta, d-\beta/2]$ and $g'(x)$ does not change sign over this interval. Hence,

$$\begin{aligned}
|g'(d-\beta)| &\leq \frac{2}{\beta}\left|\int_{d-\beta}^{d-\beta/2} g'(x)\,dx\right| \\
&= \frac{2}{\beta}|g(d-\beta/2) - g(d-\beta)| \\
&\leq \frac{2}{\beta}\left(2\frac{K}{\beta/2} + 2\frac{K}{\beta}\right) \\
&= \frac{2}{\beta}\left(\frac{4K+2K}{\beta}\right) \\
&= 12\frac{K}{\beta^2}.
\end{aligned}$$

Hence

$$|g'(x)| \leq \max(|g'(c+\beta)|, |g'(d-\beta)|) \leq 12\frac{K}{\beta^2}, \tag{3.114}$$

for $x \in [c+\beta, d-\beta]$.

(3.17): Since $|g''(x)|$ has no local maxima, either $g''(x)$ has constant sign or it is monotonic, and its graph looks like one of the graphs in Figure 3.1 on page 38. Then the graph of $g'(x)$ looks like one of the 24 different graphs in Figure 3.4. In all cases $|g''(x)| \leq \max(|g''(c+\beta)|, |g''(d-\beta)|)$ for $x \in [c+\beta, d-\beta]$.

- Suppose $|g''(c+\beta)| \geq |g''(d-\beta)|$: Then

$$\begin{aligned}
|g''(c+\beta)| &\leq \frac{2}{\beta}\left|\int_{c+\beta/2}^{c+\beta} g''(x)\,dx\right| \\
&= \frac{2}{\beta}\left|g'(c+\beta) - g'(c+\beta/2)\right|.
\end{aligned} \tag{3.115}$$

- Suppose $|g''(c+\beta)| \leq |g''(d-\beta)|$: Then

$$\begin{aligned}
|g''(d-\beta)| &\leq \frac{2}{\beta}\left|\int_{d-\beta}^{d-\beta/2} g''(x)\,dx\right| \\
&= \frac{2}{\beta}\left|g'(d-\beta/2) - g'(d-\beta)\right|.
\end{aligned} \tag{3.116}$$

To find bounds for $|g''(c+\beta)|$ and $|g''(d-\beta)|$ we will need similar bounds for $|g'(c+\beta)|$, $|g'(c+\beta/2)|$, $|g'(d-\beta)|$ and $|g'(d-\beta/2)|$.

- For the graphs in a) and b), iii) in c), i) in d), iii) in e), i) in f), iii) in g) and i) in h) in Figure 3.4, $|g'(x)|$ has no maxima, and by (3.16),

$$\max(|g'(c+\beta)|, |g'(d-\beta)|) \leq 12\frac{K}{\beta^2}. \tag{3.117}$$

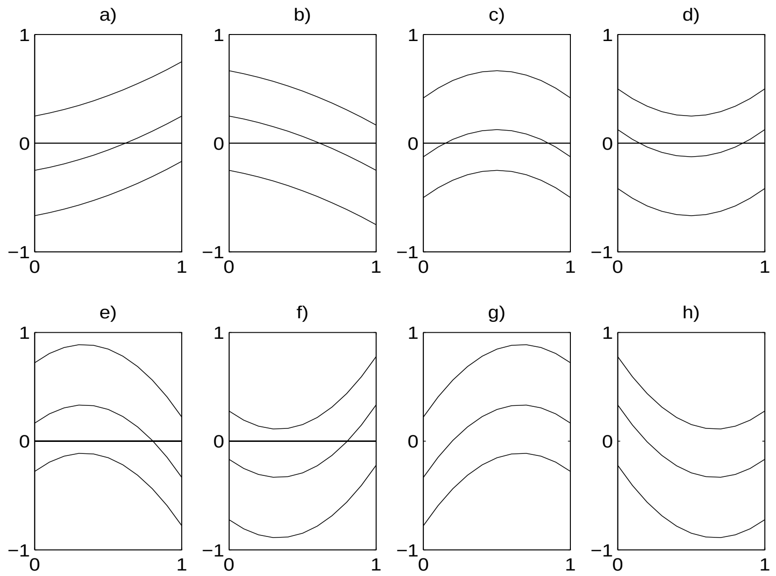

Figure 3.4 The possible graphs of $g'(x)$ when $|g''(x)|$ has no maxima.

- For i) in c), iii) in d), i) in e), iii) in f), i) in g) and iii) in h) in Figure 3.4, we have:

$$\begin{aligned} |\int_{c+\beta}^{d-\beta} g'(x)\,dx| &= |g(d-\beta)-g(c+\beta)| \\ &\leq 2\frac{K}{\beta}+2\frac{K}{\beta} \\ &= 4\frac{K}{\beta}, \end{aligned}$$

and

$$\begin{aligned} |\int_{c+\beta}^{d-\beta} g'(x)\,dx| &\geq 1/2|g'(c+\beta)-g'(d-\beta)|((d-\beta)-c+\beta) \\ &+ ((d-\beta)-(c+\beta))|g'(d-\beta)| \\ &= 1/2|g'(c+\beta)+g'(d-\beta)|((d-\beta)-(c+\beta)). \end{aligned}$$

This implies that

$$1/2|g'(c+\beta)+g'(d-\beta)|((d-\beta)-(c+\beta)) \leq 4\frac{K}{\beta},$$

and

$$\begin{aligned} |g'(c+\beta)|+|g'(d-\beta)| &\leq 8\frac{K}{\beta}\frac{1}{((d-\beta)-c+\beta))} \\ &\leq 4\frac{K}{\beta^2}. \end{aligned} \tag{3.118}$$

From (3.118) we have

$$\begin{aligned}\max(|g'(c+\beta)|, |g'(d-\beta)|) &\leq (|g'(c+\beta)| + |g'(d-\beta)|) \\ &\leq 4\frac{K}{\beta^2},\end{aligned}$$

meaning

$$|g(x)| \leq 4\frac{K}{\beta^2}, \tag{3.119}$$

- For the graphs ii) in c), ii) in d), ii) in e), ii) in f), ii) in g) and ii) in h) in Figure 3.4, we have:

$$\begin{aligned}|g'(c+\beta)| &\leq \frac{2}{\beta}\left|\int_{c+\beta/2}^{c+\beta} g'(x)\ dx\right| \\ &= \frac{2}{\beta}\left|g(c+\beta) - g(c+\beta/2)\right| \\ &\leq \frac{2}{\beta}\left(2\frac{K}{\beta} + 2\frac{K}{\beta/2}\right) \\ &= \frac{2}{\beta}\left(\frac{6K}{\beta}\right) \\ &= 12\frac{K}{\beta^2}.\end{aligned} \tag{3.120}$$

and

$$\begin{aligned}|g'(d-\beta)| &\leq \frac{2}{\beta}\left|\int_{d-\beta}^{d-\beta/2} g'(x)\ dx\right| \\ &= \frac{2}{\beta}\left|g(d-\beta/2) - g(d-\beta)\right| \\ &\leq \frac{2}{\beta}\left(2\frac{K}{\beta/2} + 2\frac{K}{\beta}\right) \\ &= \frac{2}{\beta}\left(\frac{6K}{\beta}\right) \\ &= 12\frac{K}{\beta^2}.\end{aligned} \tag{3.121}$$

From (3.117), (3.119), (3.120 and (3.121), we have

$$\begin{aligned}|g''(x)| &\leq \frac{2}{\beta}\left|12\frac{K}{(\beta/2)^2} - 12\frac{K}{\beta^2}\right| \\ &\leq \frac{2}{\beta}\left(\frac{48K + 12K}{\beta^2}\right) \\ &= 120\frac{K}{\beta^3}\end{aligned} \tag{3.122}$$

proving (3.17).

We now have proven the upper bounds on $g(x)$, $\frac{d\,g}{dx}$ and $\frac{d^2\,g}{dx^2}$ given in Lemma 3.5 on page 36.

$\square$

Proposition 3.6 (Mallat, Hwang (A.1)). *Let ψ be a wavelet that can be written $\psi(x) = \frac{d^n\phi(x)}{dx^n}$, where $\phi(x)$ is a continuous function with compact support. Let $f(x)$ be a function and suppose that for any $\epsilon > 0$, there exists a constant K_ϵ, such that at all scales s,*

$$\int_{a+\epsilon}^{b-\epsilon} |f * \phi_s(x)| \, dx \leq K_\epsilon. \tag{3.123}$$

If $|\mathcal{W}_\psi f(s,x)|$ has no local maxima for $x \in (a,b)$ and $s < s_0$, then for any $\epsilon > 0$, there exists a constant $A_{\epsilon,n}$ such that for any $x \in (a+\epsilon, b-\epsilon)$ and $s < s_0$

$$|\mathcal{W}_\psi f(s,x)| \leq A_{\epsilon,n} s^n. \tag{3.124}$$

Proof. In the following we will suppose that $\mathrm{supp}(f) \subset [a,b]$. We prove the proposition by induction on n. Let $g(x) = (f * \phi_s)(x)$.

$\mathbf{n = 1:}$ Since $\psi(x) = \frac{d\phi(x)}{dx}$, $\mathcal{W}_\psi f(s,x) = (f * \frac{d}{dx}\phi_s)(x) = s\frac{d}{dx}(f * \phi_s)(x)$. The hypothesis supposes that $|g'(x)| = \frac{1}{s}|\mathcal{W}_\psi f(s,x)|$ has no maxima on (a,b) and that $g(x)$ satisfies (3.14) in Lemma 3.5 on page 36, for $c = a + \epsilon/2$ and $d = b - \epsilon/2$. The result of Lemma 3.5, for $\beta = \epsilon/2$ and $s < s_0$, yields

$$|\mathcal{W}_\psi f(s,x)| \leq s C_{\epsilon/2} = A_{\epsilon,1} s. \tag{3.125}$$

$\mathbf{n = 2:}$ Since $\psi(x) = \frac{d^2\phi(x)}{dx^2}$, we have that

$$\mathcal{W}_\psi f(s,x) = s^2 \frac{d^2}{dx^2}(f * \phi_s)(x) = s^2 g''(x). \tag{3.126}$$

We then apply Lemma 3.5 to $g(x) = f * \phi_s(x)$, $\beta = \epsilon/2$, $c = a + \epsilon/2$ and $d = b - \epsilon/2$. Equation (3.17) yields

$$|\mathcal{W}_\psi f(s,x)| \leq s^2 D_{\epsilon/2} = A_{\epsilon,2} s^2. \tag{3.127}$$

$\mathbf{n = k:}$ Suppose the proposition is valid for $n = k \geq 2$

$\mathbf{n = k+1:}$ Let ψ be a wavelet with $k+1$ vanishing moments. The wavelet $\psi(x)$ can be written $\psi(x) = \frac{d\chi(x)}{dx}$ where the wavelet $\chi(x)$ has k vanishing moments. Let $\frac{df(x)}{dx}$ be the derivative of f in the sense of distributions. Then $\mathcal{W}_\psi f(s,x) = s\frac{df}{dx} * \chi_s(x)$. Since ψ has 2 (at least) vanishing moments, we have already proven that $|\mathcal{W}_\psi f(s,x)| \leq A_{\epsilon,2} s^2$. By Theorem 3.3 we know that f is uniformly Hölder α on $(a+\epsilon, b-\epsilon)$, for $\alpha < 2$. Then by Lemma 2.16, $\frac{df}{dx}(x)$ is uniformly Hölder α for $\alpha < 1$. Hence $\frac{df}{dx}(x)$ is uniformly bounded on any compact interval $[a+\epsilon, b-\epsilon]$. Then $h(x) = (\frac{df}{dx} * \phi_s)(x)$ satisfies (3.123). By the induction hypothesis for $n = k$,

$$\begin{aligned} |\mathcal{W}_\psi f(s,x)| &= s\left|\frac{df}{dx} * \chi_s(x)\right| && (3.128)\\ &\leq s A'_{\epsilon,k} s^k && (3.129)\\ &= A_{\epsilon,n} s^n, && (3.130) \end{aligned}$$

which proves the proposition. $\square$

Theorem 3.7 (Mallat, Hwang (Theorem 5.2)). *Let $n \in \mathbb{N}$ and $\psi \in C^n(a,b) \cap M_n(a,b)$ be a wavelet with compact support. If there exists a scale $s_0 > 0$ such that for all scales $s < s_0$ and $x \in (a,b)$, $|\mathcal{W}_\psi f(s,x)|$ has no local maxima, then for any $\epsilon > 0$ and $\alpha < n$, f is uniformly Hölder α on $(a+\epsilon, b-\epsilon)$.*
If $\psi = \frac{d^n}{dx^n}\Theta(x)$ where $\Theta(x)$ is a smoothing function*, then f is uniformly Hölder n on any such interval $(a+\epsilon, b-\epsilon)$.*

Proof. $\psi(x) = \frac{d^n}{dx^n}\phi(x) \in C^n((a,b))$. Then $\phi(x) \in C^n((a,b))$ and $(f * \phi_s)(x)$ is continous, and consequently bounded on $[a+\epsilon, b-\epsilon]$. So

$$\int_{a+\epsilon}^{b-\epsilon} |f * \phi_s(x)| dx \leq K_\epsilon, \text{ for all } s < s_0.$$

Since $|\mathcal{W}_\psi f(s,x)|$ has no maxima we have from Proposition 3.6 that

$$|\mathcal{W}_\psi f(s,x)| \leq A_\epsilon s^\alpha. \tag{3.131}$$

$\alpha < n$: $|\mathcal{W}_\psi f(s,x)| \leq A_\epsilon s^\alpha$ and $\psi \in C^n([a,b])$ implies that $f(x)$ is uniformly Hölder α, from Theorem 3.3 on page 32.

$\alpha = n$: $\psi(x) = \frac{d^n}{dx^n}\Theta(x)$, where $\Theta(x)$ is a smooting function. We have that $|\mathcal{W}_\psi f(s,x)| = s^n|\frac{d^n f}{dx^n} * \Theta_s(x)| \leq A_{\epsilon,n}s^n$ and this implies that

$$|\frac{d^n f}{dx^n} * \Theta_s(x)| \leq A_{\epsilon,n}. \tag{3.132}$$

Since $\Theta(x)$ is a smoothing function, $\hat{\Theta}(0) \neq 0$, and by Lemma A.51,

$$\left|\frac{d^n f}{dx^n}\right| \leq A_{\epsilon,n}. \tag{3.133}$$

Then, by Lemma 2.16, $\frac{d^n f}{dx^n}$ is Hölder 0, and by the same lemma, $f(x)$ is Hölder n.

$\square$

Corollary 3.8 (Mallat, Hwang (Corollary to Theorem 5.2)). *The closure of the set of points where f is not Hölder n is included in the closure of the wavelet transform maxima of f.*

Proof. This is a simple consequence of Theorem 3.7 $\square$

Theorem 3.9 (Mallat, Hwang (Theorem 5.3)). *Let $\psi \in C^n([a,b])$ be compactly supported, $\Theta(x)$ a smoothing function such that $\psi(x) = \frac{d^n}{dx^n}\Theta(x)$. Let f be a tempered distribution whose wavelet transform is well defined over (a,b) and let $x_0 \in (a,b)$. We suppose that there exists a scale $s_0 > 0$ and a constant C such that for $x \in (a,b)$ and $s < s_0$,*
$\mathrm{Max}(\mathcal{W}_\psi f(s,x)) \subset \mathrm{Cone}(x_0, C)$.

- $x_1 \in (a,b)$, $x_1 \neq x_0 \Rightarrow f$ *is uniformly Hölder* n *in a neighborhood of* x_1.
- $\alpha < n$ *non-integer.* f *is Hölder* α *at* x_0 *if and only if there exists a constant* A *such that at each local maxima* $(s,x) \in \mathrm{Cone}(x_0, C)$*, we have*

$$|\mathcal{W}_\psi f(s,x)| \leq A s^\alpha. \tag{3.134}$$

Proof. We prove the two points separately:

- Let $x_1 \in (a, x_0)$. For $s < s_0$, $\mathrm{Max}(\mathcal{W}_\psi f(s,x)) \subset \mathrm{Cone}(x_0, C)$. Hence, for $\epsilon > 0$, such that $a + \epsilon < x_0 - \epsilon$, there exists a s_ϵ such that for $s < s_0$ and $x \in (a + \epsilon/2, x_0 - \epsilon/2)$, $|\mathcal{W}_\psi f(s,x)|$ has no maxima. By Theorem 2.50, $f(x)$ is uniformly Hölder n in $[a + \epsilon, x_0 - \epsilon]$. With the same argument, $f(x)$ is uniformly Hölder n in $[x_0 + \epsilon, b - \epsilon]$.
- $\Rightarrow$: By Theorem 3.2, $|\mathcal{W}_\psi f(s,x)| \leq A(s^\alpha + |x - x_0|^\alpha) \leq A' s^\alpha$ inside the Cone of Influence when $f(x)$ is Hölder α at x_0.

  $\Leftarrow$: Let $x_1 \in (a, x_0)$ and $x_2 \in (x_0, b)$. Then, from the first part of this proof, we know that $f(x)$ is Hölder n in neighborhoods of x_1 and x_2. By Theorem 3.1, there exists a $s_0 > 0$ such that for $x \in (x_1, x_2)$,

$$\begin{aligned} |\mathcal{W}_\psi f(s,x)| &\leq \max_{\substack{y = x_1 \\ y = x_2 \\ y \in \mathrm{Cone}(x_0, C)}} (|\mathcal{W}_\psi f(s,y)|)) \\ &= \max(A_1 s^n, A s^\alpha, A_2 s^n) \\ &\leq B s^\alpha. \end{aligned}$$

  By Theorem 3.3, $f(x)$ is Hölder α.

□

Theorem 3.10 (Mallat, Hwang (Theorem 5.4)). *Let* $\psi \in C^n(\mathbb{R})$ *be a wavelet such that* $\mathrm{supp}(\psi) \subset [-K, K]$ *and* $\psi(x) = \frac{d^n}{dx^n}\Theta(x)$ *where* Θ *is a strictly positive function on* $(-K, K)$*. Let* $x_0 \in \mathbb{R}$ *and* $f \in L^2(\mathbb{R})$*. Suppose:*

- *There exists an interval* (a,b)*, with* $x_0 \in (a,b)$*, and a scale* $s_0 > 0$ *such that the wavelet transform* $\mathcal{W}_\psi f(s,x)$ *has constant sign for* $s < s_0$ *and* $x \in (a,b)$*.*
- *There exists constants* $B, \epsilon > 0$ *such that for all points* $x \in (a,b)$ *and any scale* s*,*

$$|\mathcal{W}_\psi f(s,x)| \leq B s^\epsilon. \tag{3.135}$$

- *Let* $x = X(s)$ *be a curve such that*

$$(s, X(s)) \in \mathrm{Cone}(x_0, K),\ \forall s < s_0,$$

  with $K < C$ *(i.e the curve* $\{(s, X(s)\}$ *is in a cone strictly smaller than the Cone of Influence). Then there exists a constant* A *such that for any scale* $s < s_0$*, the wavelet transform satisfies*

$$|\mathcal{W}_\psi f(s, X(s))| \leq A s^\gamma \text{ with } 0 \leq \gamma \leq n, \tag{3.136}$$

Then f is Hölder α at x_0, for any $\alpha < \gamma$.

Proof. In order to apply Theorem 3.4, we want to prove that there exists a scale s_1 and $\epsilon > 0$ such that if $s < s_1$ and $x \in (x_0 - \epsilon, x_0 + \epsilon)$,

$$|\mathcal{W}_\psi f(s,x)| \leq B(s^\gamma + |x - x_0|^\gamma). \tag{3.137}$$

We prove this by showing separately that there exists two constants B_1 and B_2 such that

$$|\mathcal{W}_\psi f(s,x)| \leq B_1 s^\gamma, \tag{3.138}$$

when (x, s) is *inside* the Cone of Influence of x_0, and

$$|\mathcal{W}_\psi f(s,x)| \leq B_2 |x - x_0|^\gamma, \tag{3.139}$$

when (x, s) is *outside* the Cone of Influence of x_0.
Once (3.137) is proved, Theorem 2.55 is a simple consequence of Theorem 3.4, for $\alpha < \gamma$. We shall suppose that the constant sign of $\mathcal{W}_\psi f(s,x)$ in a neighborhood of x_0 is positive. For $s < s_0$ and $|X(s) - x_0| < Cs$, we have

$$\mathcal{W}_\psi f(s, X(s)) \leq A s^\gamma. \tag{3.140}$$

We first prove (3.138) and then (3.139) for

$$\epsilon = \frac{1}{4}(K - C)s_0$$

and

$$s_1 = \frac{1}{4K}(K - C)s_0.$$

- $|\mathcal{W}_\psi f(s,x)| \leq B_1 s^\gamma$ when (x, s) is in the Cone of Influence:

$$\begin{aligned} 0 &\leq \mathcal{W}_\psi f(s,x) \\ &= (f * \psi_s)(x) \\ &= (f * \frac{d^n}{dx^n}\Theta_s)(x) \\ &= s^n (f^{(n)} * \Theta_s)(x) \\ &= s^{n-1} \int_{\mathbb{R}} f^{(n)}(u)\Theta(\frac{x-u}{s})\, du \\ &= s^{n-1} \int_{x_0 - 2Ks}^{x_0 + 2Ks} f^{(n)}(u)\Theta(\frac{x-u}{x})\, du. \end{aligned} \tag{3.141}$$

  The derivative of f is in the sense of distributions, and is always defined. The last equality is valid because $\operatorname{supp}\left(\Theta(\frac{x-u}{s})\right) \subset [x_0 - 2Ks, x_0 + 2Ks]$. Let $0 < M = \max_{x\in[-K,K]} \Theta(x)$ and $x \in \left[-\frac{K+C}{2}, \frac{K+C}{2}\right]$. Then there exists a $\lambda > 0$ such that $\Theta(x) > \lambda M$. Let $s' = \frac{4Ks}{K-C}$ and $u \in [x_0 - 2Ks, x_0 + 2Ks]$. Then

$\frac{|X(s')-u|}{s'} \leq C \leq \frac{K+C}{2}$, and consequently $\Theta\left(\frac{X(s')-u}{s'}\right) > \lambda M$. Since $0 \leq \Theta(\frac{x-u}{s}) \leq M$ and by Lemma A.51, $f^{(n)} \geq 0$, in the sense of distributions, we have by (3.141)

$$\begin{aligned}
\mathcal{W}_\psi f(s,x) &\leq s^{n-1}\int_{x_0-2Ks}^{x_0+2Ks} f^{(n)}(u)\Theta(\frac{x-u}{x})\ du. \\
&\leq \frac{s^{n-1}}{\lambda}\int_{x-2Ks}^{x_0+2Ks} f^{(n)}(u)\Theta\left(\frac{X(s')-u}{s'}\right)\ du \qquad (3.142)\\
&= \frac{1}{\lambda}\mathcal{W}_\psi f(s', X(s')) \\
&\leq \frac{1}{\lambda}A(s')^\gamma \\
&= \frac{a(4K)^\gamma}{(K-C)^\gamma}s^\gamma \\
&= B_1 s^\gamma.
\end{aligned}$$

- $|\mathcal{W}_\psi f(s,x)| \leq B_2|x-x_0|^\gamma$ when (x,s) is below the Cone of Influence:

$$\begin{aligned}
0 &\leq \mathcal{W}_\psi f(s,x) \qquad (3.143)\\
&= s^{n-1}\int_{\mathbb{R}} f^{(n)}(u)\Theta\left(\frac{x-u}{s}\right)\ du \\
&= s^{n-1}\int_{x_0-2Ks_2}^{x_0+2Ks_2} f^{(n)}(u)\Theta\left(\frac{x-u}{s}\right)\ du
\end{aligned}$$

since $\text{supp}(\Theta\left(\frac{x-u}{s}\right)) \subset [x_0-2Ks_2, x_0+2Ks_2]$ when $s_2 = \frac{|x-x_0|}{K} \geq s$ since $|x-x_0| \geq Ks$

Define $s_2' = \frac{4Ks_2}{(K-C)}$. We then have by (3.143)

$$\begin{aligned}
\mathcal{W}_\psi f(s,x) &\leq s^{n-1}\int_{x_0-2Ks_2}^{x_0+2Ks_2} f^{(n)}(u)\Theta\left(\frac{x-u}{s}\right)\ du \\
&\leq \frac{s^{n-1}}{\lambda}\int_{x_0-2Ks_2}^{x_0+2Ks_2} f^{(n)}(u)\Theta\left(\frac{X(s_2')-u}{s_2'}\right)\ du \qquad (3.144)\\
&= \frac{1}{\lambda}\mathcal{W}_\psi f(s_2', X(s_2')) \\
&\leq A(s_2')^\gamma \\
&= A\left(\frac{4Ks_2}{(K-C)}\right)^\gamma \\
&= \frac{A4^\gamma}{(K-C)^\gamma}|x-x_0|^\gamma \\
&= B_2|x-x_0|^\gamma.
\end{aligned}$$

3.142 and 3.144 together gives

$$|\mathcal{W}_\psi f(s,x)| \leq B(s^\gamma + |x-x_0|^\gamma). \qquad (3.145)$$

Then, by Theorem 3.4 with (3.137) and (3.135), $f(x)$ is Hölder α at x_0 for any $\alpha < \gamma$.

$\square$

4 WTMM ALGORITHM

4.1 Introduction

In this Chapter we will describe an implementation of a WTMM (Wavelet Transform Modulus Maxima) algorithm to find edges in data and an approximation of the Hölder regularity at the edge-points.

Laser altimetry data collected with an oscillating mirror laser scanner on an airplane scans the ground in a zig-zag pattern along the flight direction. First we rotate the data so that the flight direction was approximately along the y-coordinate axis. To resample these data to a regular grid, we use a *Delauney Triangulation*, linear interpolate on the triangles and resample the triangulated surface to the regular grid we desire. A grid size of 25 by 25 centimeters was used. ‘Outliers’, meaning ‘wild values’ might then result in several data points in the resampled data set, giving the impression that there really *is* a high structure (a tall tree, chimney or tower etc). We do not expect to create an algorithm that completely and automatically analysis datasets in this thesis so we leave these topics unanswered.

In a laser altimetry data set along one direction, buildings are characterized by a step-function when we ‘hit’ the building and an other step function when we ‘leave’ the building, whereas trees behave more like a peak with relatively few samples with approximately the same height. The issue of ‘preprocessing’ the data might be important here but we will not go into that.

In Section 4.2 we describe our data set more accurately. In Section 4.3 we will describe the WTMM algorithm with one line of the data set as example data, and in Section 4.4 we will test the algorithm on synthetic data with known properties, for example step functions, spikes, cusps etc. In Section 4.5 a full ’analysis’ of two laser altimetry datasets will be presented.

4.2 Data Description

Our test data set is from the town of Sandvika, outside Oslo, Norway. The ’raw’ data from Sandvika is a set of 615 000 points organized in columns, where the first column is the latitude, the second is the longitude and the third is the height of the sampled point. The first few lines in the Sandvika file are here:

```
584502.400 6641006.650 119.610
584501.820 6641006.200 119.390
584501.180 6641005.700 119.500
584500.620 6641005.270 119.590
584499.960 6641004.760 119.470
584499.400 6641004.320 119.440
   ...          ...         ... ,
```

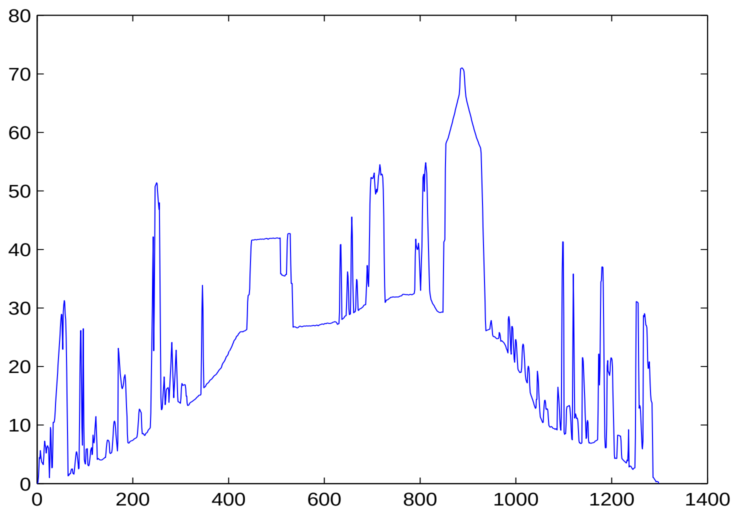

Figure 4.1 Our test signal

where everything is in meters, relative to some coordinate system which we do not worry about. After a rotation of the dataset of 61.8 degrees, the flight direction was approximately along the y-axis, so after subtracting the minimum value from the Xs and the Ys, we get a new, rotated data-set with the origin close to the corner of the data set and all positional values positive but small, but keeping the relative distances between them. The total area covered by the data set is approximately 360 meters wide and 1600 meters along the flight distance.

4.3 WTMM Algorithm

To use the mathematical theory we have been studying so far for edge detection, we selected one (arbitrary) line from the laser altimetry dataset. Then we subtracted a line that intersected the curve in the first and the last sample, to avoid getting large values of $|\mathcal{W}_\psi f(s, x)|$ at each end, which would hide some of the features near the start and the end of the data set. The line is plotted in Figure 4.1.

Theorem 2.54 shows us how to find points x_0 where the function f is regular and where x_0 is surrounded by points where f is possibly less regular. What we want to do, for edge detection, is the opposite; to find points where f is irregular surrounded by points where f is more regular. This means that the conditions in Theorem 2.54 not quite are satisfied. And Theorem 2.55, which introduces the curves to calculate the decay of $|\mathcal{W}_\psi f(s, X(s))|$ on, has very stringent demands on ψ and on the behavior of $\mathcal{W}_\psi f(s, x)$, but we still might use the methods illustrated in these theorems as a tool for edge detection. We decided to use the continuous wavelet transform, cwt, or more precisely, a linear sampling in both the x and s variables of the cwt. This is a more directly approach of the theorems presented in this thesis than the use of the dyadic, discrete wavelet transform, dwt. The various pros and cons of this approach will not be discussed here. The choice of the wavelet ψ of course affects $|\mathcal{W}_\psi f(s, x)|$, but testing

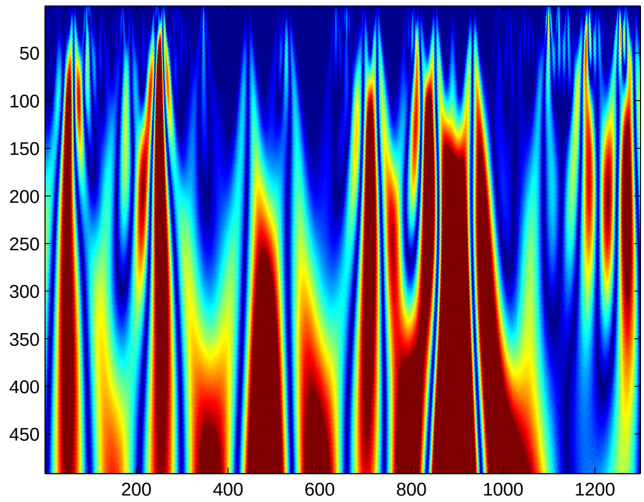

Figure 4.2 The magnitude of $|\mathcal{W}_\psi f(s,x)|$

on both the Haar wavelet and the Mexican Hat wavelet, which represents two wavelets with almost opposite properties (the Haar wavelet being compactly supported and discontinuous and the Mexican Hat wavelet being $C^\infty(\mathbb{R})$ and infinitely supported. See Figure 2.1 on page 14) shows us that the results are similar, though not identical, on a wide range of wavelets. The Mexican Hat wavelet is the one used in the figures in this section.

To implement this algorithm in Matlab, we use the *cwt* function of the Matlab Toolbox for continuous wavelet transform. The intensity and the surface plots of $|\mathcal{W}_\psi f(s,x)|$ are plotted in Figure 4.2 and Figure 4.3. The maxima lines in $|\mathcal{W}_\psi f(s,x)|$ consists of the point cloud of maxima of the one dimensional functions $g(x) = |\mathcal{W}_\psi f(s_0,x)|$ we get when we fix s_0. To actually find these lines in datasets, and especially in real and noisy data offers some challenges. What looks like a nice and easily selectable maximum of $g(x) = |\mathcal{W}_\psi f(s_0,x)|$, often consists of many small peaks, when we zoom in to pixel level, as in Figure 4.4.

To find the 'true' maxima, and since observing that a typical plot of $g(x) = |\mathcal{W}_\psi f(s_0,x)|$ consists of small bumps resembling second or 4th order polynomials on the intervals between zero-crossings of $\mathcal{W}_\psi f(s_0,x)$ as in Figure 4.5, we approximate by a 4. order polynomial using the Matlab function *polyfit*. To catch more than two bumps in intervals between zero-crossings, we would need higher order polynomials.

The set of all the maxima-points, found using the local polynomial approximation on the intervals between the zero-crossings of $g(x)$, resembles lines going from coarse scales that splits into two lines, with a possible gap between the 'splitting point' and the start of the new maxima line, making the 'splits' not into 'Y-s', but into patterns as in figure 4.7. Figure 4.6 shows *all* the points found in the polynomial approximation.

Observing that the zero-crossings, which are the red dots plotted together with the yellow maxima points in Figure 4.8 resembles droplets with a more or less open top and with a

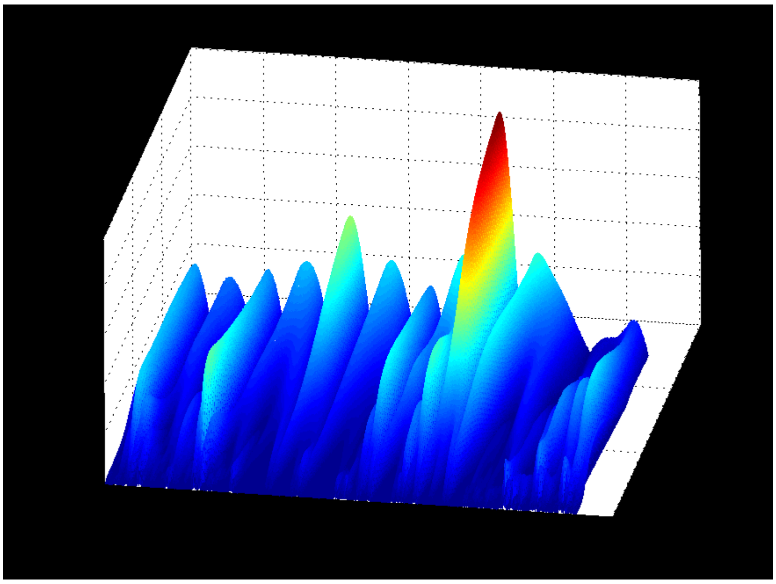

Figure 4.3 The magnitude surface of $|\mathcal{W}_\psi f(s,x)|$

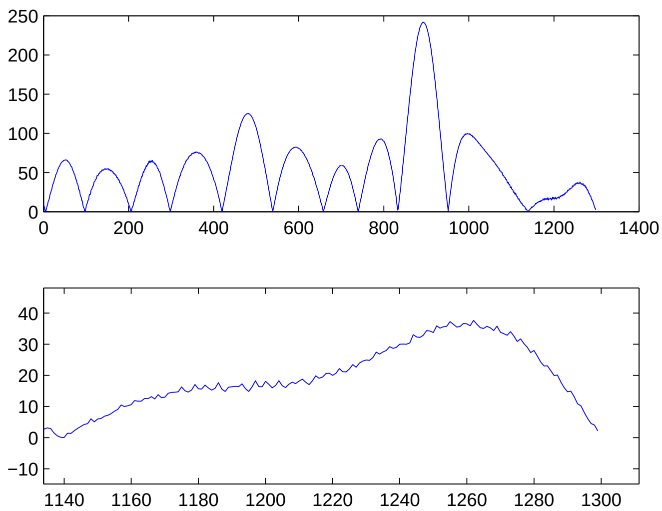

Figure 4.4 Ripples that makes it difficult to find maxima points

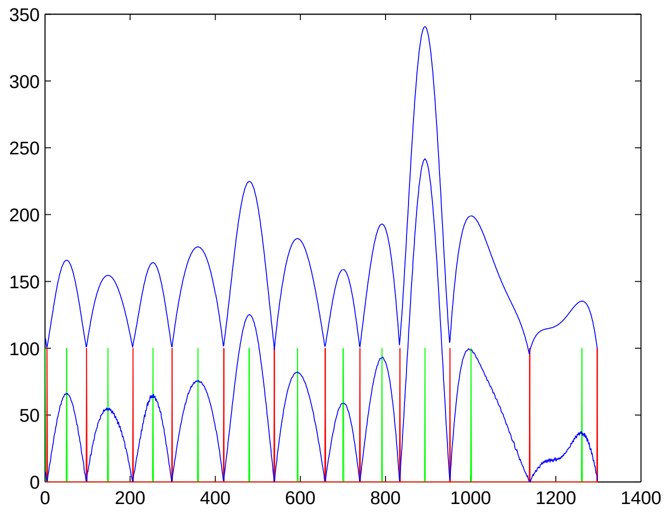

Figure 4.5 *Polynomial approximation (of* $|\mathcal{W}_\psi f(s,x)|+100$*) between the zero-crossings (red lines) gives the maxima (green lines).*

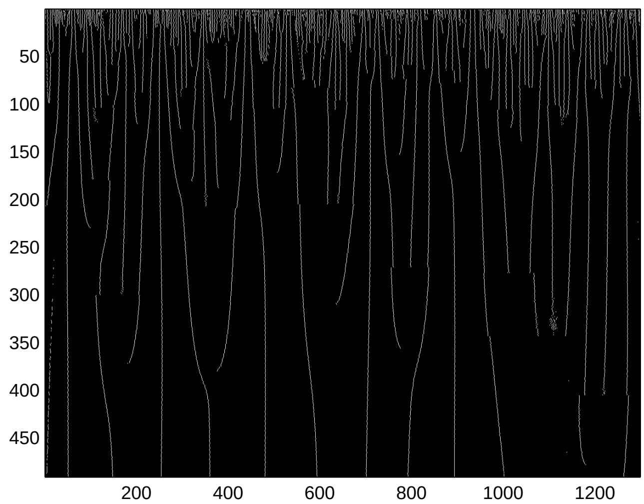

Figure 4.6 *The point clouds of the maxima.*

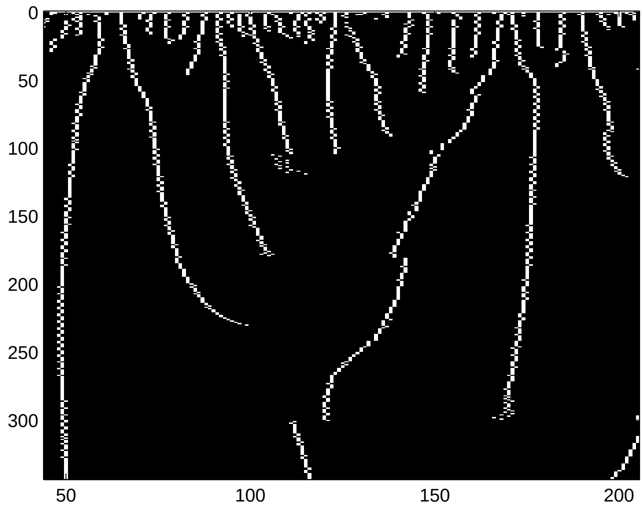

Figure 4.7 The splitting of lines.

maxima line and one or more new such droplets inside, might help us understand the structure of the continuous wavelet transform $\mathcal{W}_\psi f(s, x_0)$.

Now we have a matrix with all the found 'candidates' of points belonging to the maxima-lines of $|\mathcal{W}_\psi f(s, x)|$. The next thing we want to do is to group these points together in line-objects, $X(s)$ on which we can calculate the decay of $|\mathcal{W}_\psi f(s, X(s))|$ when $s \to 0$. This grouping of points into line-objects might cause some challenges. The maxima found are not the exact 'true' maxima of $|\mathcal{W}_\psi f(s, x)|$, and exact *how* we go about in this grouping will affect our final results.

One way to group the points into lines is to start at the coarsest scale (the bigger s), and traverse the maxima-matrix searching indexes representing a maximum. When found, continue from this point toward finer scales as long as there are maxima points, either directly down, meaning at the exact same x, or allowing a certain angle or certain gaps in the line. There probably isn't any totally fool-proof way of doing this. The maxima points found along such 'lines' are marked so that we don't include any maxima points in several lines. The resulting line-objects are stored. We have to decide how long a line has to be, relative the scale, s, where it starts, and how close to the finest scale a line has to reach to be stored as a 'line'. If a line stops before the finest scale, we also have to decide which x this line is to be associated with. This is not handled in a satisfactory way in the current code. According to Proposition 2.53, and given the conditions on ψ therein, the maxima lines should never be interrupted, and they should continue all the way down to the finest scale, but our method of the maxima points and avoiding points to belong more than one maxima line might cause some difficulties here. The result of the grouping of points into contiguous line object is shown in Figure 4.9.

After the grouping of maxima points into connected lines, we are able to use the lines $\{(s, X_i(s))\}$ for estimating the decay of $|\mathcal{W}_\psi f(s, X_i(s))|$ along these lines, as in Figure 4.10.

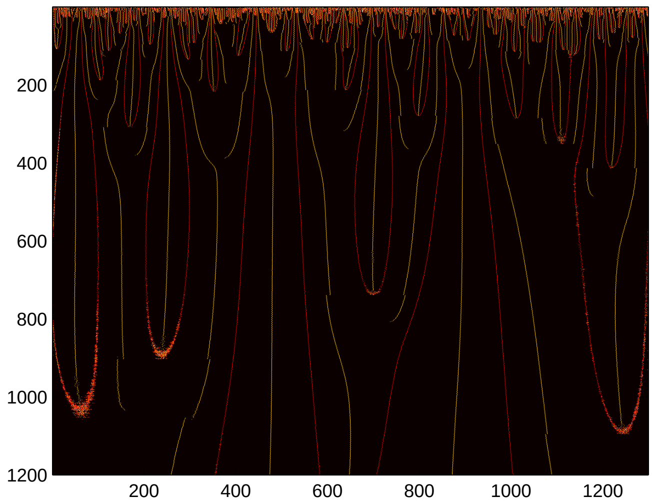

Figure 4.8   *Zero-crossings (red points) and maxima lines (yellow points).*

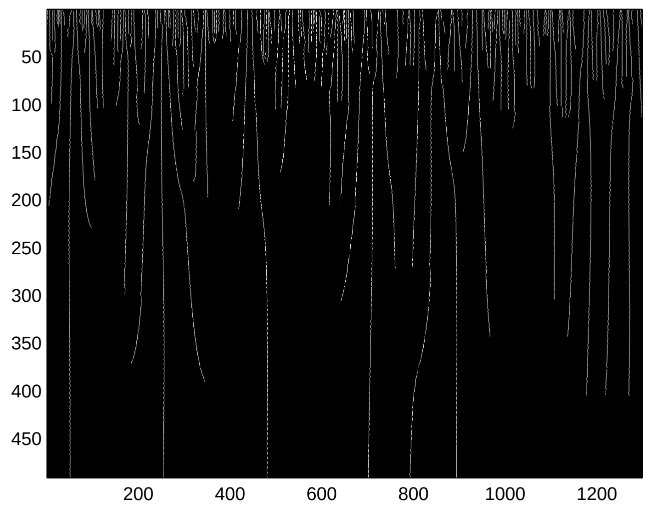

Figure 4.9   *Maxima points connected to lines.*

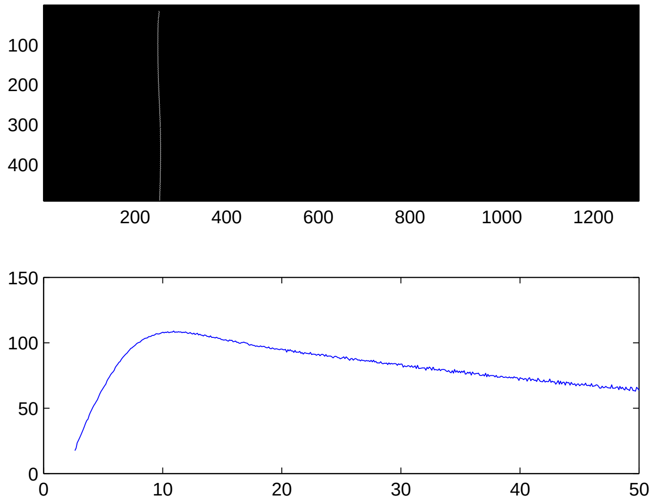

Figure 4.10 *A long line in the maxima matrix and* $|\mathcal{W}_\psi f(s, X(s))|$.

The inequality (2.27) of Theorem 2.55 is what gives us α (which is any $\alpha < \gamma$). To estimate α, we observe that taking logarithms of both sides of $|\mathcal{W}_\psi f(s, X(s))| \leq As^\alpha$ gives us

$$\log(|\mathcal{W}_\psi f(s, X(s))|) \leq \log A + \log(s)\ \alpha \tag{4.1}$$

which reads

$$W(s') \leq B + s'\ \alpha. \tag{4.2}$$

Plotting this $W(s')$, which is a $\log$-$\log$ plot of $|\mathcal{W}_\psi f(s, X(s))|$, and searching for α, which is the slope of this 1. order polynomial bounding $W(s)$ is again a matter of difficulties. Since CWT is undefined at $s = 0$, and since we operate with a discrete sampling of $|\mathcal{W}_\psi f(s, x)|$, this estimation of α will always be a guess, rather than the exact value. The $\log$-$\log$ plot of $|\mathcal{W}_\psi f(s, X(s))|$ blows up the small oscillations of $\mathcal{W}_\psi f(s, X(s))$ at small scales, as shown in Figure 4.13. This makes the estimation of α especially ambiguous when the line $X(s)$ is short, consisting only of the small values of s' and $W(s')$. Instead of searching for a line that is completely above the graph of $W(s')$, we want to find an approximation of the slope of $W(s')$ at small scales. This can be done in several ways, each with some advantages and disadvantages. We ended up using the derivative at the smallest scale of the 2nd degree polynomial approximating the whole of $W(s')$, again using the Matlab function *polyfit*, as in figure 4.11.

If a line consists of few points and at small scales, the $\log$-$\log$ approach and the small number of samples makes the polynomial approximation unstable. What we *really* are searching for is the slope, α of a line that is completely above all the samples when $s \to 0$. Our method gives strange results, as indicated in Figure 4.12 and Figure 4.13.

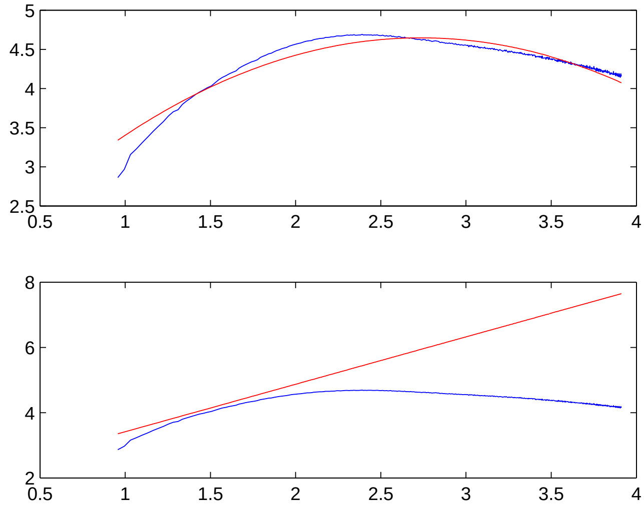

Figure 4.11 *The* log-log *plot of* $|\mathcal{W}_\psi f(s, X(s))|$ *(blue curve) and the 2nd order polynomial approximation (red parabola) and the derivative with slope* α*, representing the* As^α.

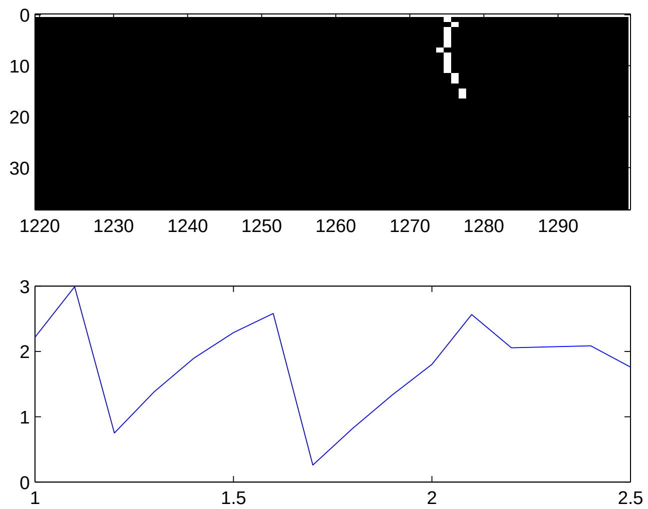

Figure 4.12 *A short line in the maxima matrix and* $|\mathcal{W}_\psi f(s, X(s))|$.

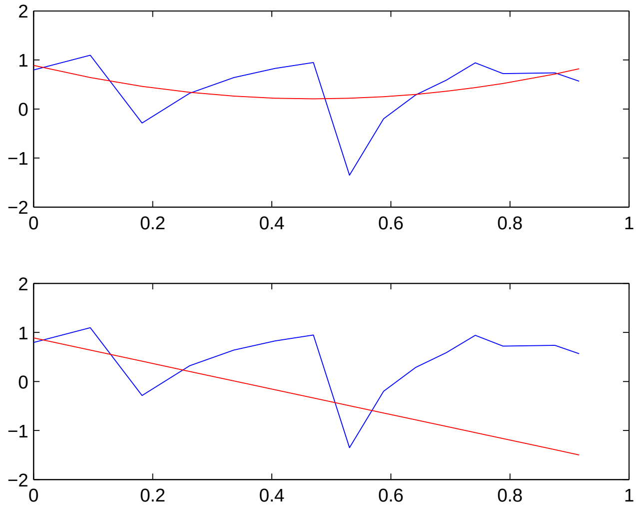

Figure 4.13 The log-log *plot and the α of a short line*

There might be several maxima lines pointing at a particular x_0, and we are interested in the one with the smallest α, since α gives us the upper bound of the regularity of the function $f(x)$ at x_0. Actually, CWT with a complex wavelet might give us the α-regularity at a point both from the left and from the right by following maxima lines that are to the left and to the right of the point, which might differ, as explained in Tu, Hwang (56), but that is beyond the scope here. Following all the stored maxima lines toward finer scales, and storing the minimum α's in a vector at the positions of the original x's, makes us able to plot the original signal $f(x)$ together with the Hölder regularity of each point that has a maxima line pointing at it, as in figure 4.14. The α's plotted are actually the plot of $(25\ (4 - \alpha)\ +\ 140)$, showing small α's as tall bars.

4.4 Synthetic Data

To test the resulting α values, we need to analyse synthetic data with known regularity at the points of interests. We know that a continuous function is Hölder α for $\alpha < 1$, and that a cusp is the limiting case, i.e. it is continuous but it is not Hölder $\alpha = 1$. The functions $f(x) = A\ |x|^\alpha$ for $0 < \alpha < 1$ are Hölder α which follows trivially from the definition of Hölder regularity, as does the function $f(x) = 1 - |x|^\alpha$ for $0 < \alpha < 1$ and $x \in [-1, 1]$. Since our data is discretely sampled, strictly speaking we do not know whether the original signal is $C^\infty(\mathbb{R})$ or discontinuous at any point, but the best we get is what it *looks like* at the finest scale we have at our disposal. In this section we will be analysing several interesting synthetic cases, and see what α-s our WTMM-algorithm gives us.

In the following, we present some examples of synthetic signals with particularly interesting properties. The first example represents an abrupt change in the level of a function, a ‘step’, plotted in Figure 4.15 and Figure 4.16. The *Analysis Plots* of all these examples consists of 5

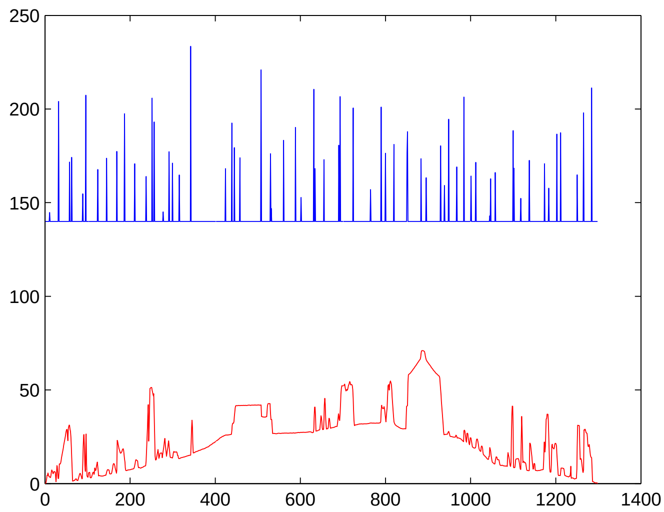

Figure 4.14 *Signal (Red curve) and Hölder regularity (*$(25\ (4-\alpha)\ +\ 140)$ *blue bars)*

parts:

1. The signal.
2. The α-values.
3. $|\mathcal{W}_\psi f(s,x)|$ as grey-level.
4. All the maxima points.
5. The maxima lines found by grouping of maxima points.

The next example is two such steps, which could represent a building, plotted in Figure 4.17 and Figure 4.18.

The third example is a 'spike'. A spike which is a sudden 'wild value', which could be a tree, but which also could be an outlier, plotted in Figure 4.19 and Figure 4.20.

Then, in the next two examples, we have cusps with known non-integer α-regularity, which would be the ultimate test of our algorithm for finding these α-s. The result is plotted in Figure 4.21 and Figure 4.22 and in Figure 4.23 and Figure 4.24.

4.5 The Sandvika Data Set

Running the algorithm described in the previous sections on all the lines and all the columns in the Sandvika data set in Figure 4.25 gives the results in Figure 4.26.

The histogram with 100 equally spaced containers of the alpha values of all the rows is showed in Figure 4.28, and the same for all the columns is showed in Figure 4.27. The value $\alpha = 0$ is

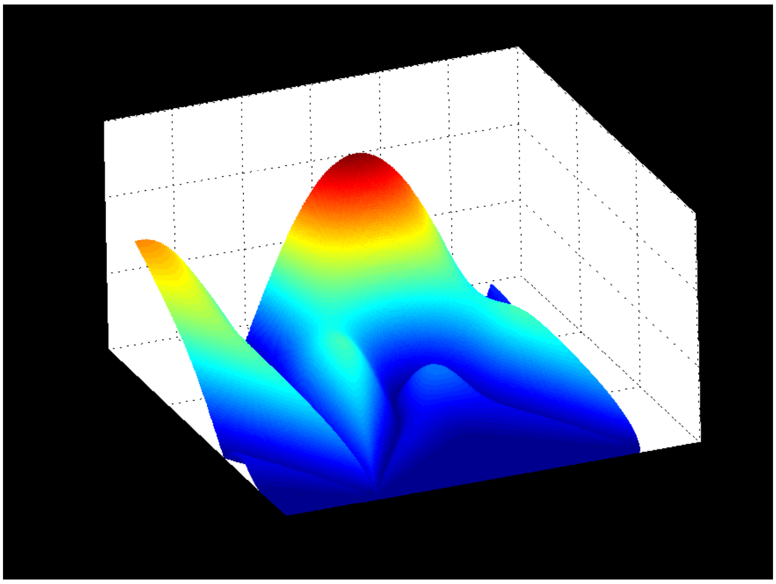

Figure 4.15 $|\mathcal{W}_\psi f(s,x)|$ *of a step function*

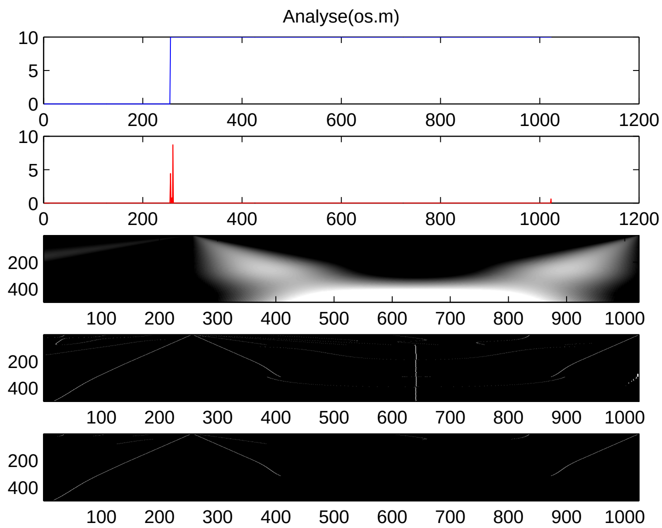

Figure 4.16 *The analysis of the step function*

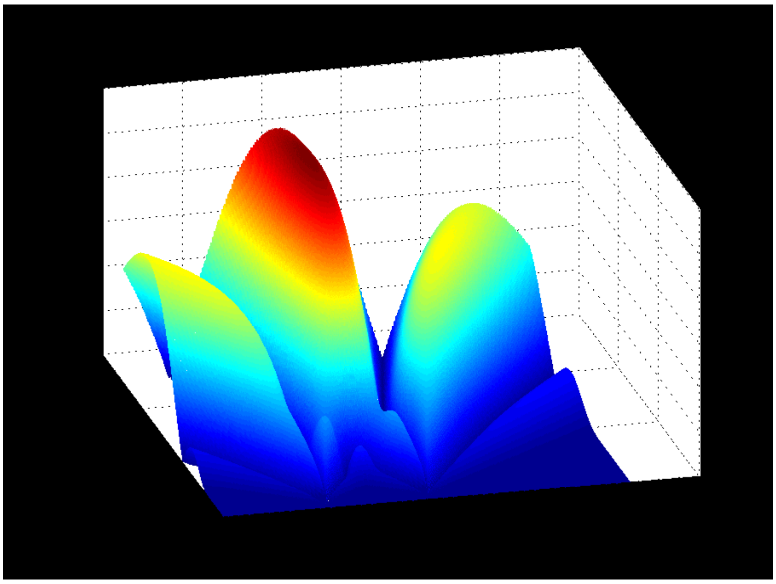

Figure 4.17 $|\mathcal{W}_\psi f(s,x)|$ *of a box function*

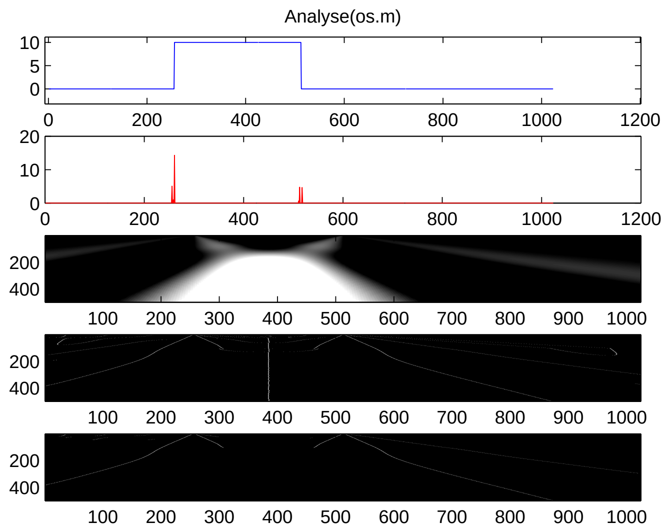

Figure 4.18 *The analysis of the box function*

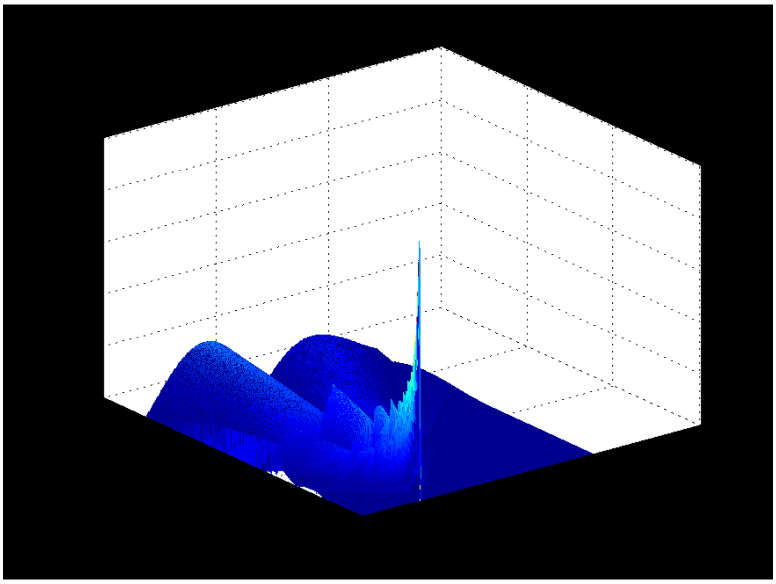

Figure 4.19 $|\mathcal{W}_\psi f(s,x)|$ *of a spike function*

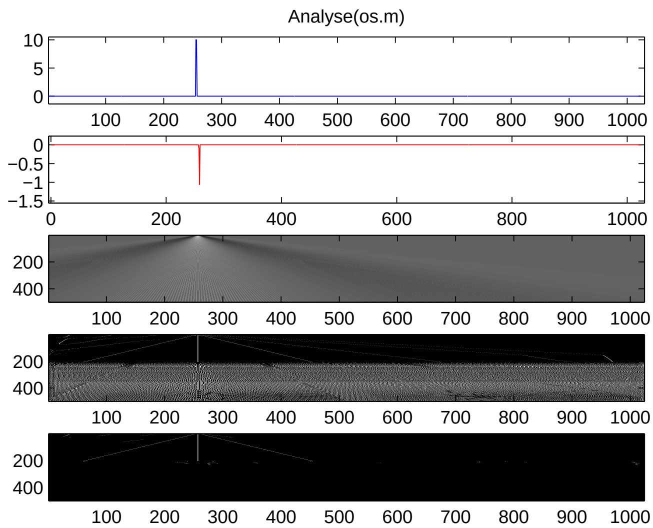

Figure 4.20 *The analysis of the spike function*

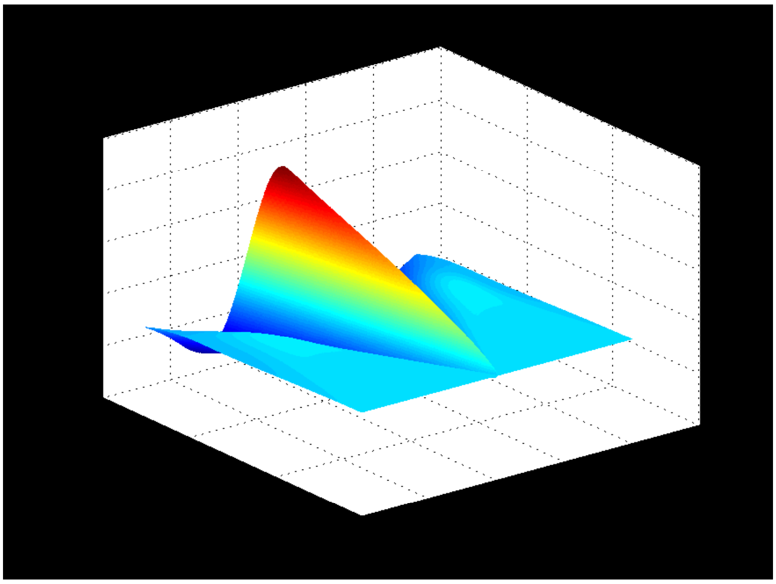

Figure 4.21 $|\mathcal{W}_\psi f(s,x)|$ *of a cusp (*$\alpha = 0.25$*) function*

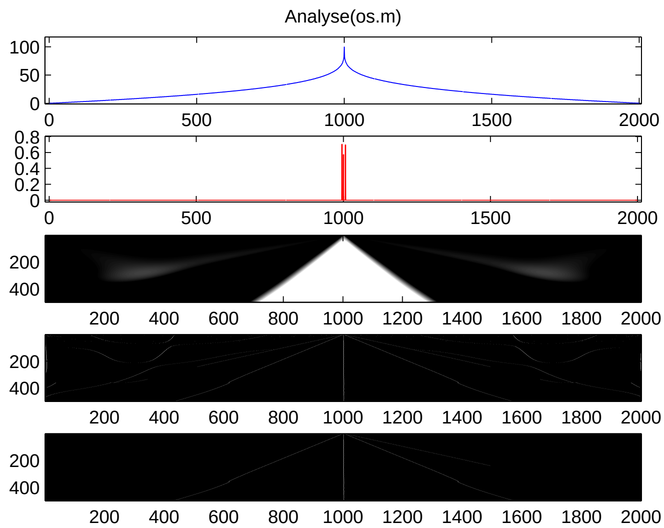

Figure 4.22 *The analysis of the cusp (*$\alpha = 0.25$*) function*

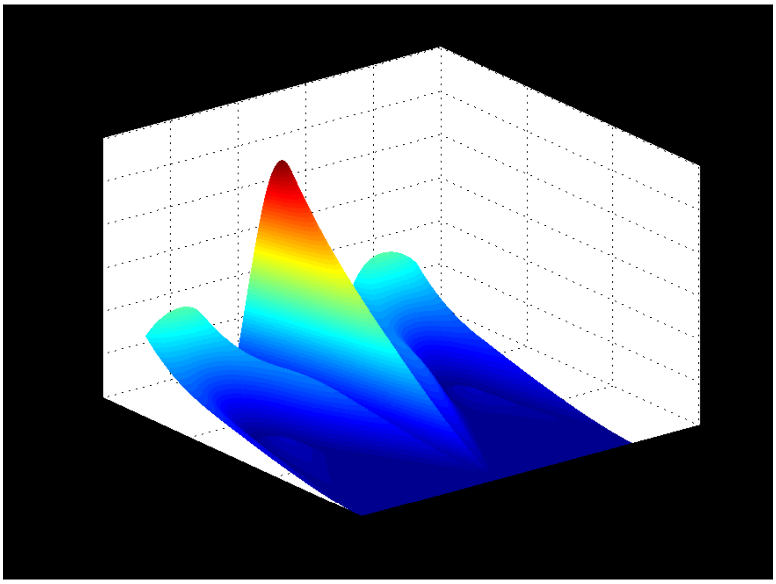

Figure 4.23 $|\mathcal{W}_\psi f(s,x)|$ *of a cusp (*$\alpha = 0.75$*) function*

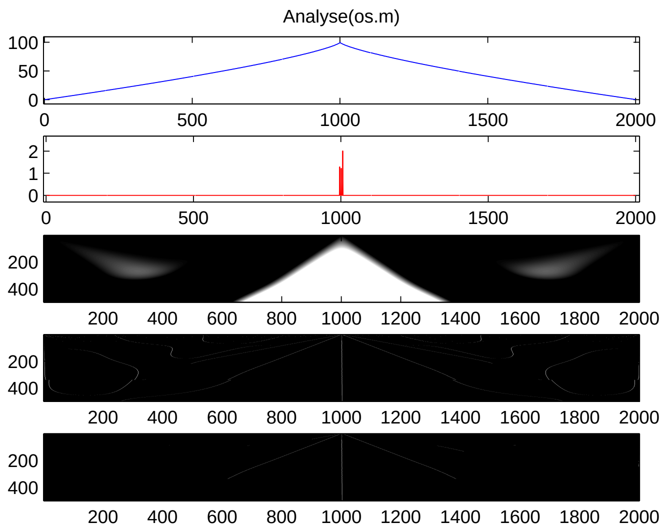

Figure 4.24 *The analysis of the cusp (*$\alpha = 0.75$*) function*

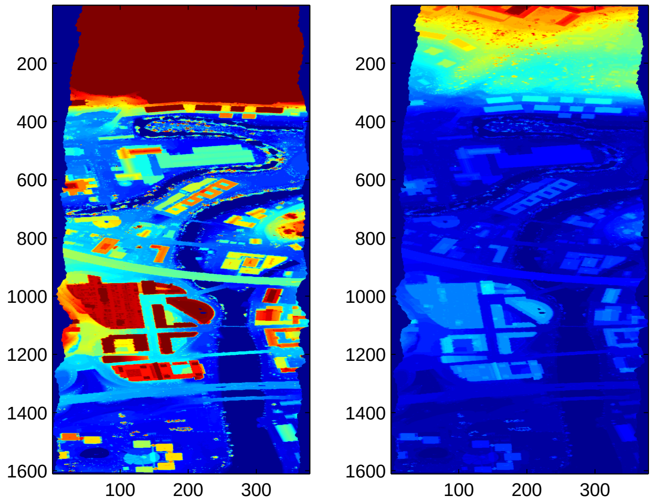

Figure 4.25 The Sandvika data set with two different scalings

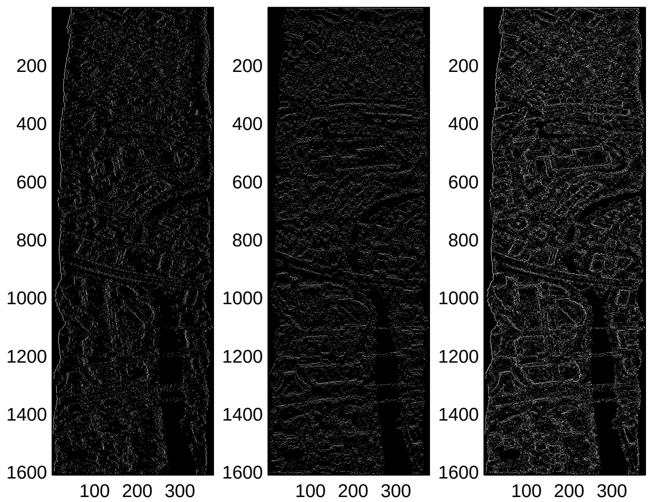

Figure 4.26 All the alpha values found along maxima lines in the rows, the columns and either in the dataset

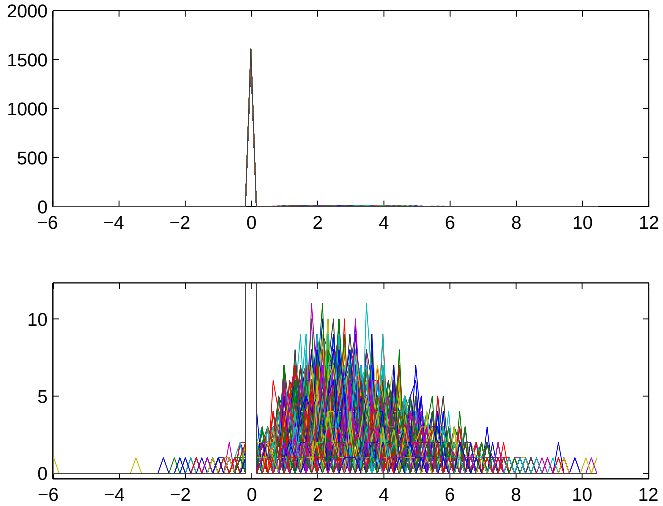

Figure 4.27 All the alpha values found along maxima lines in all the columns

used for an initial value of all the points, and is therefore (probably!) all the points with *no* maxima line pointing at them

We see that the alpha values range over a wide interval, including negative values, which shows that our algorithm does not work completely.

4.6 Conclusions

The example functions with known Hölder regularity in Section 4.4 shows us that our algorithm has some weaknesses. The most serious such is probably the process of finding and gathering the maxima points of the function $g(x) = |\mathcal{W}_\psi f(s_0, x)|$ for all x_0 into contiguous line objects. Failing in finding all the *true* maxima lines and 'following' them all the way down to the right scale and position (s_0, x_0) or by making 'false' maxima lines will make the algorithm produce strange results. Also the process of approximating the slope of the $\log - \log$ curve of $g(s) = |\mathcal{W}_\psi f(s, X(s))|$ when the curve consists of few points, or if it is oscillating, offers some challenges. This is particularly so if we have made some wrong choices for the maxima lines.

5 THESIS SUMMARY

We have seen that mathematically, the CWT with surprisingly weak conditions on ψ gives us a tool for approximating α-regularity locally and globally. But the discretization of CWT (and of the data to be analysed) for computer application offers some challenges. For a more thorough analysis of 2D data, similar results should be studied with 2 dimensional CWT, or possibly the 2 dimensional dyadic discrete wavelet transform. Also, the theorems studied here only

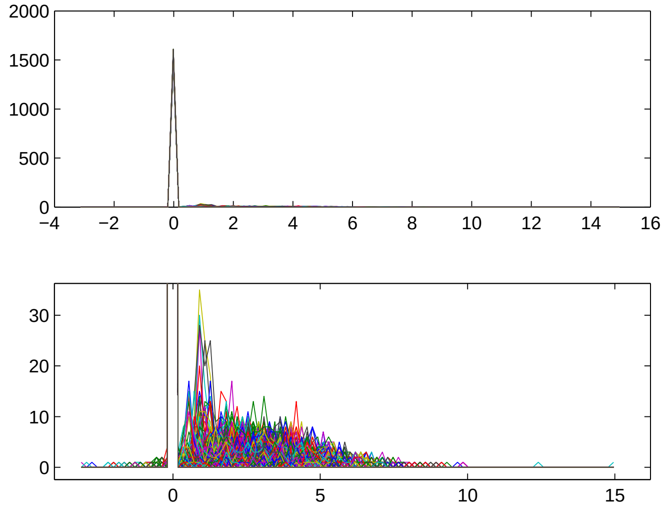

Figure 4.28   All the alpha values found along maxima lines in all the rows

considers the *regularity*, but in real datasets the noise will usually make the signal very ‘un-regular’ even though the noise ‘structures’ are very small compared to the real objects or structures in the data.

Also the *amplitude* of the wavelet transform should be considered. For instance, a paved road with width 3 meters and height 10 centimeters would give us the same results as a building with the same width but with height 10 meters, o nly that the house will have 100 times larger $|\mathcal{W}_\psi f(s,x)|$. But the α-values here will be the same.

Furthermore, the *scale* at which the maxima lines occur or have some sort of maxima or center shows us the extent of the structure, which normally would be very interesting information to analyse further, since we for instance normally would want to separate a matchbox from a large building...

Different kinds of thresholding, which means altering the value of $\mathcal{W}_\psi f(s,x)$ according to some rule, for instance zeroing out $\mathcal{W}_\psi f(s_0,x_0)$ where $|\mathcal{W}_\psi f(s_0,x_0)|$ is smaller than some threshold or *soft thresholding* which means that we just reduce the values instead of zeroing it, is also a useful tool in analysis of data. Also stopping at a scale larger than the smallest scale “filters out” the high frequency contribution, which often is noise or structures to small for our interest.

APPENDIX

A PRELIMINARIES

A.1 Introduction

In this appendix, we want to include all the background material needed in the thesis not included in the text. This is also the place to find general definitions and notations etc. Most of the material is from Folland (11), Pedersen (51), from Richards, Youn (52) or from either of the three major sources in the rest of the thesis: Mallat, Hwang (39), Holschneider (14) and Mallat (37).

A.2 Integration Theory

Definition A.1 (The Power Set $\mathcal{P}(\mathbb{R}^n)$).

$\mathcal{P}(\mathbb{R}^n) = \{A | A \subset \mathbb{R}^n\}$.

Definition A.2 (σ-algebra). *A* (set-) σ-algebra *in a non-empty set X is a family $\mathcal{A} \subset \mathcal{P}(X)$ which is closed under countable unions and compliments.*

Definition A.3 (The σ-algebra generated by $\mathcal{E}$). *Let $\mathcal{E} \subset \mathcal{P}(X)$ be a subset. The smallest σ-algebra which contains $\mathcal{E}$ is the σ-algebra* generated *by $\mathcal{E}$, and is written $\mathcal{M}(\mathcal{E})$.*

Definition A.4 (Borel σ-algebra). *The σ-algebra generated by the open (or equivalently by the closed) subsets of X is called the* Borel σ-algebra *on X and is denoted $\mathcal{B}_X$.*

Lemma A.5. *If $f : X \to Y$ is a function and $\mathcal{N} \subset \mathcal{P}(Y)$ is a σ-algebra on Y, then $\mathcal{M} = \{f^{-1}(E) : E \in \mathcal{N}\}$ is a σ-algebra on X.*

Proof. This is obvious, since $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ commutes with unions, intersections and complements. □

Definition A.6 (Measure, Measurable space, Measure space). *A* measure *on $(X, \mathcal{M})$ is a function $\mu : \mathcal{M} \to \mathbb{R}_+ = [0, \infty) \subset \mathbb{R}$ such that*

- $\mu(\emptyset) = 0$,
- $\{E_j\}_{j=1}^{\infty}$ *disjoint in $\mathcal{M}$* $\Rightarrow \mu(\cup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu(E_j)$.

A measurable space, $(X, \mathcal{M})$*, is a set X equipped with a σ-algebra $\mathcal{M}$.*
A measurable space $(X, \mathcal{M})$ with a measure μ is called a measure space.

Definition A.7 (Null-set, Complete measure). *A* null-set *is a measurable set X, where $\mu(X) = 0$. A measure whose domain contains all subsets of null-sets is called* complete.

Theorem A.8. *Suppose $(X, \mathcal{M}, \mu)$ is a measure space. Let $\mathcal{N}$ be the set of null-sets in $\mathcal{M}$, i.e. $\mathcal{N} = \{N \in \mathcal{M} | \mu(N) = 0\}$. Define $\overline{\mathcal{M}} = \{E \cup F | E \in \mathcal{M}$ and $F \subset N$ for some $N \in \mathcal{N}\}$. Then $\overline{\mathcal{M}}$ is a σ-algebra, and there is a unique extension $\overline{\mu}$ of μ to a complete measure on $\overline{\mathcal{M}}$. $\overline{\mu}$ is called the* completion *of μ.*

Proof. Since $\mathcal{M}$ and $\mathcal{N}$ are closed under countable unions, so is $\overline{\mathcal{M}}$. If $E \cup F \in \overline{\mathcal{M}}$ where $F \subset N \in \mathcal{N}$, we can assume that $E \cap N = \emptyset$ (otherwise, replace F, N by $F \backslash E,\ N \backslash E$). Then $E \cup F = (E \cup N) \cap (N^c \cup F)$, so $(E \cup F)^c = (E \cup N)^c \cup (N \backslash F)$. But $(E \cup N)^c \in \mathcal{M}$ and $N \backslash F \subset N$, so that $(E \cup F)^c \in \overline{\mathcal{M}}$. Thus $\overline{\mathcal{M}}$ is a σ-algebra.

If $(E \cup F) \in \overline{\mathcal{M}}$ $(F \subset N \in \mathcal{N})$, set $\overline{\mu}(E \cup F) = \mu(E)$. This is well defined, since if $(E_1 \cup F_1) = (E_2 \cup F_2)$ $(F_j \subset N_j \in \mathcal{N})$, then $E_1 \subset E_2 \cup N_2$ and so $\mu(E_1) \leq \mu(E_2) + \mu(N_2) = \mu(E_2)$, and likewise $\mu(E_2) \leq \mu(E_1)$. It is easily verified that $\overline{\mu}$ is a complete measure on $\overline{\mathcal{M}}$, and that $\overline{\mu}$ is the only measure on $\overline{\mathcal{M}}$ which extends μ. □

Definition A.9 (Borel measure). *The measure generated by each of the following:*

- $\mathcal{E}_1 = \{(a, b) : a < b\}$,
- $\mathcal{E}_2 = \{[a, b] : a < b\}$,
- $\mathcal{E}_3 = \{(a, b] : a < b\}$,
- $\mathcal{E}_4 = \{[a, b) : a < b\}$,
- $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$,
- $\mathcal{E}_6 = \{(-\infty, b) : b \in \mathbb{R}\}$,
- $\mathcal{E}_7 = \{[a, \infty) : a \in \mathbb{R}\}$,
- $\mathcal{E}_8 = \{(-\infty, b] : b \in \mathbb{R}\}$.

is called the Borel measure, *and is denoted $\mathcal{B}_{\mathbb{R}}$.*

Definition A.10 (Lebesgue and Lebesgue-Stieltjes measure).
If $F : \mathbb{R} \to \mathbb{R}$ is an increasing, right continuous function (i.e. $F(a) = \lim_{x \to a} F(x), \forall a \in \mathbb{R}$), then the completion of the measure μ_F defined on $\mathcal{B}_{\mathbb{R}}$ by $\mu((a, b]) = F(b) - F(a)$ is called the Lebesque-Stieltjes measure *associated to F. The complete measure associated to $F(x) = x$ is called the* Lebesgue measure *on $\mathbb{R}$ and is denoted m. The domain of m is denoted $\mathcal{L}$.*

Definition A.11 (Measurable functions). *If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable spaces, a mapping $f : X \to Y$ is called* $(\mathcal{M}, \mathcal{N})$-measurable, *(or just* measurable*) if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.*

Definition A.12 (Lebesgue measurable function). *A function $f : \mathbb{R} \to \mathbb{R}$ is called* Lebesgue measurable *if it is $(\mathcal{L}, \mathcal{B}_{\mathbb{R}})$-measurable.*

Proposition A.13. *If $\mathcal{N}$ is a σ-algebra generated by $\mathcal{E}$, then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$.*

Proof. The "if" part follows from the fact that $\{E \subset Y : f^{-1}(E) \in \mathcal{M}\}$ is a σ-algebra which contains $\mathcal{E}$, and therefore contains $\mathcal{N}$.
The "only if" implication is trivial. □

Corollary A.14. *Let $f : X \to \mathbb{R}$ be a function, where $(X, \mathcal{M})$ is a measurable space. The following are equivalent:*

- *f is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable,*
- *$f^{-1}((a, \infty)) \in \mathcal{M}$, $\forall a \in \mathbb{R}$,*
- *$f^{-1}([a, \infty)) \in \mathcal{M}$, $\forall a \in \mathbb{R}$,*
- *$f^{-1}((-\infty, a)) \in \mathcal{M}$, $\forall a \in \mathbb{R}$,*
- *$f^{-1}(-\infty, a])) \in \mathcal{M}$, $\forall a \in \mathbb{R}$.*

Proof. This follows trivially from the definition of the Borel σ-algebra. □

Definition A.15 (Characteristic- and Simple functions). *The* characteristic function $\chi_A(x)$ *of a set A is defined*

$$\chi_A(x) = \begin{cases} 1 & \text{for } x \in A, \\ 0 & \text{for } x \notin A \end{cases}$$

A simple function *is a finite linear combination of characteristic functions of sets in $\mathcal{M}$, $f(x) = \sum_{n=0}^{N} a_n \chi_{E_n}(x)$.*

Theorem A.16. *Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to (0, \infty]$ is measurable, there is a sequence $\{\phi_n\}$ of simple functions such that $0 \le \phi_1 \le \phi_2 \le \cdots \le f$, $\phi_n \to f$ pointwise and $\phi_n \to f$ uniformly on any set on which f is bounded.*

Proof. We prove this by constructing the sequence ϕ_n.
Let $n \in \mathbb{N}$, and $0 \le k \le (2^{2n} - 1)$. Define

$$E_n^k = f^{-1}(\,(\frac{k}{2^n}, \frac{k+1}{2^n}]),$$

and

$$F_n = f^{-1}(\,(2^n, \infty]).$$

Define

$$\phi_n(x) = \sum_{k=0}^{(2^{2n}-1)} \frac{k}{2^n} \chi_{E_n^k}(x) + 2^n \chi_{F_n}(x).$$

Then $\phi_n \le \phi_{n+1}$ for all n, and $0 \le f - \phi_n \le 2^{-n}$ on the set where $f(x) \le 2^n$. □

Definition A.17 (Lebesgue integral). *Let* $s(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x)$ *be a simple function, where* $\{E_i\}_{i=0}^{N}$ *is a partition of a measurable set* E*, with measure* μ*. Define* $I_E(s) = \sum_{i=1}^{n} c_i \mu(E \cap E_i)$*. We then define the* Lebesgue integral *of a measurable positive function* f *over the set* E *as*

$$\int_E f \; d\mu = \sup_{0 \leq s \leq f} I_E(s).$$

Theorem A.18 (Monotone convergence theorem). *If* $\{f_n\} : X \to [0, \infty]$ *are measurable,* $0 \leq f_n \leq f_{n+1}$ *and*$f_n \to f$*, then*

$$\int f \; d\nu = \lim_{n\to\infty} \int f_n d\nu.$$

Proof. $\{\int f_n \; d\nu\}$ is monotone, so the limit does exist (possibly ∞). $\int f_n \leq \int f \Rightarrow \lim \int f_n \leq \int f$. Let $0 < \alpha < 1$ and let $0 \leq \phi \leq f$ be a simple function. Define $E_n = \{x \in E | \; f_n(x) \geq \alpha\phi(x)\}$. Then $E_n \subset E_{n+1} \to E$ and

$$\int_E f_n \geq \int_{E_n} f_n \geq \alpha \int_{E_n} \phi \underset{n\to\infty}{\longrightarrow} \alpha \int_E \phi \underset{\alpha\to 1}{\longrightarrow} \int_E \phi \underset{\sup(\phi| \; \phi \text{ simple})}{\longrightarrow} \int_E f.$$

$\square$

Lemma A.19 (Fatou's lemma). *If* $\{f_n\}$ *is measurable and* $0 \leq f_n(x) \leq \infty, \; \forall x$*, then*

$$\int (\liminf f_n) \leq \liminf \int f_n.$$

Proof. $\int \inf f_n \leq \inf \int f_n$. By the Monotone Convergence theorem: $\int (\liminf f_n) = \lim \int (\inf f_n) \leq \liminf \int f_n$ $\square$

If we have a function which is complex or real and negative we define

$$f_R^+(x) = \max(Re(f), 0),$$

$$f_R^-(x) = \max(-Re(f), 0),$$

$$f_I^+(x) = \max(Im(f), 0),$$

$$f_I^-(x) = \max(-Im(f), 0).$$

All of these are positive, real-valued functions and

$$f(x) = f_R^+(x) - f_R^-(x) + i \left(f_I^+(x) - f_I^-(x) \right).$$

A.3 General Theory

Definition A.20 (Convolution). *The* convolution $f * g$ *of two functions* $f, g \in L^2(\mathbb{R})$ *is defined by*

$$f * g(x) = \int_{\mathbb{R}} f(u)g(x-u)du. \tag{A.1}$$

Definition A.21 (Involution, Dilation and Translation). *Let* $f(x)$ *and* $\psi(x)$ *be functions. Then*

- *The* involution f^{inv} *of* f *is defined by*

  $$f^{inv}(x) = f(-x) \tag{A.2}$$

- $f_s(x)$ *denotes the* dilation *of* $f(x)$ *by the factor* s*:*

  $$f_s(x) = \frac{1}{s} f(x/s). \tag{A.3}$$

  Then $||f||_{L^1(\mathbb{R})} = ||f_s||_{L^1(\mathbb{R})}$ *and* $||f||_{L^2(\mathbb{R})} = \sqrt{s}||f_s||_{L^2(\mathbb{R})}$.

- $\psi^{a,b}(x)$ *denotes the* dilation and translation *of* ψ *by the factors* a *and* b*;*

  $$\psi^{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) \tag{A.4}$$

  Then $||f^{a,b}||_{L^2(\mathbb{R})} = ||f||_{L^2(\mathbb{R})}$ *and* $||f^{a,b}||_{L^1(\mathbb{R})} = \sqrt{|a|}||f||_{L^1(\mathbb{R})}$.

For wavelets with compact support, the Cone of Influence is important, because it gives the cone in the time-frequency plane of the continuous wavelet transform, pointing at x_0 which influences the wavelet transform at x_0.

Definition A.22 (The Cone of Influence). *We define the* Cone of Influence *of a point* x_0*, for a constant* C *by*

$$\mathrm{Cone}(x_0, C) = \{(s, x) \in \mathbb{R}^2 \mid |x - x_0| \leq Cs\}.$$

A.4 Vector Spaces

Definition A.23 (Abelian group). *An* Abelian group $(G, +)$ *is a set* G *with a binary operation* $+$ *(*$a, b \in G \Rightarrow a + b \in G$*) on* G*, such that the following are satisfied:*

- *The binary operation is associative and commutative.*
  ($a + (b + c) = (a + b) + c$ *and* $a + b = b + a, \ \forall a, b, c \in G$*)*

- *There exist an element $e \in G$ (an* identity*) such that $e + x = x, \ \forall x \in G$*
- $\forall a \in G, \ \exists a' \in G$*, such that $a + a' = e$. (an* inverse*)*

Definition A.24 (Vector space, Vectors, Span, Basis). *A real (or complex)* vector space *is an Abelian group V, with addition as the binary operation and an operation of scalar multiplication of each element in V with each element in $\mathbb{R}$ (or $\mathbb{C}$), such that for all $a, b \in \mathbb{R}$ (or $\mathbb{C}$) and all $\alpha, \ \beta \in V$:*

- $a\alpha \in V$
- $a(b\alpha) = (ab)\alpha$
- $(a + b)\alpha = (a\alpha) + (b\alpha)$
- $a(\alpha + \beta) = (a\alpha) + (a\beta)$
- $1\alpha = \alpha$

The elements of V are called vectors. *If $\{a_n\}_{n\in F} \subset V$ for an index set F, we define the* span *of $\{a_n\}_{n\in F}$, $Span(\{a_n\})$, as*

$$Span(\{a_n\}) = \left\{ a \mid a = \sum_{n\in F} c_n a_n, \text{ for some } \{c_n\}_{n\in F} \subset \mathbb{R} \text{ (or } \mathbb{C}\text{)} \right\}.$$

A basis *for V is such a set $\{a_n\}_{n\in F} \subset V$, such that $V = Span(\{a_n\}_{n\in F})$, and the set $\{c_n\}_{n\in F}$ is unique for each $a \in V$.*

Definition A.25 (Norm, Normed Vector space). *A* norm *on a vector space V is a function $\| \cdot \| : V \to [0, \infty)$ such that, for all $x, y \in V$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$),*

- $||x + y|| \leq \|x\| + \|y\|$
- $\|\lambda x\| = |\lambda| \ \|x\|$
- $\|x\| = 0 \Leftrightarrow x = 0$ *(an identity element of V)*

A vectorspace with a norm is called a normed vector space.

Definition A.26 (The L^p-spaces). *Let X be a measure space with σ-algebra $\mathcal{M}$ and measure μ and let $0 < p < \infty$. Define the equivalence relation $\sim$ by*

$$f \sim g \Leftrightarrow \int_X |f - g|^p \, d\mu = 0.$$

Let $||f||_p = \left(\int_X |f|^p \, d\mu\right)^{1/p}$, where $\int$ is the Lebesgue integral. Then

$$L^p(X, \mathcal{M}, \mu) = L^p(X) = \{f : X \to \mathbb{C} | \ f \text{ is measurable and } ||f||_p < \infty\}/\sim .$$

Definition A.27 (Cauchy sequence, Complete vector space). *A sequence* $\{x_n\}_{n\in\mathbb{N}}$ *in a vector space* V *is called a* Cauchy sequence *if* $\|x_n - x_m\| \to 0$ *when* $n, m \to \infty$. *A vectorspace* V*, where every Cauchy sequence converges to a point in* V *is called a* complete vector space.

Definition A.28 (Banach Spaces). *A normed vector space, which is complete with respect to the norm is called a* Banach space.

The $L^p(\mathbb{R})$-spaces, for $p \geq 1$, are Banach spaces, with the norm given by

$$\|f\|_p = \left(\int |f(x)|^p \, dx \right)^{(1/p)}.$$

Definition A.29 (Inner product, Inner product space). *Let* X *be a complex vector space. An* inner product *on* X *is a map* $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$, $(x, y) \mapsto \langle x, y \rangle$ *such that:*

- $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$ *for all* $x, y, z \in X$ *and* $a, b \in \mathbb{C}$.
- $\langle x, y \rangle = \overline{\langle y, x \rangle}, \ \forall x, y \in X.$
- $\langle x, x \rangle > 0, \ \forall x \in X, x \neq 0.$

Every inner product defines a norm given by $\|x\| = \langle x, x \rangle^{1/2}$. *A complex vector space with an inner product and a norm defined by the inner product is called an* inner product space.

Definition A.30 (Orthogonal/Orthonormal vectors/set). *We say that two vectors* $x, y \in X$ *are* orthogonal *if* $\langle x, y \rangle = 0$. *This is denoted* $x \perp y$. *If, in addition* $\langle x, x \rangle = \langle y, y \rangle = 1$ *we say that* x *and* y *are* orthonormal. *If* $\{x_\alpha\}_{\alpha\in\Lambda}$ *satisfies* $\alpha \neq \beta \Rightarrow \langle x_\alpha, x_\beta \rangle = 0$ *and* $\langle x_\alpha, x_\alpha \rangle = 1, \ \forall \alpha, \ \beta \in \Lambda$ *then* $\{x_\alpha\}$ *is an* orthonormal set.

Lemma A.31 (The Schwarz inequality). *Let* X *be a innerproduct space, with* $x, y \in X$ *and* $a \in \mathbb{C}$. *Then*

- $|\langle x, y \rangle| \leq ||x|| \ ||y||$
- $|\langle x, y \rangle| = ||x|| \ ||y|| \Leftrightarrow x = ay.$

Proof.

- Let $y \neq 0$ and $\lambda = \frac{\langle x,y \rangle}{||y||^2}$.

Then

$$\begin{aligned}0 &\leq ||x-\lambda y||^2\\ &= \langle x-\lambda y, x-\lambda y\rangle\\ &= \langle x,x\rangle - \langle x,\lambda y\rangle - \langle \lambda y, x\rangle + |\lambda|^2\langle y,y\rangle\\ &= \langle x,x\rangle - \overline{\lambda}\langle x,y\rangle - \lambda\overline{\langle x,y\rangle} + |\lambda|^2\langle y,y\rangle\\ &= \langle x,x\rangle - \frac{\overline{\langle x,y\rangle}}{||y||^2}\langle x,y\rangle - \frac{\langle x,y\rangle}{||y||^2}\overline{\langle x,y\rangle} + \frac{|\langle x,y\rangle|^2}{||y||^4}\langle y,y\rangle\\ &= ||x||^2 - 2\frac{|\langle x,y\rangle|^2}{||y||^2} + \frac{|\langle x,y\rangle|^2}{||y||^2}\\ &= ||x||^2 - \frac{|\langle x,y\rangle|^2}{||y||^2}.\end{aligned}$$

Hence $\frac{|\langle x,y\rangle|^2}{||y||^2} \leq ||x||^2$.

- Equality holds in the previous calculations if and only if $x = \lambda y$.

□

Definition A.32 (Hilbert Spaces). *A complex vector space with an inner product, which is complete with respect to the norm* $||x|| = \sqrt{\langle x,x\rangle}$ *is called a* Hilbert space.

$LtoR$ is a Hilbert space, with the innerproduct given by

$$\langle f,g\rangle = \int f(x)\ \overline{g(x)}\ dx.$$

$\mathbb{R}^n$ and $\mathbb{C}^n$ are also a Hilbert spaces, with the usual inner products.

Lemma A.33 (The parallelogram law). *Let* x, y *be two elements in a Hilbert space* $\mathcal{H}$. *Then*

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2).$$

Proof.

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle\\ &= \left(\|x\|^2 + 2\Re e\langle x,y\rangle + \|y\|^2\right) +\\ &\quad\ \left(\|x\|^2 - 2\Re e\langle x,y\rangle + \|y\|^2\right)\\ &= 2\left(\|x\|^2 + \|y\|^2\right).\end{aligned}$$

□

Theorem A.34 (The Pythagorean theorem). *Let* $\mathcal{H}$ *be a Hilbert space. If* $\{x_i\}_{i=0}^n \subset \mathcal{H}$ *is an orthogonal set, then*

$$||\sum_{i=1}^n x_i||^2 = \sum_{i=1}^n ||x_i||^2.$$

Proof.

$$||\sum x_i||^2 = \left\langle \sum x_i, \sum x_i \right\rangle = \sum_{i,j} \langle x_i, x_j \rangle = \sum_i \langle x_i, x_i \rangle = \sum ||x_i||^2.$$

□

Theorem A.35 (Bessel's ineqality). *Let $\mathcal{H}$ be a Hilbert space. If $\{n_\alpha\}_{\alpha \in A} \subset \mathcal{H}$ is a orthonormal set, then*

$$\sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2 \leq ||x||^2.$$

Proof. It suffices to show the inequality for any *finite* $F \subset A$. Then, by the Pythagorean Theorem,

$$\begin{aligned} 0 &\leq \|x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle\|^2 \\ &\leq \|x\|^2 - 2\Re e \left\langle x, \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha \right\rangle + \| \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha \|^2 \\ &= \|x\|^2 - 2 \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 + \sum_{a \in F} |\langle x, u_\alpha \rangle|^2 \\ &= \|x\|^2 - \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 \end{aligned}$$

□

Definition A.36 (Orthonormal Basis). *An orthonormal set $\{u_\alpha\}_{\alpha \in A}$ in a Hilbert space $\mathcal{H}$, is called an* orthonormal basis *for $\mathcal{H}$ if the following equivalent properties are satisfied;*

- *(Completeness) If $\langle x, u_\alpha \rangle = 0, \forall \alpha$, then $x = 0$,*
- *(Parseval's equation) $||x||^2 = \sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2, \forall x \in V$,*
- *$\forall x \in \mathcal{H}, x = \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha$, where the sum has only countably many nonzero terms, and converges in norm to x, no matter how these terms are ordered.*

Definition A.37 (Dual basis, Biorthogonal bases). *Given a basis $\{u_\alpha\}_{\alpha \in A}$ in a Hilbert space $\mathcal{H}$. A set $\{\tilde{u}_\alpha\}_{\alpha \in A} \subset \mathcal{H}$ is a* dual basis *of $\{u_\alpha\}_{\alpha \in A}$, if*

$$\langle u_\alpha, \tilde{u}_\beta \rangle = \delta(\alpha - \beta) = \begin{cases} 1 & \text{for } \alpha = \beta, \\ 0 & \text{for } \alpha \neq \beta \end{cases} .$$

$\{u_\alpha\}_{\alpha \in A}$ and $\{\tilde{u}_\alpha\}_{\alpha \in A}$ together is called a biorthogonal basis *for $\mathcal{H}$.*

Theorem A.38. *If $\{u_\alpha\}_{\alpha \in A}$ and $\{\tilde{u}_\alpha\}_{\alpha \in A}$ are dual bases in a Hilbert space $\mathcal{H}$, then, $\forall a \in \mathcal{H}$,*

$$a = \sum_{\alpha \in A} \langle a, u_\alpha \rangle \tilde{u}_\alpha = \sum_{\alpha \in A} \langle a, \tilde{u}_\alpha \rangle u_\alpha.$$

Proof. □

Definition A.39 (Frames, Tight frames). *A set* $\{u_\alpha\}_{\alpha\in A}$ *in a Hilbert space* $\mathcal{H}$ *is called a* frame *if, for given* $0 < A \leq B < \infty$*,*

$$A\|g\|^2 \leq \sum_{\alpha\in A} |\langle u_\alpha, g\rangle|^2 \leq B\|g\|^2,$$

for all $g \in \mathcal{H}$*. If* $A = B$*,* $\{u_\alpha\}_{\alpha\in A}$ *is called a* tight frame.

Lemma A.40 (Dominated Convergence Theorem). *Let* $\{f_n\} \subset L^1(\mathbb{R})$ *such that*

- $f_n \to f$ *almost everywhere.*
- $\exists g \in L^1(\mathbb{R})$ *such that* $|f_n| \leq g,\ \forall n \in \mathbb{N}$.

Then $f \in L^1(\mathbb{R})$ *and* $\int f = \lim \int f_n$*.*

Proof. By taking real and imaginary parts it suffices to assume that f_n and f are real valued. We have that $g + f_n \geq 0$ and $g - f_n \geq 0$ almost everywhere. By Fatou's lemma

$$\int f + \int g \leq \liminf \int (g + f_n) = \int g + \liminf \int f_n,$$

and

$$\int g - \int f \leq \liminf \int (g - f_n) = \int g - \limsup \int f_n.$$

Hence

$$\limsup \int f_n \leq \int f \leq \int f_n.$$

□

Lemma A.41 (The Fubini-Tonelli theorem). $f : \mathbb{R}^2 \to \mathbb{R}$. *Define* $g(y) = \int f(x,y)\,dx$*, and* $h(x) = \int f(x,y)\,dy$

- *Let* $0 \leq f(x,y)$*. Suppose* $g(y) = \int f(x,y)\,dx$ *is measurable* $\forall y \in \mathbb{R}$ *and* $h(x) = \int f(x,y)\,dy$ *is measurable* $\forall x \in \mathbb{R}$*. Then*

  $$\int\int f(x,y)\,dx\,dy = \int\int f(x,y)\,dy\,dx.$$

- *If* $f \in L^1(\mathbb{R}^2),\ g, h \in L^1(\mathbb{R})$ *then*

  $$\int\int f(x,y)\,dx\,dy = \int\int f(x,y)\,dy\,dx.$$

Proof. The proof is on page 65 in Folland (11). □

Lemma A.42. *Let* $a, b \geq 0$ *and* $0 < \lambda < 1$. *Then*

$$a^\lambda b^{(1-\lambda)} \leq \lambda a + (1-\lambda)b.$$

Proof. If $b = 0$ the result is obvious; otherwise, setting $t = a/b$, we need to show that $t^\lambda \leq \lambda t + (1-\lambda)$ with equality if and only if $t = 1$. But by elementary calculus, $t^\lambda - \lambda t$ is strictly increasing for $t < 1$ and strictly decreasing for $t > 1$, so its maximum value, namely $1 - \lambda$, occur at $t = 1$. □

Theorem A.43 (Hölder's ineqality). *Let* $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. *Then*

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q.$$

Hence, if $f \in L^p(X) g \in L^q(X)$, *then* $fg \in L^1(X)$.

Proof. Letting

$$a = \left|\frac{f(x)}{\|f\|_p}\right|^p, b = \left|\frac{g(x)}{\|g\|_q}\right|^q, \text{ and } \lambda = \frac{1}{p},$$

in the lemma above, we get

$$\frac{|f(x)g(x)|}{\|f\|_p \|g\|_p} \leq \frac{|f(x)|^p}{p \int |f|^p d\mu} + \frac{|g(x)|^q}{q \int |g|^q d\mu}.$$

Integrating both sides yields

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

□

Theorem A.44 (Minkowski's ineqality). *Let* $1 \leq p < \infty$ *and* $f, g \in L^p(X)$. *Then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. The proof is on page 175 in Folland (11). □

A.5 Distribution theory

Definition A.45 ($C^\infty(\Omega)$ and $C^\infty_\downarrow(\Omega)$). .

$$C^\infty(\Omega) = \{f | \frac{d^n}{dx^n} f(x) \textit{ exists forall } n \in \mathbb{N}\}.$$

$$C^\infty_\downarrow(\Omega) = \{f \in C^\infty(\Omega) | \text{ ...}\}.$$

Definition A.46 (Test Functions). *Let* $\Omega \subset \mathbb{R}^n$ *be non-empty.*

$$\mathcal{D}(\Omega) = C_0^\infty(\Omega) = \{f \in C^\infty(\Omega) : \operatorname{supp}(f) \textit{ compact}\}.$$

Definition A.47 (The Schwarz-class).

$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) : |\mathbf{x}|^k |D^\alpha \phi(\mathbf{x})| < \infty \forall k \in \mathbb{N},\ \alpha = (\alpha_1, \dots, \alpha_m)\}.$

Definition A.48 (Distribution). *A* distribution *or* generalized function *is a linear mapping* $T_f =< f, \cdot >: \mathcal{D}(\Omega) \to \mathbb{R}$, $\phi \mapsto (f, \phi)$, *which is continuous in the following sense: If* $\phi_n \to \phi$ *in* $\mathcal{D}(\Omega)$, *then* $< f, \phi_n > \to < f, \phi >$. *The set of all distributions is called* $\mathcal{D}'(\Omega)$.

Definition A.49 (Convergence in the Sense of Distributions). *Let* $\{T_n\}_{n\in\mathbb{N}}$ *be distributions. We say that* $T_n \to T$ *if* $< T_n, \phi > \to < T, \phi >$ *for all test functions* ϕ.

Definition A.50. *Let* S *and* T *be distributions,* $g \in C^\infty(\mathbb{C})$ *and* $a \in \mathbb{C}$. *Then we define the following new distributions:*

- $S + T$: $< S + T, \phi >=< S, \phi > + < T, \phi >$.
- aT: $< aT, \phi >= a < T, \phi >$.
- $\frac{\partial}{\partial x} T$: $< T', \phi >= - < T, \phi' >$.
- $D^\alpha f$: $< D^\alpha, \phi >= (-1)^{|\alpha|} < f, D^\alpha \phi >$.
- $T(ax)$: $< T(ax), \phi >= \frac{1}{|a|} < T, \phi(x/a) >$.
- $T(x - a)$: $< T(x - a), \phi >=< T, \phi(x + a) >$.
- $g(x)T(x)$: $< g(x)T(x), \phi >=< T, g(x)\phi(x) >$, *where* ϕ *is any test function.*

Lemma A.51. *Suppose* $\Theta(x) \geq 0,\ \forall x \in \mathbb{R}$, $\mathrm{supp}(\Theta) \subset (a, b)$, f *distribution and* $(f * \Theta_s)(x) \geq 0,\ \forall s$. *Then* $f(x) \geq 0$ *in the sense of distributions.*

Definition A.52 (Tempered distribution). *A* tempered distribution*is a linear mapping* $(f, \cdot) : \mathcal{S}(\mathbb{R}^m) \to \mathbb{R}$, $\phi \mapsto (f, \phi)$, *which is continuous in the following sense: If* $\phi_n \to \phi$ *in* $\mathcal{S}(\mathbb{R}^m)$, *then* $(f, \phi_n) \to (f, \phi)$. *The set of all tempered distributions is denoted* $\mathcal{S}'(\mathbb{R})$.

Definition A.53 (Approximative identity on a set $\mathbf{A}$). *An* approximative identity *on a set* $\mathbf{A}$ *is a family of functions* $\{\phi_n\}_{n\in\mathbb{Z}} \subset C(\mathbf{A})$ *such that:*

- $\int_{\mathbf{A}} \phi_n(x)\, dx = 1, \forall n \in \mathbf{N}$,
- $\int_{\mathbf{A}} |\phi_n(x)|\, dx = 1$ ($\Rightarrow \phi_n(x) \geq 0$ *almost everywhere*),
- $\int_{|x|>\delta \mod (\mathbf{A})} |\phi_n(x)|\, dx \to 0$ *for all* $\delta > 0$ *when* $n \to \infty$.

Example (Approximative identity of class $C^\infty(\mathbb{R})$).
Let $h(x) = \begin{cases} e^{-1/x} & , x > 0 \\ 0 & , x \leq 0. \end{cases}$ *Then* $\frac{d^n}{dx^n} h(x) = h^{(n)}(x) = \frac{P_n(x)}{t^{2n}} e^{-1/x}$, *where* $P_n(x)$ *is a polynomial of degree* $(n - 1)$, *proven easily by induction on* n. *Now, define* $\phi(x) = h(1 - x^2)$. *Let* $\phi_1(x) = \frac{\phi(x)}{\int_{\mathbb{R}} |\phi(x)|\, dx}$, *and* $\phi_n(x) = n\phi_1(nx)$. *Then* $\{\phi_n\}_{n\in\mathbb{N}}$ *is a set of approximative identities of class* $C^\infty(\mathbb{R})$, *and* $\mathrm{supp}(\phi_n) \subset [-1/n, 1/n]$.

Example (The delta function). *The* Delta Function *is a distribution (not a function), and is defined as* $\delta(x) = \lim_{a\to\infty} af(ax)$*, where* $f(x)$ *is a approximative identity.*
We have $< \delta(x-a), \phi >=< \delta, \phi(x+a) >= \phi(a)$.

Definition A.54 (Convolution of Distribution and Test Function). *Let* T *be a distribution and* ϕ *a test function. We define* $(T * \phi)(x) =< T(y), \phi(x-y) >=< T(x-y), \phi(y) >$.

There probably is no sensible definition of convolution of two arbitrary distributions.

Lemma A.55. *Let* T *be a distribution and* ϕ *a test function. We then have:*

- $T * \phi$ *is a test function.*
- *If* $\operatorname{supp}(T) \subset [-a, a]$ *and* $\operatorname{supp}(\phi) \subset [-b, b]$*, then*

  $\operatorname{supp}(T * \phi) \subset [-(a+b), a+b]$.
- *If* T *have compact support, then*

  $T'(x) = \delta'(x) * T(x)$*, and*

  $T(x-a) = \delta(x-a) * T(x)$.

Theorem A.56. *Every distribution* T *is the limit in the distribution sense of a sequence* $(\phi_n) \subset C^\infty$*. If* T *has compact support, then* (ϕ_n) *will be test functions.*

Lemma A.57 (The $\phi(x)/x$ lemma). *Let* $\phi(x) \in C^\infty(\mathbb{R})$ *such that* $\phi(0) = 0$*. Then* $\phi(x)/x \in C^\infty(\mathbb{R})$.

A.6 Fourier Transforms

The Fourier transform is the classical ’frequency transform’ which gives us the ’contribution’ of each frequency to the total signal. It is ’totally un-localized’ in that a small change to a small part of the signal gives contribution to the whole transform.

Definition A.58 (Fourier Transform). *We define the* Fourier-transform $\hat{f}$ *of* f *as* $\hat{f}(\xi) = \int_{\mathbb{R}} f(t)e^{-it\xi}\, dt$.

Definition A.59 (Weak derivative). $f \in L^1(\mathbb{R})$ *has a* weak derivative $g \in L^1(\mathbb{R}) \Leftrightarrow \int_{\mathbb{R}} g(y)\phi(y)\, dy = -\int_{\mathbb{R}} f(y)\phi'(y)\, dy,\ \forall \phi \in C_0^1(\mathbb{R})$.

The motivation for this definition comes from the theory of distributions, and by *Integration by Parts*.

Lemma A.60. *Let* $f \in L^1(\mathbb{R})$.

1. *Inversion formula:* $f(x) = (\hat{f})^\wedge(x) = \int \hat{f}(\gamma)e^{2\pi i\gamma x}\, d\gamma$.
2. $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi)e^{i\xi x}\, d\xi$ *almost everywhere.*

3. *Parseval's identity:* $\langle f, g\rangle = \left\langle \hat{f}, \hat{g} \right\rangle$.

4. *The Convolution Theorem:* $\widehat{f * g} = \hat{f}\hat{g}$.

5. $\wedge : C^{\infty}_{\downarrow}(\mathbb{R}) \to C^{\infty}_{\downarrow}(\mathbb{R})$ *is surjective.*

6. $|\hat{f}(\xi)| \leq \int_{\mathbb{R}} |f(x)|\, dx = ||f||_{L^1(\mathbb{R})}$.

7. $\hat{f}(\xi)$ *is continuous on* $\mathbb{R}$.

8. f *has a weak derivative* $g \in L^1(\mathbb{R}) \Rightarrow \hat{g}(\xi) = i\xi \hat{f}(\xi)$.

9.
$$yf(y) \in L^1(\mathbb{R}) \Rightarrow \hat{f} \text{ differentiable, and } \hat{f}'(\xi) = -i(yf(y))\hat{} \tag{A.5}$$

10. $\psi \in C([a,b]),\ f \in L^1([a,b]) \Rightarrow g(x) = (f * \psi_s(x)) \in L^1([a,b])$.

Lemma A.61. $f \in L^2(\mathbb{R}) \Rightarrow \hat{f} \in L^2(\mathbb{R})$.

Lemma A.62 (Heisenberg's unequality). *A function cannot be both band- and time-limited.*

Proof. If f is band limited, then f is the restriction to $\mathbb{R}$ of an entire analytic function. If f is time limited as well, then $f \equiv 0$. □

Lemma A.63 (Shannon's theorem). *Let* $f \in L^2(\mathbb{R})$ *be band limited,* $\operatorname{supp}(\hat{f}) \subset [-\pi, \pi]$. *Then* $f(x) = \sum_{n\in\mathbb{Z}} f(n) \frac{\sin((x-n)\pi)}{\pi(x-n)}$.

Proof. $\hat{f}(\xi) = \sum_{n\in\mathbb{N}} < f, \frac{e^{-inx}}{\sqrt{2\pi}} > e^{-in\xi}$, where $c_n =< f, \frac{e^{-inx}}{\sqrt{2\pi}} >= \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{f}(\xi)e^{in\xi}\, d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi)e^{in\xi}\, d\xi = \frac{1}{\sqrt{2\pi}} f(n)$.
We then have:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\xi) e^{ix\xi}\, d\xi \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \left(\sum_{n\in\mathbb{N}} c_n e^{-in\xi} \right) e^{ix\xi}\, d\xi \\ &= \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{N}} c_n \int_{-\pi}^{\pi} e^{i(x-n)\xi}\, d\xi \\ &= \sum_{n\in\mathbb{N}} f(n) \frac{\sin(\pi(x-n))}{\pi(x-n)}. \end{aligned}$$

□

Definition A.64 (Nyquist sampling density). *When* $\operatorname{supp}(\hat{f}) \subset [-\Omega, \Omega] \subset \mathbb{R}$, *the sampling partition of* $\mathbb{R}$ *determined by the sample points* $\{n\frac{\pi}{\Omega}\}$ *is called the* Nyquist sampling density.

A.7 Function Spaces

Definition A.65 ($\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{Z}$, $\mathbb{C}$, $\mathbb{C}^n$, $\mathbb{N}$, $\mathbb{H}$).

- *The real numbers :* $\mathbb{R} = \{x : x$ *real*$\}$.
- *The integers :* $\mathbb{Z} = \{n : n$ *integer*$\}$.
- *The Complex numbers :* $\mathbb{C} = \{z : z$ *Complex*$\}$.
- *The natural numbers :* $\mathbb{N} = \{n \in \mathbb{Z} : n > 0\}$.
- *The Euclidean spaces :* $n \in \mathbb{N},\ \mathbb{C}^n,\ \mathbb{R}^n, n \in \mathbb{N}$.
- *The Half Plane :* $\mathbb{H} = \{(b, a) \in \mathbb{R}^2 : a > 0\}$.

Definition A.66 (Differentiable Functions, $C^n(\Omega)$).
Let $n \in \mathbb{N},\ 0 \leq n \leq \infty$ *and* $\Omega \in \mathbb{C}^n$

$$C^n(\Omega) = \{f : \Omega \to \mathbb{C} : \frac{d^k}{dx^k} f(x) = f^{(k)}(x)) \text{ exists for } k \in \mathbb{N},\ 0 \leq k \leq n\}.$$

Definition A.67 (The L^p-spaces). *Let X be a measure space with σ-algebra $\mathcal{M}$ and measure μ and let $0 < p < \infty$. Define the equivalence relation $\sim$ by*

$$f \sim g \Leftrightarrow \int_X |f - g|^p \, d\mu = 0.$$

Let $||f||_p = \left(\int_X |f|^p \, d\mu\right)^{1/p}$. *Then*

$$L^p(X, \mathcal{M}, \mu) = L^p(X) = \{f : X \to \mathbb{C} |\ f \text{ is measurable and } ||f||_p < \infty\}/\sim .$$

Definition A.68 (The Schwarz-class or 'Tempered Functions').

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) : |\mathbf{x}|^k |D^\alpha \phi(\mathbf{x})| < \infty,\ \forall k \in \mathbb{N},\ \alpha = (\alpha_1, \ldots, \alpha_m)\}.$$

$$\mathcal{S}_0(\mathbb{R}) = \{f \in \mathcal{S}(\mathbb{R}) : \mathrm{supp}(f) \text{ compact}\}.$$

Definition A.69 (Test function). *Let Ω be a non-empty set in $\mathbb{R}^n$. A function f defined on Ω is called a* test function *if $f \in C^\infty(\Omega)$, and f is compactly supported. The set of test functions is denoted $\mathcal{D}(\Omega) = C_0^\infty(\Omega)$.*

Definition A.70 (Distributions). *A* distribution *or* generalized function *is a linear mapping $T_f =< f, \cdot >: \mathcal{D}(\Omega) \to \mathbb{R}$, $\phi \mapsto (f, \phi)$, which is continuous in the following sense: If $\phi_n \to \phi$ in $\mathcal{D}(\Omega)$, then $< f, \phi_n > \to < f, \phi >$. The set of all distributions is called $\mathcal{D}'(\Omega)$.*

Definition A.71 (Local W^r-regularity at x_0). *Let r be a monotonic, non-negative, submultiplicative function which satisfies $r(x) = \mathbf{O}(1 + x^2)^{\gamma/2}$ for some $\gamma > 0$. Suppose $\psi \in S_0(\mathbb{R})$ is admissible ($C_\psi < \infty$).*

$$W^r(x_0) = \{f : \mathcal{W}_\psi f(a, b + x_0)| = \mathbf{O}(r(a) + r(b)),\ (a, b \to 0)\}.$$

Theorem A.72 (Holschneider (2.5.2)). *$W^r(x_0)$ in definition A.71 is independent of ψ.*

References

(1) Berman, Baras (1993): *Properties of the multiscale maxima and zero-crossings representations.* (IEEE Trans. Sign. Proc. 41)

(2) Buckheit, Chen, Donoho, Johnstone, Scargle (1995): *Wavelab reference manual.* (http://www-stat.stanford.edu/˜ wavelab/)

(3) Byrnes, Byrnes, Hargreaves, Berry (1994): *Wavelets and their applications.* (NATO ASI Series ISBN 0-7923-3078-1)

(4) Chui (1992): *An introduction to wavelets.* (Wavelets 1. Academic Press INC. LCCCN 91-58831)

(5) Chui (1992): *Wavelets: A tutorial in theory and applications.* (Wavelets 2. ISBN 0-12-174590-2)

(6) Chui, Montefusco, Puccio (1994): *Wavelets: Theory, algorithms, and applications.* (Wavelets 5. Academic Press, INC. ISBN 0-12-174575-9)

(7) Combes, Grossmann, Tchamitchian (1987): *Wavelets. Time-frequency methods and phase space.* (Springer-Verlag, ISBN 3-540-51159-8)

(8) Daubechies(1993): *Orthonormal bases of compactly supported wavelets II: Variations on a theme.* (SIAM J. math. anal. 24)

(9) Daubechies *Ten Lectures on Wavelets.* (CBMS 61. 1992. Regional conference series in applied mathematics.)

(10) Farge, Hunt, Vassilicos (1993): *Wavelets, fractals, and Fourier transforms.* (OXFORD UNIVERSITY PRESS, ISBN 019853647X)

(11) Folland, *Real Analysis, Modern Techniques and Their Applications* (Wiley-Intersciense publication. 1984 ISBN 0-471-80958-6.)

(12) Fournier (1995). *Wavelets and their applications in computer graphics.* (SIGGRAPH ’95 Course notes.)

(13) Graps (1995): *An introduction to wavelets.* (http://www.amara.com/IEEEwave/IEEEwavelet.html)

(14) Holschneider (1995): *Wavelets: An analysing tool.* (OXFORD MATHEMATICAL MONOGRAPHS, ISBN 0 19 853481 7)

(15) Holschneider, Tchamitchian (1989): *Regularite locale de la fonction "non-differentiable" de Riemann.* (Lecture notes in Mathematics, P.G.Lemarie, Ed. New York, Springer-Verlag, 1989)

(16) Holschneider, Tchamitchian (1990): *Pointwise analysis of Riemann’s "nondifferentiable" function.* (Lecture notes in Mathematics, P.G.Lemarie, Ed. New York, Springer-Verlag, 1989)

(17) Hwang, Mallat (1994): *Characterization of self-similar multifractals with wavelet maxima.* (Appl. and comp. harm. anal. 1)

(18) Jaffard (1989): *Exsposants de Hölder en des points donnès et coefficients d'ondelettes.* (C.R. Acad. sci. Paris)

(19) Jaffard (1991): *Pointwise smoothness, Two-microlocalization and wavelet coefficients.* (Publ. math. 35)

(20) Jaffard (1993): *Formalisme multifractal pour les functions.* (C.R. Acad. sci. Paris, t. 317)

(21) Jaffard (1993): *Multifractal formalism for functions. Part I: Results valid for all..* (SIAM J. of Mathematical Analysis)

(22) Jaffard (1993): *Multifractal formalism for functions. Part II: Self-similar functions.* (SIAM J. of Mathematical Analysis)

(23) Jaffard (1993): *Orthonormal and continuous wavelet transform: Algorithms and applications to the study of pointwise properties of functions.* (Wav. Frac. Fou. (10))

(24) Jaffard (1994): *Regularity analysis of functions and random processes using wavelets.* (Byrnes et. al. (3))

(25) Jaffard (1994): *Some mathematical results about the multifractal formalism for functions.* (Chui, Wavelets 5 (6))

(26) Jaffard (1995): *Functions with prescribed Hölder exponent.* (Appl. Comp. Harm. Anal. 2.4)

(27) Jaffard, Meyer (1993): *Pointwise behavior of functions.* ( )

(28) Jaffard, Meyer (1996): *Wavelet MEthods for Pointwise Regularity and Local Oscillations of Functions* (Memoirs of AMS, 587)

(29) Jaffard, Meyer, Ryan (2001):*Wavelets. Tools for Science and Technology* (SIAM Book, ISBN 0-89871-448-6)

(30) Louis, Maas, Riedler (1994): *Wavelets: Theorie und Anwendung.* ()

(31) Louis, Maas, Riedler (1997): *Wavelets: Theory and Applications* (Wiley & Sons. Pure and Apllied Mathematics. A Wiley-Interscience Series of Texts, Monographs and Tracts. ISBN 0 471 96792 0)

(32) Lu (1993): *Signal recovery and noise reduction with wavelets.* (Ph.D. thesis, Thayer scool of engineering, Dartmouth college, Hanover, New Hampshire)

(33) Mallat (1989): *A theory for multiresolution signal decomposition : The wavelet Representation.* (IEEE Trans. PAMI 11)

(34) Mallat (1989): *Multiresolution approximations and wavelet orthonormal bases of* $L^2(\mathbb{R})$. (Trans AMS 315)

(35) Mallat (1989): *Multifrequency channel decompositions of images and wavelet models.* (IEEE Trans. ASSP 37)

(36) Mallat (1991): *Zero-crossings of a wavelet transform.* (IEEE Trans. inf. th. 37)

(37) Mallat (1999) *A Wavelet Tour of Signal Processing* (ISBN 0-12-466606-X, Academic Press, 1999)

(38) Mallat, Hwang (1992): *Characterization of singularities.* (IEEE Trans. inf. th. 37)

(39) Mallat, Hwang (1992): *Singularity detection and processing with wavelets.* (IEEE Trans. inf. th. 38)

(40) Mallat, Papanicolaou, Zhang (1998): *Adaptive covariance estimation of locally stationary processes.* (Annals of Statistics 26)

(41) Mallat, Zhang (1993): *Matching pursuits with time-frequency dictionaries.* (IEEE Trans. Sign. Proc. 41)

(42) Mallat, Zhong (1992): *Wavelet transform maxima and Multiscale edges.* (Ruskai et. al (54))

(43) Mallat, Zhong (1992): *Characterization of signals from multiscale edges.* (IEEE Trans. PAMI 14)

(44) Meyer (1987): *Orthonormal Wavelets.* (Combes et. al (7))

(45) Meyer (1992): *Wavelets and Operators.* (Cambridge Univ. Press. Vol 37 Cambridge Studies in advanced Mathematics. ISBN 0521 42000 8)

(46) Meyer (1993): *Wavelets. Algorithms and Applications.* (SIAM ISBN 0-89871-309-9)

(47) Meyer (1993): *Wavelets and Operators.* (Proc. of symposia in appl. math.)

(48) Meyer, Roques (1993): *Progress in wavelet analysis and applications.* Proceedings of the international conference "Wavelets and Applications, Toulouse, France - june 1992.

(49) Meyer (1993): *Wavelets: Their past and their future* (Meyer, Roques (48))

(50) Morlet, Grossmann (1993): *Wavelets, ten years ago.* (Meyer, Roques (48))

(51) Pedersen (1980): *Analysis Now* (Graduate Texts in Mathematics 118, Springer-Verlag ISBN 0-387-96788-5)

(52) Richards, Youn (1990): *Theory of distributions: a non-technical introduction.* (Cambridge University Press. ISBN 0 521 37149 X)

(53) Rudin (1976): *Principles of Mathematical Analysis* (McGraw-Hill, ISBN 0-07-Y85613-3)

(54) Ruskai, Beylkin, Coifman, Daubechies, Mallat, Meyer, Raphael (1992): *Wavelets and their applications.* (Jones and Bartlett Publishers, Inc. ISBN 0-86720-225-4)

(55) Tchamitchian, Torresani (1992): *Ridge and skeleton extraction from the wavelet transform.* (Ruskai et. al)

(56) Chun-Liang Tu, Wen-Liang Hwang (2005): *Analysis of Singularities From Modulus Maxima of Complex Wavelets* (IEEE Trans. Inf. Theory, Vol51, NO3, MARCH 2005)

(57) Vu, Tokunaga (2001): *Wavelet and Scale-Space Theory in Segmentation of Airborne Laser Scanner Data.* (22nd Asian Conference on Remote Sensing, November 2001)

(58) Vu, Tokunaga (2004): *Filtering Airborne Laser Scanner Data: A Wavelet-Based Clustering Method.* (Photogrammetric Engineering and Remote Sensing, Vol. 70, No. 11, November 2004)

(59) Vu, Tokunaga, Yamazaki (2003): *Wavelet-based extraction of building features from airborne laser scanner data.* (Can. J. Remote Sensing, Vol. 29, No. 6, 2003)

(60) Wickerhauser (1994): *Adapted wavelet analysis from theory to software.* (A. K. Peters, Wellesley, Massachusetts. ISBN 1-56881-041-5)

(61) Wickerhauser, Farge, Goirand (1997): *Theoretical dimension and the complexity of simulated turbulence.* (Vol 6 of Wavelet Analysis and its Applications: Multiscale Wavelet Methods for PDE. ISBN 0-12-200675-5)