UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Introduction to Wavelet Theory

Autor: Olaia Bilbao Vega

Director: Dr. F. Xavier Massaneda Clares
Realitzat a: Departament de Matemàtiques
i Informàtica

Barcelona, 16 de gener de 2024

Contents

Abstract

The main goal of this work is to cover fundamental wavelet theory concepts, MRA construction, and Mallat's theorem's proof. Practical applications in signal processing and image analysis are highlighted, emphasizing wavelets' versatile role in diverse fields.

2020 Mathematics Subject Classification. 42C40, 42C99

Acknowledgements

I would firstly like to thank my advisor, Dr. F. Xavier Massaneda Clares, for his tireless help and advice during the weekly meetings we have held over the past months. Thanks for all is knowledge, corrections, patience, and all the time devoted to this work, it would not have been possible without him.

I would also like to thank my parents and my sister for their support during these years. A sincere thank you to all my friends and colleagues, past and present, old and new, for making these years more enjoyable. In particular, I would like to thank Iker and Pablo for all the patience and help they have always given me. I could not have done it without all of you.

Introduction

Within the intricate tapestry of mathematics lies a sometimes overlooked yet pervasive tool: wavelets. This work explores how wavelets work in math and how they are used in different areas like processing images, analyzing signals, and interpreting data. By exploring Wavelet Theory, this work aims to show how useful they are in solving real-world challenges.

The inception of wavelet theory traces its roots back to the latter half of the 20th century, although its antecedents can be found in diverse areas of mathematics and signal analysis that span several centuries.

The beginnings of wavelets can be associated with the work of the French mathematician Joseph Fourier in the early 19th century, who introduced the Fourier transform (Section 1.2). This transformative mathematical tool enabled the representation of signals in the frequency domain, revolutionizing the study of signals and systems. However, the Fourier transform, while powerful, faced limitations when dealing with non-stationary signals whose frequency content evolves over time.

The quest to overcome these limitations led to the development of wavelets. The term "wavelet" was first coined by the French mathematician and physicist Jean Morlet in the mid-20th century in his exploration of seismic signals. Morlet's work laid the groundwork for the understanding and application of wavelets in the analysis of signals with varying frequencies over time, addressing the shortcomings of traditional Fourier methods. To establish a relationship between time precision and frequency we will look at the Heisenberg uncertainty principle (Section 1.11), which states that the more precisely a signal's frequency is determined, the less precisely its time can be known and vice versa.

However, it was the pioneering work of the mathematicians Yves Meyer, Ingrid Daubechies, Stéphane Mallat, and others in the late 1970s and early 1980s that solidified the mathematical foundation of wavelets. One of the fundamental milestones was the introduction of the "mother wavelet" by Daubechies in the 1980s, which provided a solid basis for the construction of wavelets with specific properties, such as compactness in the time and frequency domain, essential for practical applications. They established the theoretical framework, defining the properties of wavelets, exploring their multiresolution analysis (MRA), and unveiling their applications in signal processing and data analysis.

The fundamental breakthrough came with the formulation of the discrete wavelet

transform (DWT) and the introduction of orthogonal wavelet bases. These innovations, along with the development of fast algorithms for wavelet computations by Stéphane Mallat and others, propelled wavelet theory into the spotlight of scientific and technological advancement.

The evolution of wavelet theory from its nascent stages to its present-day applications across diverse disciplines like image processing, data compression, signal analysis, and beyond, exemplifies its profound impact on modern mathematics and its practical significance in addressing real-world challenges.

As we delve deeper into this work, we will explore the mathematical underpinnings and versatile applications of wavelet theory, showcasing its journey from a theoretical concept to an indispensable tool in the realm of mathematics and signal analysis.

The initial Chapter introduces the basic ideas needed to understand Wavelet Theory thoroughly. Beginning with an examination of Hilbert Spaces, the study progresses to investigate the Fourier Transform and its key properties. Subsequently, attention is directed toward the Haar basis of $L^2[0,1]$ and an exploration of Heisenberg's Uncertainty Principle, forming the foundational pillars for the subsequent exploration of Wavelet Theory.

The core of this work revolves around a detailed analysis of Wavelet Theory. Starting with an in-depth study of the Haar wavelet and its Multiresolution Analysis (MRA), a mathematical framework that examines signals or functions at multiple scales simultaneously, capturing both coarse and fine details efficiently. The work unfolds to introduce broader concepts within Wavelet Theory. This includes presenting some examples of wavelets, such as Shannon or Daubechies Wavelets. It is also shown how to construct an MRA through Mallat's theorem, which states that any function can be decomposed into different scales and positions using wavelet transformations, allowing for an efficient representation with sparse coefficients. We will explore two different proofs of this theorem, one directly and the other through the Fourier perspective. Additionally, the study extends to analyze wavelets in $\mathbb{R}^2$ and investigates separable multiresolutions.

The final section of this work focuses on the practical applications of Wavelet Theory in various fields. From its utilization in signal and audio processing to its role in image and video compression, including specific applications such as medical image processing, the scope extends further into data and time series analysis. More-

over, the thesis explores its application in geospatial data compression and analysis, showcasing the versatility of Wavelet Theory across multiple disciplines.

In summary, this work aims not only to build a solid theoretical base in Wavelet Theory, but also to emphasize its various and significant uses in solving practical problems across many different areas.

Chapter 1

Preliminaries

Before delving into the theory of wavelets, we will provide a brief summary of concepts and results essential for their construction. We will begin by examining what a Hilbert space is and what constitutes a Hilbert basis. Additionally, we will define the Fourier transform and explore some properties that will prove useful later on. In the next section we will introduce a basis of $L^2[0,1]$, different from the Fourier basis, the Haar basis. Finally, we will discuss Heisenberg's uncertainty principle, which is crucial because it underlines the trade-off between time and frequency precision in signal analysis.

In this section, we will not delve extensively as it serves as an introduction to specific concepts that are essential and which are known for understanding the core subject of this work. We present concise notions of essential elements that are either necessary or beneficial to comprehend the principal theme of this work. Should further depth or more expansive explanations be sought, or if demonstrations are required, you can refer to the following sources: [10], [14].

1.1 Hilbert Spaces

Definition 1.1. *A Hilbert space is a vector space H on $\mathbb{C}$ with a scalar product $\langle x, y\rangle$, $x, y \in H$ such that:*

1. *$\langle x, y\rangle$ is bilinear.*
2. *$\langle x, y\rangle = \overline{\langle x, y\rangle}$, i.e. it is Hermitian.*
3. *$\langle x, x\rangle \geqslant 0$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.*

4. *H is a complete metric space with the distance included by the norm $\|x\| = \langle x, x\rangle^{1/2}$.*

Prior to establishing a Hilbert basis, we shall first define the orthogonal space.

Definition 1.2. *Given a closed subspace $M \subset H$ its orthogonal space is:*

$$M^{\perp} = \{x \in H : \langle x, y\rangle = 0 \, \forall y \in M\}.$$

In this situation any $x \in H$ can be projected onto M, as if H were a space with finite dimension.

Theorem 1.3. (Projection theorem) *Let $M \subset H$ be closed. Then $H = M \oplus M^{\perp}$, i.e., for any $x \in H$ there exist unique $y \in M$ and $z \in M^{\perp}$ such that $x = y + z$. Furthermore,*

$$d(x, M) = \|x - y\| = \|z\| = \sup_{w \in M^{\perp}, \|w\|=1} \|\langle x, y\rangle\|.$$

Definition 1.4. *A Hilbert basis consists of elements $\{e_i\}_{i \in I} \in H$, forming a complete orthonormal system. Completeness implies that the closure of the span, $\overline{\langle e_i\rangle_{i \in I}}$, covers the entire H, which is equivalent to $V^{\perp} = \{0\}$, i.e., that if $\langle x, e_i\rangle = 0 \, \forall i \in I$, then $x = 0$ necessarily.*

Projections on subspaces spanned by orthonormal systems are straightforward.

Theorem 1.5. *Let $\{e_n\}_{n=1}^{\infty}$ be a countable orthonormal system in a Hilbert space H and let $V = \langle e_n\rangle_{n=1}^{\infty}$ be its span. Given $x \in H$,*

(a) $P_V(x) = \sum_{n=1}^{\infty} \langle x, e_n\rangle \, e_n$,

(b) $\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 \leq \|x\|^2$ *(Bessel's inequality).*

In particular, given a countable orthonormal basis, the norm of any $x \in H$ can be discretized.

Corollary 1.6. *Let $\{e_n\}_{n=1}^{\infty}$ be a countable orthonormal system. The following are equivalent:*

(a) $\{e_n\}_{n=1}^{\infty}$ *is a Hilbert basis.*

(b) If $\langle x, e_n\rangle = 0$ for all $n \geq 1$, then $x = 0$.

(c) $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2$ *(Parseval's identity).*

1.2 Fourier transform and main properties

In this section we will focus on the Fourier transform, which decomposes a function, usually a signal in the time domain, into its frequency components. This transform allows the function to be expressed in terms of sinusoidal functions, which facilitates the analysis of its frequency components.

Consider the normalized Hilbert space $L^2[0, 2\pi]$, with inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}dt, \qquad f, g \in L^2[0, 2\pi].$$

Then, it is well known that the system $\{e^{int}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2[0, 2\pi]$. Now, let's define the Fourier series, which is essentially the decomposition of any $f \in L^2[0, 2\pi]$ in this basis.

Definition 1.7. *Given $n \in \mathbb{Z}$, the n-th Fourier coefficient of a function $f \in L^2[0, 2\pi]$ is*

$$\widehat{f}_n = \langle f, e_n\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{int}dt.$$

As the system $\{e^{int}\}_{n\in\mathbb{Z}}$ is a basis, we can represent any function $f \in L^2[0, 2\pi]$ as

$$f(t) = \sum_{n\in\mathbb{Z}} \widehat{f}_n e^{int}. \tag{1.1}$$

In the equation (1.1), the right hand side is called Fourier series of f. This representation allows the digitisation of the continuous signal $f(t)$, since all the information of f is encoded in the discrete sequence $\{\widehat{f}_n\}_{n\in\mathbb{Z}}$.

Now, to define the Fourier transform in L^1, we should consider the space of integrable functions in $\mathbb{R}$:

$$L^1(\mathbb{R}) = \left\{ f : \mathbb{R} \longrightarrow \mathbb{C} : \|f\|_1 = \int_{\mathbb{R}} |f(t)|dt < +\infty \right\}.$$

Definition 1.8. *Given $f \in L^1(\mathbb{R})$, its Fourier transform is the function $\widehat{f} : \mathbb{R} \longrightarrow \mathbb{C}$ defined by*

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(t)e^{-2\pi i\xi t}dt, \quad \xi \in \mathbb{R}.$$

The Fourier transform is well-defined and bounded: $|\widehat{f}(\xi)| \leq \int_{\mathbb{R}} |f(t)|dt = \|f\|_1$.

We will now list some properties of the Fourier transform. In the column "Time" we will see some operations and properties in the time domain, where the variable $t \in \mathbb{R}$ is temporal. In the "Frequency" column, where the variable $\xi \in \mathbb{R}$ refers to the frequency, we have the corresponding operations in the frequency domain. The following table shows how these operations are transformed from the time domain to the frequency domain and vice versa.

	Time	**Frequency**
(1)	linear properties $af + bg$	linear properties $a\widehat{f} + b\widehat{g}$
(2)	conjugation $\bar{f}(t) := \overline{f(t)}$	conjugate reflection $\widehat{\bar{f}}(\xi) = \overline{\widehat{f}(-\xi)} = \overline{(f^\vee)}(\xi)$
(3)	translation $\tau_h f(t) := f(t-h)$	modulation $\widehat{\tau_h f}(\xi) = M_{-h}\widehat{f}(\xi) = e^{-2\pi i h\xi}\widehat{f}(\xi)$
(4)	modulation $M_h f(t) := e^{2\pi i h t} f(t)$	translation $\widehat{M_h f}(\xi) = \tau_h \widehat{f}(\xi) = \widehat{f}(\xi - h)$
(5)	dilation $D_s f(t) := s f(st)$	inverse dilation $\widehat{D_s f}(\xi) = s D_{s^{-1}}\widehat{f}(\xi)$
(6)	derivative $f'(t)$ $f^{(k)}(t)$	multiply by polynomial $\widehat{f'}(\xi) = (2\pi i\xi)\widehat{f}(\xi)$ $\widehat{f^{(k)}}(\xi) = (2\pi i\xi)^k \widehat{f}(\xi)$
(7)	multiply by polynomial $-2\pi i t f(t)$	derivative $[-2\pi i t f(t)]\widehat{\ }(\xi) = \frac{d}{d\xi}\widehat{f}(\xi)$
(8)	convolution $(f * g)(t) := \int f(t-x)g(x)dx$	product $\widehat{(f*g)}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$
(9)	product $f(t)g(t)$	convolution $\widehat{fg}(\xi) = (\widehat{f} * \widehat{g})(\xi)$

Table 1. Fundamental properties of the Fourier transform.

We want to use the Hilbert structure of $L^2(\mathbb{R})$, but L^2 functions are not necessarily in L^1, so first we have to explain how the definition of Fourier transform is extended to L^2 functions.

Definition 1.9. *The Fourier transform of a function $f \in L^2(\mathbb{R})$ is the only signal $\widehat{f}(\xi)$ such that it is a limit, in the sense of $L^2(\mathbb{R})$, of a sequence of transforms $\{\widehat{f}_n(\xi)\}$, where $\{f_n(t)\}_n \subseteq L^2(\mathbb{R})$ satisfy $f_n(t), \widehat{f}_n(\xi) \in L^1(\mathbb{R})$ for all $n \leq 0$ and $\lim_{n\to\infty} \|f - f_n\|_2 = 0$.*

Note that it is well defined, since the L^2 limit does not depend on the representative, $\{f_n\}_n$. This can be seen in more detail in Chapter 12 from [4].

Remember that in L^2 we have the Hermitian product

$$\langle f, g\rangle = \int_{\mathbb{R}} f(t)\,\overline{g(t)}\,dt\ , \qquad f, g \in L^2(\mathbb{R}),$$

which gives the norm

$$\|f\|_2 = \left(\int_{\mathbb{R}} |f(t)|^2 dt\right)^{1/2}.$$

In L^2 the roles of f and $\hat{f}$ are equivalent, and this symmetry is often quite useful. This is clear in the following result.

Theorem 1.10. (Plancherel theorem) *Let $f \in L^2(\mathbb{R})$. Then $\hat{f} \in L^2(\mathbb{R})$ and $\|f\|_2 = \|\hat{f}\|_2$. In particular, if $f, g \in L^2(\mathbb{R})$*

$$\int_{\mathbb{R}} f(t)\,\overline{g(t)}\,dt = \int_{\mathbb{R}} \hat{f}(\xi)\,\overline{\hat{g}(\xi)}\,d\xi.$$

1.3 The Haar basis of $L^2[0,1]$

Let us define another orthogonal basis of $L^2[0,1]$ (where for ease of calculation the normalisation has been changed from $[0,2\pi]$ to $[0,1]$), the Haar basis, which will be relevant in the next chapter. This basis is of a completely different nature: its elements are compact support functions, rather than e^∞ functions, such as e^{int}.

Let us start with the constant function $\chi \equiv 1$, $\chi \in L^2[0,1]$. Let also

$$\psi_0(t) = \begin{cases} -1 & \text{if } t \in [0,1/2) \\ 1 & \text{if } t \in [1/2,1) \end{cases}$$

Note that the constant functions on each half of the interval $[0,1]$ are a linear combination of χ and ψ_0, since $\chi_{[0,1/2)} = \frac{1}{2}(\chi - \psi_0)$ and $\chi_{[1/2,1)} = \frac{1}{2}(\chi + \psi_0)$. Observe that we have divided the interval into 2 equal sized intervals.

Repeat the same procedure dyadic step by step. In each half of $[0,1]$ consider a function with the shape of ψ_0.

In the interval $[0,1/2]$ we have:

$$\psi(2t) = \begin{cases} -1 & \text{if } 2t \in [0,\frac{1}{2}) \Rightarrow t \in [0,\frac{1}{4}) \\ 1 & \text{if } 2t \in [\frac{1}{2},1) \Rightarrow t \in [\frac{1}{4},\frac{1}{2}) \end{cases}$$

Likewise for the interval $[1/2,1]$, we get:

$$\psi(2t-1) = \begin{cases} -1 & \text{if } 2t-1 \in [0,\frac{1}{2}) \Rightarrow t \in [\frac{1}{2},\frac{3}{4}) \\ 1 & \text{if } 2t-1 \in [\frac{1}{2},1) \Rightarrow t \in [\frac{3}{4},1) \end{cases}$$

For each dyadic interval of all generations we rescale the functions so that their L^2 norm is 1. So we get:

$$\psi_{1,0}(t) = \frac{1}{\sqrt{2}}\psi(2t) = \begin{cases} -\frac{1}{\sqrt{2}} & \text{if } t \in [0,\frac{1}{4}) \\ \frac{1}{\sqrt{2}} & \text{if } t \in [\frac{1}{4},\frac{1}{2}) \end{cases}$$

$$\psi_{1,1}(t) = \frac{1}{\sqrt{2}}\psi(2t-1) = \begin{cases} -\frac{1}{\sqrt{2}} & \text{if } t \in [\frac{1}{2},\frac{3}{4}) \\ \frac{1}{\sqrt{2}} & \text{if } t \in [\frac{3}{4},1) \end{cases}$$

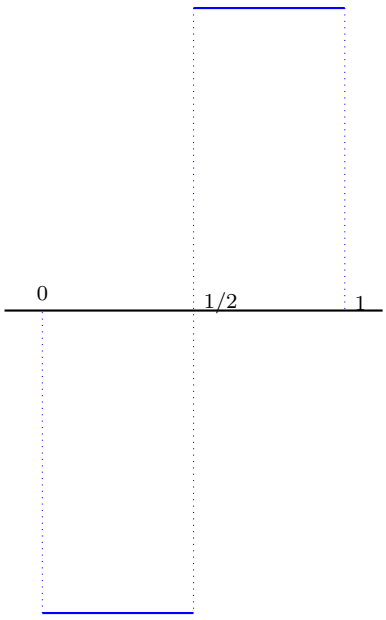

Figure 1.1: Plot of $\psi_0(t)$,

By repeating this same process, we can write, for all $n \geq 0$ and $k = 0, \ldots, 2^n - 1$, the functions

$$\psi_{n,k}(t) = 2^{n/2}\psi_0(2^n t - k) = \begin{cases} -2^{n/2} & \text{if } t \in \left[\frac{k}{2^n}, \frac{k+1/2}{2^n}\right) \\ 2^{n/2} & \text{if } t \in \left[\frac{k+1/2}{2^n}, \frac{k+1}{2^n}\right). \end{cases}$$

So, the Haar wavelet $\psi_0(t)$ on the unit interval is defined as

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

and the family $\{\psi_{j,k} := 2^{j/2}\psi(2^j t - k)\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(\mathbb{R})$. We have seen that this family generates all $L^2[0,1]$, but to see that it forms a basis of $L^2(\mathbb{R})$, it is necessary to make sure that the set is complete. This proof can be found in Chapter 9 of [14].

1.4 Heisenberg's Uncertainty Principle

Fourier emphasises frequency accuracy at the expense of time accuracy, however, wavelets offer localized information in time and frequency domains simultaneously. Now, we see Heisenberg's uncertainty principle, which states that it is impossible to find a function that it is simultaneously well-localized in time and in frequency. This general principle can be quantified in several ways, the one in equation (1.2) is probably the best known.

Theorem 1.11. (Heisenberg's Uncertainty Principle) *Let $f \in L^2(\mathbb{R})$ and let $a, b \in \mathbb{R}$. Then*

$$\left(\int_{\mathbb{R}} (t-a)^2 |f(t)|^2 dt\right)^{1/2} \left(\int_{\mathbb{R}} (\xi - b)^2 |\hat{f}(\xi)|^2 d\xi\right)^{1/2} \geq \frac{\|f\|_2^2}{4\pi}. \tag{1.2}$$

Chapter 2

Wavelet Theory

The fundamental concept behind wavelets involves representing functions or signals as combinations of small waves, and their translations and dilations. In this context, wavelets serve a similar purpose to sines and cosines in Fourier analysis but can be more effective, for example, for some natural phenomena where high frequency events happen for short duration.

What constitutes a wavelet and its multiresolution analysis? These fundamental inquiries will be addressed in the second section of this chapter. Before, it is shown the Haar wavelet as an illustrative example, which is very useful to explore the following concepts. Post our exploration of the introductory aspects of Wavelet Theory, we will examine more intricate examples showcasing diverse properties compared to the Haar wavelet.

The subsequent section will outline the process of constructing a Multiresolution Analysis (MRA), a fundamental theorem in this endeavor being Mallat's theorem. This theorem's proof will be presented via distinct methodologies, one stemming from the outlined definitions and the other from a Fourier-based perspective.

Lastly, considering the aspiration to explore applications of wavelets, and recognizing the need to address multidimensional signals beyond one-dimensional signals like sound, a brief explanation on transitioning from one-dimensional to two-dimensional wavelets in $\mathbb{R}^2$ will be presented.

2.1 The Haar wavelet and its MRA

We will now explain a simple example that contains all the essential elements that we will define in Section 2.2. In Section 1.3 the Haar basis of $L^2[0,1]$ has been defined. Consider $\varphi(t) = \chi_{[0,1]}(t)$ and its translates $\varphi_{0,k}(t) = \varphi(t-k)$, $k \in \mathbb{Z}$, which form an orthonormal basis of the closed subspace of L^2 consisting of the functions which are constant between integers:

$$V_0 = \big\{ f \in L^2(\mathbb{R}) : f_{|[k,k+1)} = c_k \text{ constant for all } k \in \mathbb{Z} \big\}.$$

Observe that a function f is in V_0 if and only if is of the form

$$f = \sum_{k\in\mathbb{Z}} c_k \varphi_{0,k} \qquad \text{with} \qquad \|f\|^2 = \sum_{k\in\mathbb{Z}} |c_k|^2 < +\infty.$$

Observe also that, given an arbitrary $f \in L^2(\mathbb{R})$, its projection on V_0 is $P_0 f = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{0,k}\rangle \, \varphi_{0,k}$, where the coefficient

$$\langle f, \varphi_{0,k}\rangle = \int_k^{k+1} f(t)\, dt = \fint_{I_{0,k}} f$$

is the average of f on the interval $[k, k+1)$. From this point of view $P_0 f$ can be seen as a low resolution approximation of f.

Consider now the space V_1 of L^2 functions which are constant on all intervals $I_{1,k} = \big[\frac{k}{2}, \frac{k+1}{2}\big)$, that is

$$V_1 = \big\{ f \in L^2(\mathbb{R}) : f_{|I_{1,k}} = c_k \text{ constant for all } k \in \mathbb{Z} \big\}.$$

We can obtain an orthonormal basis of V_1 just by rescaling the basis of V_0: for $k \in \mathbb{Z}$ let

$$\varphi_{1,k}(t) = \sqrt{2}\, \varphi(2t-k) = \sqrt{2}\, \chi_{I_{1,k}}(t).$$

It is clear that $\|\varphi_{1,k}\|_2 = 1$ and that

$$\langle \varphi_{1,k}, \varphi_{1,j}\rangle = 0 \qquad \text{for } j \neq k.$$

As before, a function of the form $f = \sum_{k\in\mathbb{Z}} c_k \varphi_{1,k}$ is in V_1 if and only if $\|f\|^2 = \sum_{k\in\mathbb{Z}} |c_k|^2 < +\infty$.

Now an arbitrary $f \in L^2(\mathbb{R})$ projected on V_1 gives

$$P_1 f = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{1,k}\rangle \, \varphi_{1,k} = \sum_{k\in\mathbb{Z}} c_k \, \chi_{I_{1,k}},$$

where

$$c_k = \sqrt{2}\langle f, \varphi_{1,k}\rangle = 2\int_{k/2}^{(k+1)/2} f(t)\,dt = \fint_{I_{1,k}} f.$$

Thus

$$P_1 f = \sum_{k\in\mathbb{Z}} (\fint_{I_{1,k}} f)\chi_{I_{1,k}}$$

is an approximation of f with "double resolution" that $P_0 f$.

Let us examine the detail we add to $P_0 f$ when increasing the resolution to produce $P_1 f$. For the sake of simplicity we just look at the interval $I_{0,0} = [0,1)$. The detail in this interval is

$$\begin{aligned}(P_1 f - P_0 f)\chi_{[0,1)} &= (\fint_{I_{1,0}} f)\chi_{I_{1,0}} + (\fint_{I_{1,1}} f)\chi_{I_{1,1}} - (\fint_{I_{0,0}} f)\chi_{I_{0,0}} \\ &= (\fint_{I_{1,0}} f - \fint_{I_{0,0}} f)\chi_{I_{1,0}} + (\fint_{I_{1,1}} f - \fint_{I_{0,0}} f)\chi_{I_{1,1}}.\end{aligned}$$

Observe that this is a multiple of the function $\psi := -\chi_{I_{1,0}} + \chi_{I_{1,1}}$, since

$$\begin{aligned}\fint_{I_{1,0}} f - \fint_{I_{0,0}} f + \fint_{I_{1,1}} f - \fint_{I_{0,0}} f &= 2\int_{I_{1,0}} f - \int_{I_{0,0}} f + 2\int_{I_{1,1}} f - \int_{I_{0,0}} f \\ &= 2\big(\int_{I_{1,0}} f + \int_{I_{1,1}} f\big) - 2\int_{I_{0,0}} = 0.\end{aligned}$$

The same arguments are valid for all intervals $[k, k+1)$, where the detail added when passing from V_0 to V_1 is a multiple of $\psi_{0,k}(t) := \psi(t-k)$.

Let W_0 denote the orthogonal complement of V_0 in V_1 (the space of details added to $P_0 f$ to get $P_1 f$, for $f \in L^2(\mathbb{R})$), so that $V_1 = V_0 \oplus W_0$ and $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_0.

This scheme can be reproduced at all scales: consider the dyadic intervals of the n^{th} generation $I_{n,k} = \big[k/2^n, (k+1)/2^n\big)$, $k \in \mathbb{Z}$, together with the closed subspace

$$V_n = \big\{f \in L^2(\mathbb{R}) : f_{|I_{n,k}} = c_k \text{ constant for all } k \in \mathbb{Z}\big\}.$$

The system $\{\varphi_{n,k}\}_{k\in\mathbb{Z}}$, with $\varphi_{n,k}(t) = 2^{n/2}\varphi(2^n t - k)$, is an orthonormal basis of V_n, and a function of the form $f = \sum_{k\in\mathbb{Z}} \alpha_k \chi_{I_{n,k}}$ is in V_n if and only if $\|f\|_2^2 = \sum_{k\in\mathbb{Z}} |\alpha_k|^2 < +\infty$.

The orthogonal projection $P_n : L^2(\mathbb{R}) \longrightarrow V_n$ produces the best approximation (in terms of the L^2-norm) of a given $f \in L^2(\mathbb{R})$ by functions which are constant on

each dyadic interval $I_{n,k}$, $k \in \mathbb{Z}$.

The detail added when passing from resolution V_n to resolution V_{n+1} forms a close space, denoted by W_n; hence $V_{n+1} = V_n \oplus W_n$. The rescaled functions $\psi_{n,k}(t) = 2^{n/2}\psi(2^n t - k)$, $k \in \mathbb{Z}$, form an orthonormal basis of W_n.

In the way they are defined it is clear that

$$\overline{\bigcup_{n\in\mathbb{Z}} V_n} = L^2(\mathbb{R}).$$

It is also clear that

$$\bigcap_{n\in\mathbb{Z}} V_n = \{0\}.$$

Then, from the iteration ($m < n$)

$$V_n = V_{n-1} \oplus W_{n-1} = V_{n-2} \oplus W_{n-2} \oplus W_{n-1} = \cdots = V_m \oplus W_m \oplus \cdots \oplus W_{n-1}$$

we deduce that

$$V_n = \bigoplus_{j=-\infty}^{n-1} W_j$$

and therefore

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j.$$

That is, any function can be viewed as the superposition of the details at all possible resolutions. Also $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{R})$.

This is an example of what is called, in general, a multiresolution analysis (MRA); this particular one is called Haar MRA. The initial function φ is called the *scaling function* and ψ is the (*mother*) *wavelet* of the MRA. The system $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$ is the wavelet basis of $L^2(\mathbb{R})$.

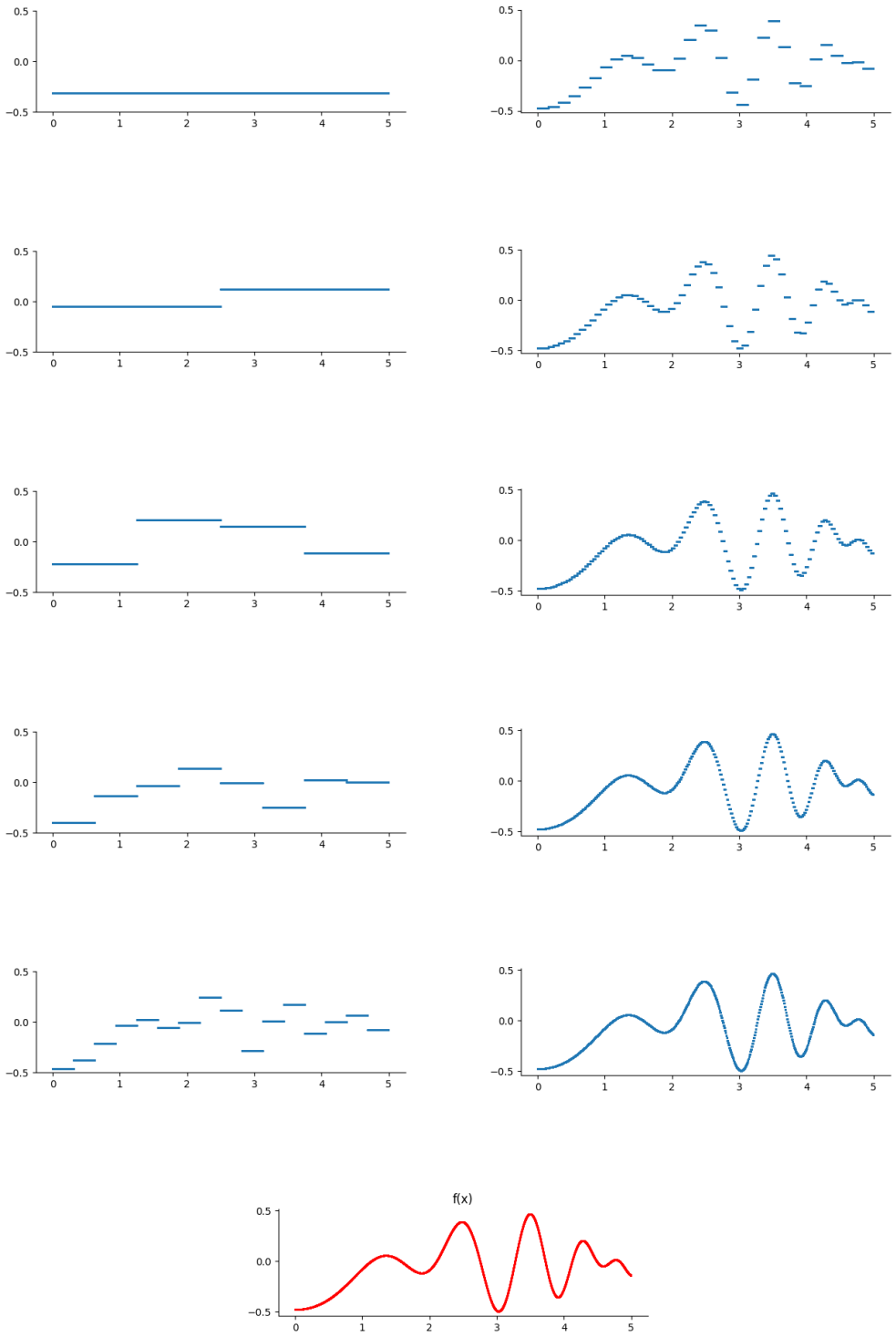

Figure 2.1: Representation of a function f with different resolutions using the Haar wavelet. The Python code to represent this f can be found in the Appendix A.

Now we will examine a specific example of how a function decomposes into its projections onto the Haar subspaces. We will do this both graphically (see Figure 2.2) and using vectors.

First of all, we will analyse this decomposition process vectorially. Consider a vector of $8 = 2^3$ samples of a function, f, which we assume to be the average value of the function on 8 intervals of length 1, so that the function is supported on the interval $[0, 8]$. In this example, we choose the vector

$$v_0 = [6, 6, 5, 3, 0, -2, 0, 6]$$

to represent f in V_0. Once we average pairs of values corresponding to taking the average value on the 4 parent intervals of length 2, we get the projection onto V_{-1}:

$$v_{-1} = [6, 6, 4, 4, -1, -1, 3, 3].$$

The difference is in W_{-1}, so we have

$$w_{-1} = [0, 0, 1, -1, 1, -1, -3, 3].$$

By repeating this process, we obtain

$$\begin{aligned} v_{-2} &= [5, 5, 5, 5, 1, 1, 1, 1], \\ w_{-2} &= [1, 1, -1, -1, -2, -2, 2, 2], \\ v_{-3} &= [3, 3, 3, 3, 3, 3, 3, 3] \text{ and} \\ w_{-3} &= [2, 2, 2, 2, -2, -2, -2, -2]. \end{aligned}$$

To compute the coefficients of the expansion, $P_j f(t) = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{j,k}\rangle \varphi_{j,k}(t)$, we must compute the inner product $\langle f, \varphi_{j,k}\rangle$ for the function $\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k)$. In terms of our vectors, we have for example

$$\langle f, \varphi_{0,3}\rangle = \langle [6, 6, 5, 3, 0, -2, 0, 6], [0, 0, 0, 1, 0, 0, 0, 0]\rangle = 3$$

and

$$\langle f, \varphi_{-1,1}\rangle = \langle [6, 6, 5, 3, 0, -2, 0, 6], [0, 0, 1/\sqrt{2}, 1/\sqrt{2}, 0, 0, 0, 0]\rangle = 8/\sqrt{2}.$$

Now, observe what happens in the example depicted in Figure 2.2. At first, there are two consecutive dyadic intervals $I_{n,k}$, $I_{n,k+1}$ of length 2^{-n}. The value of $P_{n-1}f$ on $I_{n-1,j}$ is the average of the values of $P_n f$ on $I_{n,k}$ and $I_{n,k+1}$. The difference $Q_{n-1}f = P_n f - P_{n-1}f \in W_{n-1}$ can be viewed as the detail that must be added to $P_{n-1}f$ to obtain the representation of f at resolution V_n.

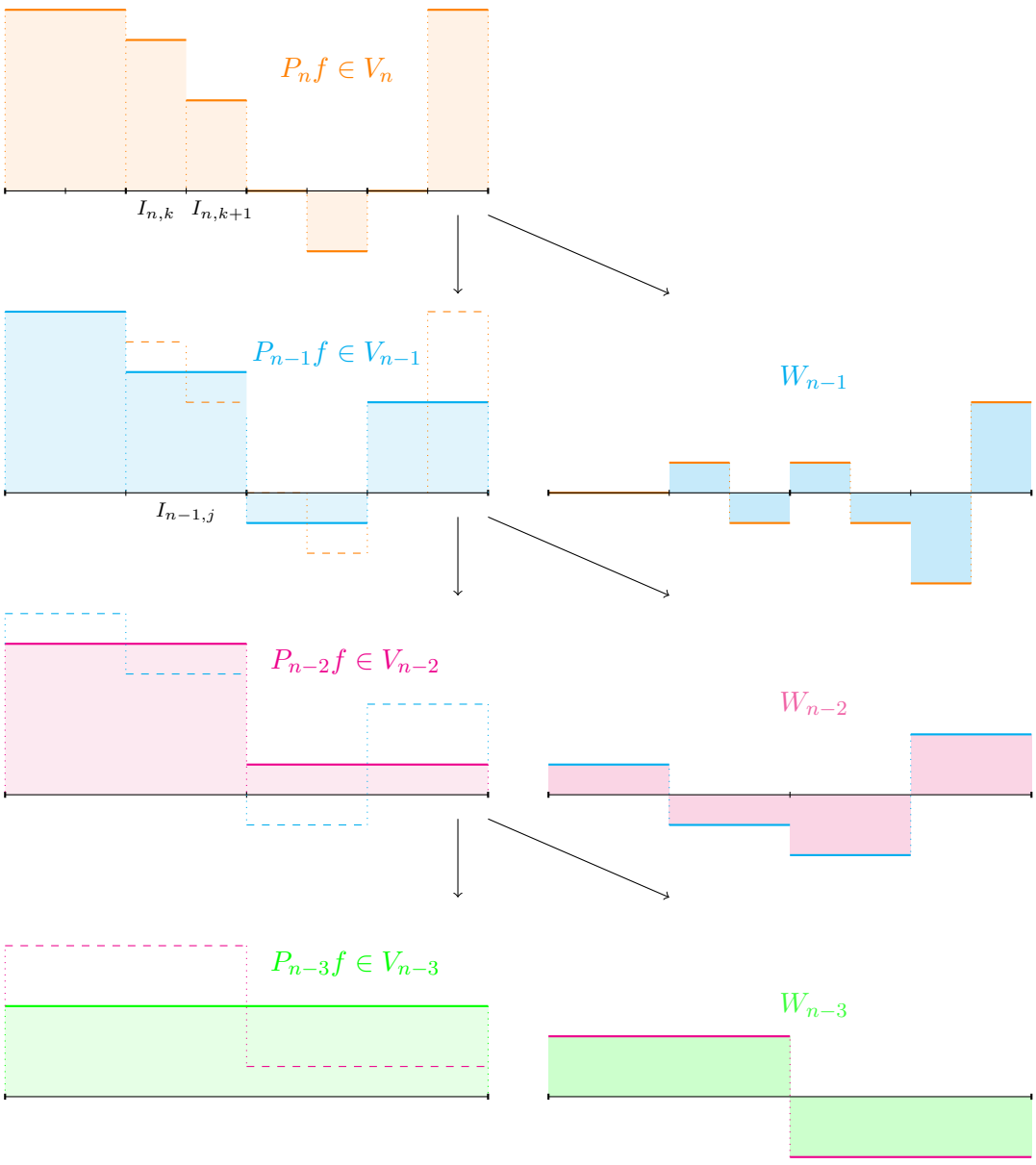

Figure 2.2: Multiresolution of a signal f.

If we look at the Figure 2.2 (to simplify the notation we will take $n = 0$) and

using what we previously we have seen that $V_{n+1} = V_n \oplus W_n$, we can see that

$$\begin{aligned} V_{-2} &= V_{-3} \oplus W_{-3} \\ V_{-1} &= V_{-2} \oplus W_{-2} = V_{-3} \oplus W_{-3} \oplus W_{-2} \\ V_0 &= V_{-1} \oplus W_{-1} = V_{-3} \oplus W_{-3} \oplus W_{-2} \oplus W_{-1} \end{aligned}$$

In summary, we have selected the coarser space V_{-n} and the finest scale V_0, truncated the chain to

$$V_{-n} \subset ... \subset V_{-2} \subset V_{-1} \subset V_0,$$

and obtained

$$V_0 = V_{-n} \oplus W_{-n} \oplus W_{-n+1} \oplus W_{-n+2} \oplus ... \oplus W_{-1}.$$

2.2 Introduction to Wavelet Theory

We will start by looking at the basic concepts of Wavelet Theory.

Definition 2.1. *An orthonormal wavelet of* $\mathbb{R}$ *is a function* $\psi \in L^2(\mathbb{R})$ *such that* $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ *is an orthonormal basis of* $L^2(\mathbb{R})$*, where*

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k). \tag{2.1}$$

Note that ψ has been translated by integers and dilated to generate an orthonormal basis. It is multiplied by $2^{j/2}$ to make each element have L^2-norm equal to one. Just by a substitution of the variable we can see that $\|\psi_{j,k}\|_2 = \|\psi\|_2 = 1$.

Moreover, if we compute the Fourier transform

$$\widehat{\psi}_{j,k}(\xi) = 2^{-j/2} e^{-2\pi i \xi 2^{-j} k} \widehat{\psi}(2^{-j}\xi)$$

we see that translations have become modulations (as we have seen in the Table 1).

Although the Fourier transforms of a wavelet basis form another orthonormal basis, it is not a wavelet basis. Since it does not have the same form as (2.1)

Definition 2.2. *A (orthogonal) multiresolution analysis is a decomposition of* $L^2(\mathbb{R})$*, into a chain of closed subspaces*

$$... \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset ... \subset L^2(\mathbb{R})$$

such that

1. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ *and* $\bigcup_{j\in\mathbb{Z}} V_j$ *is dense in* $L^2(\mathbb{R})$.

2. $f(t) \in V_j$ *if and only if* $f(2t) \in V_{j+1}$.

3. $f(t) \in V_0$ *if and only if* $f(t-k) \in V_0$ *for any* $k \in \mathbb{Z}$.

4. *There exists a function* $\varphi \in V_0$ *such that* $\{\varphi(t-k)\}_{k\in\mathbb{Z}}$, *is an orthonormal basis of* V_0. *The function* φ *is called the scaling function.*

Note that $\varphi_{j,k} \in V_j$ for all $k \in \mathbb{Z}$. Also for all $j \in \mathbb{Z}$ let P_j be the orthogonal projection into V_j, i.e.

$$P_j f(t) = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{j,k}\rangle \varphi_{j,k}.$$

The function $P_j f(t)$ can be seen as an approximation to the original function f at scale 2^{-j}. It is the best approximation of f in the subspace V_j. If we want to get a better approximation we should compute $P_{j+1}f$. In order to accomplish this, it is necessary to add to $P_j f$ the difference, $Q_j f = P_{j+1}f - P_j f$. This defines Q_j to be the orthonormal projection onto the closed subspace W_j defined as the orthogonal complement of V_j in V_{j+1}. This closed space is defined by the identity $V_{j+1} = V_j \oplus W_j$, so that, as in the Haar case,

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j. \tag{2.2}$$

It is shown [9] that the scaling function φ determines the wavelet ψ, such that $\{\psi(t-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_0. It is possible to define $\psi_{j,k}(t) = 2^{j/2}\psi\left(2^j t - k\right)$, since W_j is a dilation of W_0. So, we get $W_j = \text{span}\left(\{\psi_{j,k}(t)\}_{k\in\mathbb{Z}}\right)$.

A calculation in [9] shows that $Q_j f = \sum_{k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle \psi_{j,k}$.

The wavelet transform involves translations and dilations, and provides a zooming mechanism which is behind the multiresolution structure of these bases. The orthogonal wavelet transform is given by

$$Wf(j,k) = \langle f, \psi_{j,k}\rangle = \int_{\mathbb{R}} f(t)\overline{\psi_{j,k}(t)}dt \quad j,k \in \mathbb{Z}$$

and the reconstruction formula is,

$$f(t) = \sum_{j,k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle \psi_{j,k}(t).$$

2.3 Some examples of wavelets

The diversity of wavelets arises due to the need to address various types of signals and data characteristics. Different wavelets excel in capturing specific features such as localization in time or frequency, smoothness, or the ability to represent different types of signal behavior effectively. Moreover, applications across various fields, such as signal processing, image compression, and data analysis, often require wavelets tailored to specific needs, leading to the development of numerous types of wavelets, each optimized for particular tasks and domains.

In the previous section it is shown the Haar wavelet. It is considered basic due to its simplicity and straightforward construction, making it an elementary example of wavelet analysis. It consists of simple step functions that are easy to understand and compute, that is why we have used it for introducing the concepts of wavelet transforms. Its basic characteristics, such as discontinuity at the center and the ability to capture abrupt changes in signals, serve as a foundational basis for understanding more complex wavelet families. We will now look at some examples of other families of wavelets.

Example 2.3. (The Shannon Wavelet) This example consists of taking φ so that $\widehat{\varphi}$ is Haar's scaling function. So, let φ be given on the Fourier side by

$$\widehat{\varphi}(\xi) := e^{2\pi i\xi}\chi_{[-1,-1/2)\cup[1/2,1)}(\xi)$$

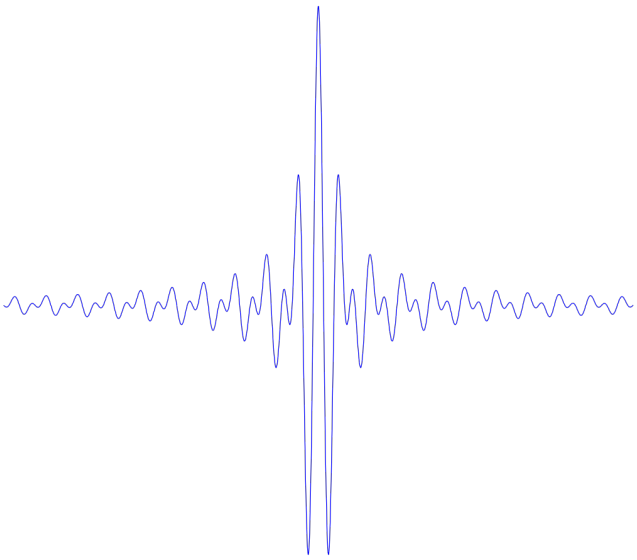

Figure 2.3: The Shannon Wavelet.

The family $\{\varphi_{j,k}\}_{j,k\in\mathbb{Z}}$ in an orthonormal basis for $L^2(\mathbb{R})$. The function φ is the Shannon wavelet, and the corresponding basis is the Shannon basis. This wavelet is C^∞ but has a slow asymptotic time decay.

Example 2.4. (The Daubechies Wavelets) In 1988, Ingrid Daubechies discovered another family of wavelets. Unlike Haar wavelets, Daubechies wavelets are continuous, so they work better with continuous signals. They also have longer supports, i.e. they use more values from the original signals to produce averages and differences. These improvements enable Daubechies wavelets to handle complicated signals more accurately.

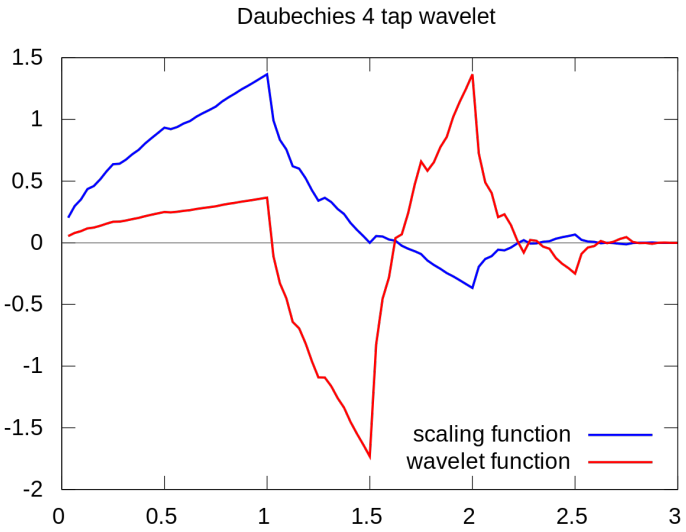

Figure 2.4: The Daubechies db4 wavelet and scaling function.

Daubechies wavelets find applications in various fields, due to their ability to compress data while retaining essential details. They have been used in medical field, for example, in the analysis and detection of community-acquired pneumonia [15], Parkinson's Disease Detection [17]; or in engineering, for example, to solve the brachistochrone problem [7] or for transient dynamic wave analysis in elastic solids [12].

Figure 2.3 has been taken from [18] and Figure 2.4 from [19]

2.4 Construction of a MRA. Mallat's theorem.

For a function $f \in L^2(\mathbb{R})$ and an integer $j \in \mathbb{Z}$ let us denote $(D_j f)(t) = 2^{j/2} f(2^j t)$, $t \in \mathbb{R}$.

In the following we propose a scheme to construct MRA.

1. Determine a $\varphi \in L^2(\mathbb{R})$ (scaling function) such that $\{\varphi_{0,k}\}_{k\in\mathbb{Z}}$ is an orthonormal system and define $V_0 = \overline{\langle \varphi_{0,k} \rangle}_{k\in\mathbb{Z}}$.
2. Check that $V_n := D_n(V_0)$ is an increasing sequence of close subspaces in $L^2(\mathbb{R})$ and that $\overline{\bigcup\limits_{n\in\mathbb{Z}} V_n} = L^2(\mathbb{R})$.
3. Find, using Mallat's theorem (see below), the associated wavelet ψ, so that $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $W_0 = V_1 \ominus V_0$.

In the third step of the procedure we use Mallat's theorem, but what does this theorem say? This theorem states that given φ satisfying 1 and 2, there exists ψ such that any function can be approximated to any desired accuracy using a linear combination of scaled and translated wavelets. We will examine the theorem and two different ways of proving it.

Firstly, we will look at some definitions, lemmas and theorems we will need. Some of their proofs can be found in Chapter 10 of [14]. We will use these lemmas mostly in the second proof, where we will prove the theorem on the Fourier side.

Definition 2.5. *A (mother) wavelet of the MRA $\{V_n\}_{n\in\mathbb{Z}}$ is a function $\psi \in W_0$ such that the translates $\psi_{0,k}(t) = \psi(t-k)$, $k \in \mathbb{Z}$, form an orthonormal basis of W_0.*

Proposition 2.6. *If ψ is a wavelet of the MRA $\{V_n\}_{n\in\mathbb{Z}}$, then the system $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$, where $\psi_{n,k} = D_n(\psi_{0,k})$, is an orthonormal basis of $L^2(\mathbb{R})$ (called the wavelet basis of the MRA).*

Proof. Since $W_n = D_n(W_0)$, $n \in \mathbb{Z}$, when ψ is a wavelet the family $\psi_{n,k} = D_n(\psi_{0,k})$, $k \in \mathbb{Z}$, is an orthonormal basis of W_n. By (2.2), the system $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$ spans then the whole $L^2(\mathbb{R})$.

It remains to see that $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$ is an orthonormal system. By definition of wavelet it is clear that $\langle \psi_{n,k}, \psi_{n,j} \rangle = \delta_{jk}$, so it will enough to prove that $W_n \perp W_m$ for $m \neq n$. But this is clear, since (assuming without loss of generality that $m < n$) one has $W_m \subseteq V_{m+1} \subseteq V_n$ and $V_n \perp W_n$. □

Observe that the following Lemma 2.7 helps us to do first step of the scheme, it tells us how the scaling function φ can be chosen.

Lemma 2.7. *Take $\varphi \in L^2(\mathbb{R})$. The family $\{\varphi_{0,k} = \tau_k\varphi\}_{k\in\mathbb{Z}}$ of integer translates of f is orthonormal if and only if*

$$\sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\xi+n)|^2 = 1 \quad \text{for a.e. } \xi \in \mathbb{R}.$$

Proof. By Plancherel, since $\widehat{\varphi_{0,k}}(\xi) = e^{2\pi ik\xi}\widehat{\varphi}(\xi)$,

$$\begin{aligned}\langle\varphi,\varphi_{0,k}\rangle = \langle\widehat{\varphi},\widehat{\varphi}_{0,k}\rangle &= \int_{\mathbb{R}} \widehat{\varphi}(\xi)\overline{\widehat{\varphi}(\xi)}\, e^{-2\pi ik\xi}d\xi = \sum_{n\in\mathbb{Z}}\int_n^{n+1} |\widehat{\varphi}(\xi)|^2 e^{-2\pi ik\xi}d\xi \\ &= \sum_{n\in\mathbb{Z}}\int_0^1 |\widehat{\varphi}(\xi+n)|^2 e^{-2\pi ik\xi}d\xi = \int_0^1 \left(\sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\xi+n)|^2\right) e^{-2\pi ik\xi}d\xi.\end{aligned}$$

Letting $F(\xi) = \sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\xi+n)|^2$, which is 1-periodic and in $L^1[0,1]$ (since $\widehat{\varphi} \in L^2(\mathbb{R})$), we have thus

$$\langle\varphi,\varphi_{0,k}\rangle = \widehat{F}(k).$$

By the uniqueness theorem for L^1 functions (see Remark **??**) the conclusion is immediate: if the system is orthonormal $\widehat{F}(k) = 0$ for all $k \neq 0$ and $\widehat{F}(0) = 1$, so $F(\xi) = 1$. On the other hand, if $F(\xi) = 1$ a.e. $\xi \in \mathbb{R}$ then $\widehat{F}(0) = 1$ and $\widehat{F}(k) = 0$ for all $k \neq 0$. □

With the above lemma we can describe determining properties of the scaling function.

Theorem 2.8. *Let φ be the scaling function of a MRA. Then:*

1. $\sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\xi+n)|^2 = 1$ *for a.e. $\xi \in \mathbb{R}$.*
2. *There exist $\{h_k\}_{k\in\mathbb{Z}} \in L^2(\mathbb{Z})$ such that*

$$\varphi(t) = \sqrt{2}\sum_{k\in\mathbb{Z}} h_k\varphi(2t-k).$$

Proof. 1. It follows from Lemma 2.7.

2. Since $\varphi \in V_0 \subseteq V_1$ and $\varphi_{1,k}(t) = \sqrt{2}\varphi(2t-k)$ are an orthonormal basis of V_1, then $\varphi(t) = \sum_{k\in\mathbb{Z}} = \langle\varphi,\varphi_{1,k}\rangle\varphi_{1,k}(t)$. Letting $h_k = \langle\varphi,\varphi_{1,k}\rangle$ we have the result, since,

$$\sum_{k\in\mathbb{Z}} |h_k|^2 = \sum_{k\in\mathbb{Z}} |\langle\varphi,\varphi_{1,k}\rangle|^2 = \|\varphi\|_2^2$$

□

Definition 2.9. *The equation*

$$\varphi(t) = \sum_{k\in\mathbb{Z}} \langle \varphi, \varphi_{1,k}\rangle \varphi_{1,k} = \sum_{k\in\mathbb{Z}} h_k \sqrt{2}\varphi(2t-k) \tag{2.3}$$

is called the dilation (or scaling) equation.

Taking Fourier transforms we get

$$\widehat{\varphi}(\xi) = \int_{\mathbb{R}} \varphi(t) e^{-2\pi i \xi t} dt = \sqrt{2} \sum_{k\in\mathbb{R}} \int_{\mathbb{R}} \varphi(2t-k) e^{-2\pi i \xi t} dt =$$
$$\overset{(s=2t-k)}{=} \frac{1}{\sqrt{2}} \sum_{k\in\mathbb{Z}} h_k e^{-\pi i \xi k} \int_{\mathbb{R}} \varphi(s) e^{2\pi i \xi (s/2)} ds.$$

Defining

$$H(\xi) = \frac{1}{\sqrt{2}} \sum_{k\in\mathbb{Z}} h_k e^{-2\pi i \xi k}$$

we get thus

$$\widehat{\varphi}(\xi) = H(\xi/2)\widehat{\varphi}(\xi/2). \tag{2.4}$$

Observe that H is 1-periodic in $L^2(\mathbb{R})$, (because $\{h_k\}_{k\in\mathbb{Z}} \in L^2(\mathbb{Z})$). It is usually called the *low pass filter*. Then h_k are called the filter coefficients.

Assume, to avoid convergence issues, that H is a finite sum, a trigonometric polynomial. These low pass filters correspond to MRA with compactly supported scaling functions, the most used in practice.

Lemma 2.10. (Quadrature mirror filter; QMF) *Given an orthogonal MRA with scaling function φ and a corresponding low-pass filter H that we assume is a trigonometric polynomial, then for every $\xi \in \mathbb{R}$,*

$$|H(\xi)|^2 + |H(\xi + 1/2)|^2 = 1$$

Proof. Insert equation (2.4) into equation 1 from Theorem 2.8, obtaining

$$1 = \sum_{n\in\mathbb{Z}} |\widehat{\varphi}(\xi+n)|^2 = \sum_{n\in\mathbb{Z}} |H((\xi+n)/2)|^2 |\widehat{\varphi}((\xi+n)/2)|^2.$$

Now separate the sum over the odd and even integers, use the fact that H has period one to factor it out from the sum, and use equation (10.5) (twice), which holds for almost every point ξ :

$$|H(\xi/2)|^2 \sum_{k\in\mathbb{Z}} |\widehat{\varphi}(\xi/2+k)|^2 + |H(\xi/2+1/2)|^2 \sum_{k\in\mathbb{Z}} |\widehat{\varphi}((\xi+1)/2+k)|^2$$
$$= |H(\xi/2)|^2 + |H(\xi/2+1/2)|^2 = 1.$$

Equality holds almost everywhere. Since H is a trigonometric polynomial, H is continuous and so equality must hold everywhere. □

Lemma 2.11. *A function $f \in W_0$ if and only if there is a function $v(\xi)$ of period one such that*

$$\widehat{f}(\xi) = e^{\pi i\xi} v(\xi) \overline{H(\xi/2+1/2)} \widehat{\varphi}(\xi/2)$$

Proof. Recall that $f \in W_0$ if and only if $f \in V_1$ and $f \perp V_0$. Since $f \in V_1$ we have $\widehat{f}(\xi) = m_f(\xi/2)\widehat{\varphi}(\xi/2)$, so we have to prove that

$$m_f(\xi) = e^{2\pi i\xi} \sigma(\xi) \overline{H(\xi+1/2)}$$

with $\sigma(\xi)$ $\frac{1}{2}$-periodic. Then $v(\xi) = \sigma(\xi/2)$ have period one.

The orthogonality $f \perp V_0$ is equivalent to $\langle f, \varphi_{0,k}\rangle = 0$ for all $k \in \mathbb{Z}$, which is by Plancharel

$$0 = \left\langle \widehat{f}, \widehat{\varphi_{0,k}} \right\rangle = \int_{\mathbb{R}} \widehat{f}(\xi) \overline{e^{-2\pi ik\xi} \widehat{\varphi}(\xi)} d\xi = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi ik\xi} \overline{\widehat{\varphi}(\xi)} d\xi =$$

$$= \int_{\mathbb{R}} e^{2\pi ik\xi} m_f(\xi/2)\widehat{\varphi}(\xi/2) \overline{H(\xi/2)\widehat{\varphi}(\xi/2)} d\xi = \int_{\mathbb{R}} e^{2\pi ik\xi} m_f(\xi/2) \overline{H(\xi/2)} |\widehat{\varphi}(\xi/2)|^2 d\xi$$

Here we break the integral over $\mathbb{R}$ into the sum of integrals over the intervals $[n, n+1)$, change variables to the unit interval, and use the periodicity of the exponential to get

$$0 = \sum_{n\in\mathbb{Z}} \int_n^{n+1} e^{2\pi ik\xi} m_f(\xi/2) \overline{H(\xi/2)} |\widehat{\varphi}(\xi/2)|^2 d\xi$$
$$= \int_0^1 e^{2\pi ik\xi} \sum_{n\in\mathbb{Z}} m_f\left(\frac{\xi+n}{2}\right) \overline{H\left(\frac{\xi+n}{2}\right)} \left|\widehat{\varphi}\left(\frac{\xi+n}{2}\right)\right|^2 d\xi$$

Now we use the 1-periodicity of m_f and H. Separating odd and even integers:

$$\begin{aligned}
0 &= \int_0^1 e^{2\pi ik\xi} \left[\sum_{k\in\mathbb{Z}} m_f\left(\frac{\xi+2k}{2}\right) \overline{H\left(\frac{\xi+2k}{2}\right)} \left|\widehat{\varphi}\left(\frac{\xi+2k}{2}\right)\right|^2 + \right. \\
&\quad \left. + \sum_{k\in\mathbb{Z}} m_f\left(\frac{\xi+2k+1}{2}\right) \overline{H\left(\frac{\xi+2k+1}{2}\right)} \left|\widehat{\varphi}\left(\frac{\xi+2k+1}{2}\right)\right|^2 \right] d\xi \\
&= \int_0^1 e^{2\pi ik\xi} \left[m_f\left(\frac{\xi}{2}\right) \overline{H\left(\frac{\xi}{2}\right)} \sum_{k\in\mathbb{Z}} \left|\widehat{\varphi}\left(\frac{\xi}{2}+k\right)\right|^2 + \right. \\
&\quad \left. + m_f\left(\frac{\xi}{2}+\frac{1}{2}\right) \overline{H\left(\frac{\xi}{2}+\frac{1}{2}\right)} \sum_{k\in\mathbb{Z}} \left|\widehat{\varphi}\left(\frac{\xi+1}{2}+k\right)\right|^2 \right] d\xi
\end{aligned}$$

By Theorem 2.8 (1), the two sums are 1 a.e. $\xi \in \mathbb{R}$; thus

$$0 = \int_0^1 e^{2\pi ik\xi} \left[m_f\left(\frac{\xi}{2}\right) \overline{H\left(\frac{\xi}{2}\right)} + m_f\left(\frac{\xi}{2}+\frac{1}{2}\right) \overline{H\left(\frac{\xi}{2}+\frac{1}{2}\right)} \right] d\xi$$

Observe that the function

$$F(\xi) = m_f(\xi/2)\overline{H(\xi/2)} + m_f(\xi/2+1/2)\overline{H(\xi/2+1/2)}$$

is 1-periodic

$$F(\xi+1) = m_f(\xi/2+1/2)\overline{H(\xi/2+1/2)} + m_f(\xi/2+1)\overline{H(\xi/2+1)} = F(\xi).$$

The above identity tells us that all Fourier coefficients $\widehat{F}(k)$, $k \in \mathbb{Z}$, are 0, so

$$F(\xi) = m_f(\xi/2)\overline{H(\xi/2)} + m_f(\xi/2+1/2)\overline{H(\xi/2+1/2)} = 0 \qquad \text{a.e. } \xi \in \mathbb{R}.$$

Equivalently, the complex vectors

$$\vec{v} = (m_f(\xi/2), m_f(\xi/2+1/2)) \qquad \vec{w} = (H(\xi/2), H(\xi/2+1/2))$$

are orthogonal.

The QMF property of H (see Lemma 2.10) tells us that $\vec{w}$ is unitary, so the space $\langle\vec{w}\rangle^\perp$ is complex 1-dimensional in $\mathbb{C}^2$. So, to characterize $\langle\vec{w}\rangle^\perp$ we just need to find one non-vanishing vector. But

$$\vec{u} = (-\overline{H(\xi+1/2)}, \overline{H(\xi)}) \in \langle\vec{w}\rangle^\perp,$$

so $\vec{v} = \lambda\vec{u}$ for some $\lambda \in \mathbb{C}$, i.e.

$$\begin{cases} m_f(\xi) = -\lambda(\xi)\overline{H\left(\xi+\frac{1}{2}\right)} \\ m_f\left(\xi+\frac{1}{2}\right) = \lambda(\xi)\overline{H(\xi)} \end{cases}$$

Here necessarily $\lambda(\xi)$ is 1-periodic (as m_f and H). Also, by adding 1/2) to ξ:

$$m_f\left(\xi+\frac{1}{2}\right) = -\lambda\left(\xi+\frac{1}{2}\right)\overline{H(\xi)},$$

hence $\lambda(\xi) = -\lambda\left(\xi+\frac{1}{2}\right)$.

Therefore $\lambda(\xi) = e^{-2\pi i\xi}$ is 1/2-periodic:

$$\lambda\left(\xi+\frac{1}{2}\right)e^{-2\pi i\left(\xi+\frac{1}{2}\right)} = -\lambda(\xi)e^{-2\pi i\xi}e^{-\pi i} = \lambda(\xi)e^{-2\pi i\xi}.$$

Defining $\sigma(\xi) = -\lambda(\xi)e^{-2\pi i\xi}$ we have finally

$$m_f(\xi) = -\lambda(\xi)\overline{H\left(\xi+\frac{1}{2}\right)} = e^{2\pi i\xi}\sigma(\xi)\overline{H\left(\xi+\frac{1}{2}\right)}$$

as required.

□

Lemma 2.12. *Let $\{V_n\}_{n\in\mathbb{Z}}$ be a MRA with scaling function φ. Then*

$$\sum_{k\in\mathbb{Z}} |\langle\varphi,\varphi_{1,k}\rangle|^2 = 1$$

and

$$\sum_{k\in\mathbb{Z}} \langle\varphi,\varphi_{1,k}\rangle\overline{\langle\varphi,\varphi_{1,k-2l}\rangle} = 0$$

for all $l\in\mathbb{Z}\setminus\{0\}$.

Proof. The first identity follows immediately from the decomposition

$$\varphi = \sum_{k\in\mathbb{Z}} \langle\varphi,\varphi_{1,k}\rangle\,\varphi_{1,k}$$

of φ in V_1.

For the second one, let $l\neq 0$ and try to express the identity $\langle\varphi,\varphi_{0,l}\rangle = 0$ in terms of the coefficients of φ and $\varphi_{0,l}$ in the basis $\{\varphi_{1,k}\}_{k\in\mathbb{Z}}$. Since

$$\varphi_{0,l}(t) = \varphi(t-l) = \sum_{k\in\mathbb{Z}} \langle\varphi,\varphi_{1,k}\rangle\,\varphi_{1,k}(t-l)$$

and

$$\varphi_{1,k}(t-l) = \sqrt{2}\varphi\big(2t-l)-k\big) = \sqrt{2}\,\varphi\big(2t-(2l+k)\big) = \varphi_{1,k+2l}(t) \tag{2.5}$$

if follows that

$$\varphi_{0,l}(t) = \sum_{k\in\mathbb{Z}} \langle \varphi, \varphi_{1,k}\rangle\, \varphi_{1,k+2l}(t) = \sum_{k\in\mathbb{Z}} \langle \varphi, \varphi_{1,k-2l}\rangle\, \varphi_{1,k}(t).$$

This and the decomposition of φ in terms of $\{\varphi_{1,k}\}_{k\in\mathbb{Z}}$ yield,

$$0 = \langle \varphi, \varphi_{0,l}\rangle = \sum_{k\in\mathbb{Z}} \langle \varphi, \varphi_{1,k}\rangle\, \overline{\langle \varphi, \varphi_{1,k-2l}\rangle},$$

as stated. □

Now, let us enunciate Mallat's theorem.

Theorem 2.13 (**Mallat's theorem**). *Given an orthogonal MRA with scaling function φ, there is a wavelet $\psi \in L^2(\mathbb{R})$ such that for each $j \in \mathbb{Z}$, the family $\{\psi_{j,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis for W_j. Hence the family $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(\mathbb{R})$.*

2.4.1 Mallat's theorem proof on the Fourier side.

Now, let us prove Mallat's theorem on the Fourier side.

Proof. The function we are looking for is $\psi \in W_0 \subseteq V_1$, hence, by the Lemma 2.11, must satisfy

$$\widehat{\psi}(\xi) = m_\psi(\xi/2)\widehat{\varphi}(\xi/2)$$

where

$$m_\psi(\xi) = e^{2\pi i\xi}\sigma(\xi)\overline{H(\xi+1/2)} \tag{2.6}$$

and $\sigma(\xi)$ is a function with period $1/2$.

Moreover, since $\psi_{0,k}$ are an orthonormal system, by Lemma 2.7,

$$\sum_{n\in\mathbb{Z}} |\widehat{\psi}(\xi+n)|^2 = 1 \qquad \text{a.e. } \xi \in \mathbb{R}.$$

As in Lemma 2.10, from here we deduce a QMF property for m_ψ:

$$|m_\psi(\xi)|^2 + \left|m_\psi\left(\xi+\frac{1}{2}\right)\right|^2 = 1 \qquad \text{a.e. } \xi \in \mathbb{R}.$$

By the form of m_ψ, this is

$$|\sigma(\xi)|^2 \left|H\left(\xi+\frac{1}{2}\right)\right|^2 + \left|\sigma\left(\xi+\frac{1}{2}\right)\right|^2 |H(\xi)|^2 = 1$$

Since σ is 1/2-periodic and H satisfies the QMF identity, this is

$$1 = |\sigma(\xi)|^2 \left[\left| H\left(\xi + \frac{1}{2}\right) \right|^2 + |H(\xi)|^2 \right] = |\sigma(\xi)|^2$$

Thus $|\sigma(\xi)| = 1$ a.e. $\xi \in \mathbb{R}$.

Choose a $\frac{1}{2}$-periodic function with $|\sigma(\xi)| = 1$ a.e., for instance $\sigma \equiv 1$ (or $\sigma(\xi) = e^{4\pi i \xi}$). Then, define ψ by

$$\widehat{\psi}(\xi) = e^{\pi i \xi} \overline{H\left(\frac{\xi}{2} + \frac{1}{2}\right)} \widehat{\varphi}\left(\frac{\xi}{2}\right),$$

i.e., choose $m_\psi(\xi) = e^{2\pi i \xi} \overline{H\left(\xi + \frac{1}{2}\right)}$.

By Lemma 2.11, $\psi \in W_0$. Its integer translates are also in W_0:

$$\psi_{0,k}(\xi) = e^{-2\pi i k \xi} \widehat{\psi}(\xi) = e^{-2\pi i k \xi} e^{\pi i \xi} \overline{H\left(\frac{\xi}{2} + \frac{1}{2}\right)} \widehat{\varphi}\left(\frac{\xi}{2}\right)$$

and $v(\xi) = e^{-2\pi i k \xi}$ is 1-periodic.

Furthermore, $H(\xi)$ satisfies the QMF property, and so $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal family.

It remains to see that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ spans W_0. By Lemma 2.11, if $f \in W_0$, there exists $v \in L^2$ 1-periodic with $v(\xi) = \sum_{k \in \mathbb{Z}} a_k e^{-2\pi i k \xi}$, where $\sum_{k \in \mathbb{Z}} |a_k|^2 < \infty$ and $(a_k) \in l^2(\mathbb{Z})$.

Thus,

$$\widehat{f}(\xi) = \sum_{k \in \mathbb{Z}} a_k e^{-2\pi i k \xi} \widehat{\psi}(\xi) = \sum_{k \in \mathbb{Z}} a_k \widehat{\psi_{0,k}}(\xi).$$

Taking the inverse Fourier transform, we see that

$$f(t) = \sum_{k \in \mathbb{Z}} a_k \psi_{0,k}(t)$$

That is, f belongs to the span of the integer translates of ψ. The integer translates of ψ form an orthonormal basis of W_0. By scale invariance, the functions $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ form an orthonormal basis of W_j. Thus the family $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ forms an orthonormal basis of $L^2(\mathbb{R})$, as required. □

2.4.2 Mallat's theorem direct proof.

In this section it is a proof of Mallat's theorem by a direct way.

Proof. Not only does the wavelet exist, it is also of the following form:

$$\psi = \sum_{k\in\mathbb{Z}} (-1)^k \, \overline{\langle \varphi, \varphi_{1,1-k}\rangle} \, \varphi_{1,k}$$

To prove it we need to see that $\psi \in V_1 \ominus V_0$ and that the translates $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ form an orthonormal basis of W_0.

It is clear by the definition that $\psi \in V_1$. In order to see that $\psi \in V_0^\perp$ it is enough to see that for all $l \in \mathbb{Z}$

$$\langle \varphi_{0,-l}, \psi\rangle = \langle \varphi, \psi_{0,l}\rangle = 0.$$

Using again that $\varphi_{1,k}(t-l) = \varphi_{1,k+2l}(t)$ (see (2.5)), denoting $c_k = \langle \varphi, \varphi_{1,k}\rangle$ and re-indexing the sum defining ψ,

$$\begin{aligned}\psi_{0,l}(t) = \psi(t-l) &= \sum_{k\in\mathbb{Z}} (-1)^k \overline{c_{1-k}} \, \varphi_{1,k}(t-l) = \sum_{k\in\mathbb{Z}} (-1)^k \overline{c_{1-k}} \, \varphi_{1,k+2l}(t) \\ &= \sum_{m\in\mathbb{Z}} (-1)^m \overline{c_{1-m+2l}} \, \varphi_{1,m}(t).\end{aligned}$$

Since

$$\varphi = \sum_{m\in\mathbb{Z}} \langle \varphi, \varphi_{1,m}\rangle \, \varphi_{1,m} = \sum_{m\in\mathbb{Z}} c_m \, \varphi_{1,m}$$

this yields

$$\langle \varphi, \psi_{0,l}\rangle = \sum_{m\in\mathbb{Z}} (-1)^m c_m c_{1-m+2l}.$$

Since $(-1)^m$ and $(-1)^{1-m+2l}$ have opposite signs this adds up to 0 (each term appers twice, and with opposite signs).

By translation, this also shows that $\psi_{0,l} \in W_0$ for all $l \in \mathbb{Z}$.

Let us see next that $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ is an orthonormal system. Since $\langle \psi_{0,m}, \psi_{0,j}\rangle = \langle \psi, \psi_{0,j-m}\rangle$, it is enough to see that $\langle \psi, \psi_{0,l}\rangle = \delta_{0l}$, $l \in \mathbb{Z}$.

Using the definition of ψ and the expression $\psi_{0,l} = \sum_{m\in\mathbb{Z}} (-1)^m \overline{c_{1-m+2l}} \, \varphi_{1,m}$ seen above, we obtain, by the identities in (a),

$$\langle \psi, \psi_{0,l}\rangle = \sum_{k\in\mathbb{Z}} (-1)^{2k} \overline{c_{1-k}} \, c_{1-k+2l} = \sum_{m\in\mathbb{Z}} \overline{c_m} \, c_{m+2l} = \delta_{0,l}.$$

It remains to prove the hard part, that $\{\psi_{0,k}\}_{k\in\mathbb{Z}}$ generates the whole W_0. To do so we want to see that any $f \in V_1 \ominus V_0$ such that $\langle f, \psi_{0,l}\rangle = 0$ for all $l \in \mathbb{Z}$ is necessarily $f = 0$.

Since all functions appearing here are in V_1, they are completely determined by their coordinates in the basis $\varphi_{1,k}$, $k \in \mathbb{Z}$. Thus

$$f \longleftrightarrow F = (f_k)_{k\in\mathbb{Z}}\ , \quad f_k = \langle f, \varphi_{1,k}\rangle.$$

Similarly, since $\varphi_{0,l} = \sum_{k\in\mathbb{Z}} \langle \varphi, \varphi_{1,k-2l}\rangle \varphi_{1,k}$ (see the proof of the first identity in (a) above), these functions can be identified with the sequence of coefficients

$$\varphi_{0,l} \longleftrightarrow \Phi_l = (c_{k-2l})_{k\in\mathbb{Z}}.$$

In the same way, from the expression above

$$\psi_{0,l} \longleftrightarrow \Psi_l = \left((-1)^k \overline{c_{1-k+2l}}\right)_{k\in\mathbb{Z}}.$$

The orthogonality assumptions are then expressed as

$$\langle F, \Phi_l\rangle_{\ell^2(\mathbb{Z})} = \langle F, \Psi_l\rangle_{\ell^2(\mathbb{Z})} = 0 \qquad \forall l \in \mathbb{Z}. \tag{2.7}$$

Consider now the infinite matrix M whose columns are the vectors Φ_l and Ψ_l writen alternately:

$$\begin{array}{ccccccccc}
\cdots & \Phi_1 & \Psi_1 & \Phi_0 & \Psi_0 & \Phi_{-1} & \Psi_{-1} & \cdots \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\cdots & c_{-1-2} & (-1)^{-1}\bar{c}_{1-1+2} & c_{-1} & (-1)^{-1}\bar{c}_{1+1} & c_{-1+2} & (-1)^{1}\bar{c}_{1+1-2} & \cdots \\
\cdots & c_{0-2} & (-1)^{0}\bar{c}_{1-0+2} & c_{0} & (-1)^{0}\bar{c}_{1-0} & c_{0+2} & (-1)^{0}\bar{c}_{1-0-2} & \cdots \\
\cdots & c_{1-2} & (-1)^{1}\bar{c}_{1-1+2} & c_{1} & (-1)^{1}\bar{c}_{1-1} & c_{1+2} & (-1)^{0}\bar{c}_{1-1-2} & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

Observe that the columns of this matrix are orthogonal.

Claim. $MM^* = I$, where M^* indicates the conjugate transpose of M.

Accepting this the proof is finished: by assumption (see (2.7)) $M^*F = 0$, and therefore

$$F = MM^*F = 0.$$

The proof of the claim is essentialy a computation. Break M into 2×2 blocks of the form

$$M_{m,l} = \begin{pmatrix} c_{2m-2l} & \bar{c}_{1-2m+2l} \\ c_{2m-2l+1} & -\bar{c}_{-2m+2l} \end{pmatrix} \qquad m, l \in \mathbb{Z}.$$

The first column is part of the vector Φ_l and the second is part of Ψ_l. Then the matrix MM^* is made up of the blocks

$$\sum_{j\in\mathbb{Z}} M_{m,j}(M_{l,j})^* = \begin{pmatrix} c_{2m-2j} & \bar{c}_{1-2m+2j} \\ c_{2m-2j+1} & -\bar{c}_{-2m+2j} \end{pmatrix} \begin{pmatrix} \bar{c}_{2l-2j} & \bar{c}_{2l-2j+1} \\ c_{1-2l+2j} & -c_{-2l+2j} \end{pmatrix}$$

$$= \begin{pmatrix} \sum\limits_j c_{2m-2j}\bar{c}_{2l-2j} + \bar{c}_{1-2m+2j}c_{1-2l+2j} & \sum\limits_j c_{2m-2j}\bar{c}_{2l-2j+1} - \bar{c}_{1-2m+2j}c_{-2l+2j} \\ \sum\limits_j c_{2m-2j+1}\bar{c}_{2l-2j} - \bar{c}_{-2m+2j}c_{1-2l+2j} & \sum\limits_j c_{2m-2j+1}\bar{c}_{2l-2j+1} + \bar{c}_{-2m+2j}c_{-2l+2j} \end{pmatrix}$$

A computation, using the orthogonality relations seen previously shows that this is

$$\begin{pmatrix} \delta_{ml} & 0 \\ 0 & \delta_{ml} \end{pmatrix}.$$

For example, for the first entry, re-index the sum by $m - j = k$ in the first term and by $k = j - l$ in the second one:

$$\begin{aligned} \sum_j c_{2m-2j}\bar{c}_{2l-2j} + \bar{c}_{1-2m+2j}c_{1-2l+2j} &= \sum_k c_{2k}\bar{c}_{2k-2m+2l} + \sum_k \bar{c}_{2k+1-2m+2l}\ c_{2k+1} \\ &= \sum_k c_k\, \bar{c}_{k-2m+2l} = \delta_{ml}. \end{aligned}$$

The other entries are dealt with similarly.

The same procedure shows that $\{\psi_{n,k}\}_{k\in\mathbb{Z}}$ is an orthonormal system and that any $f \in V_{n+1} \ominus V_n$ which is orthogonal to that system must be 0, so $\{\psi_{n,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_n. Since $\cap_n V_n = \{0\}$ and $\overline{\cup_n V_n} = L^2(\mathbb{R})$, we deduce that $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{R})$. □

2.5 Wavelets in $\mathbb{R}^2$

As previously stated, wavelets hold significance within the realm of image processing. In consideration of these practical applications, which will be elaborated on

in the forthcoming section, we hereby introduce "separable" multiresolutions. Specifically, these multiresolutions in $\mathbb{R}^2$ are derived as the products of one-dimensional multiresolutions.

A first attempt, given $\{\psi_{n,k}\}_{n,k\in\mathbb{Z}}$ wavelet orthonormal basis of $L^2(\mathbb{R})$, would be to consider the products $L^2(\mathbb{R}^2)$:

$$\{\psi_{n_1,k_1}(t_1)\,\psi_{n_2,k_2}(t_2)\}_{\substack{n_1,n_2\in\mathbb{Z}\\ k_1,k_2\in\mathbb{Z}}}\,.$$

These functions blend information from two different scales, 2^{n_1} and 2^{n_2}, across the axes t_1 and t_2. However, this arrangement lacks convenience; it's preferable to maintain consistent scaling across all directions.

This construction can be slightly modified to provide another separable wavelet basis whose elements are products of one variable functions dilated by the same factor in all coordinate directions. These multiresolution approximations have important applications in computer vision, where they are used to process images at different levels of detail.

2.5.1 Separable multiresolutions

The formal definition of a MRA in $\mathbb{R}^2$ is as in one dimension: it is an increasing collection of closed subspaces $\{V_n^{(2)}\}_{n\in\mathbb{Z}}$ with the properties previously listed (see Definition 2.2). As in dimension one, the notion of resolution is formalised with orthogonal projections on the spaces V_n; hence the approximation of $f(t_1,t_2)$ at resolution n is the orthogonal projection of f on $V_n^{(2)}$ and the space $V_n^{(2)}$ is the set of all approximations at resolution n.

We consider only the particular case of separable multiresoltions.

Definition 2.14. *Given a multiresolution $\{V_n\}_{n\in\mathbb{Z}}$ in $L^2(\mathbb{R})$, the associated separable 2-dimensional multiresolution is $\{V_n^{(2)}\}_{n\in\mathbb{Z}}$, where $V_n^{(2)} = V_n \otimes V_n$. Thus $F \in V_n^{(2)}$ if it has the form*

$$F(t_1,t_2) = \sum_{m\in\mathbb{Z}} c_m f_m(t_1)\, g_m(t_2)\ ,$$

where $f_m, g_m \in V_n$, $\|f_m\| = \|g_m\| = 1$ and $\sum_m |c_m|^2 = \|f\|^2_{L^2(\mathbb{R}^2)} < \infty$.

It is immediate to check that $\{V_n^{(2)}\}_{n\in\mathbb{Z}}$ is a multi-resolution of $L^2(\mathbb{R}^2)$.

Let φ be the scaling function of $\{V_n\}_{n\in\mathbb{Z}}$, so that $\{\varphi_{n,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_n. Then the system

$$\varphi^{(2)}_{n,k}(t_1,t_2) = \varphi_{n,k_1}(t_1)\,\varphi_{n,k_2}(t_2) = 2^n\varphi(2^n t_1 - k_1)\,\varphi(2^n t_2 - k_2)$$
$$n \in \mathbb{Z}, \quad k = (k_1,k_2) \in \mathbb{Z}^2$$

is an orthonormal basis of $V^{(2)}_n$. Notice that the scaling function of the two-dimensional multi-resolution is just

$$\varphi(t_1,t_2) = \varphi(t_1)\,\varphi(t_2).$$

Example 2.15. (Haar) Let $\{V_n\}_{n\in\mathbb{Z}}$ be the Haar MRA given in Section 2.1. Then $V^{(2)}_n$ is the approximation space consisting on the functions in $L^2(\mathbb{R}^2)$ which are constant on dyadic squares

$$Q_{n,k} = [2^{-n}k_1, 2^{-n}(k_1+1)) \times [2^{-n}k_2, 2^{-n}(k_2+1)) \ , \quad k_1,k_2 \in \mathbb{Z}.$$

The two-dimensional scaling function is therefore

$$\varphi(t_1,t_2) = \chi_{[0,1)}(t_1)\,\chi_{[0,1)}(t_2) = \chi_{[0,1)\times[0,1)}(t_1,t_2).$$

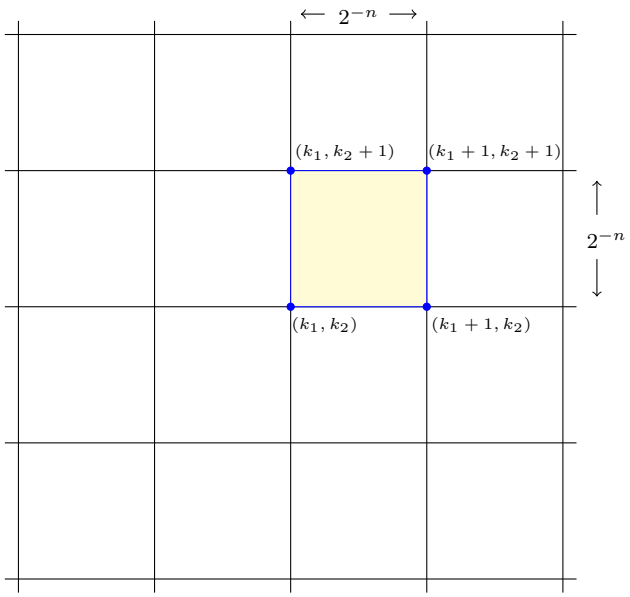

Figure 2.5: $\psi_0(t)$

Given a MRA $\{V^{(2)}_n\}_{n\in\mathbb{Z}}$ of $L^2(\mathbb{R}^2)$ let $W^{(2)}_n = V^{(2)}_{n+1} \ominus V^{(2)}_n$ denote the corresponding detail space at level n.

Theorem 2.16. *Let φ be the scaling function of a MRA $\{V_n\}_{n\in\mathbb{Z}}$ in $L^2(\mathbb{R})$ and let ψ denote the associated wavelet. Define*

$$\begin{aligned}\psi^1(t_1,t_2) &= \varphi(t_1)\,\psi(t_2),\\ \psi^2(t_1,t_2) &= \psi(t_1)\,\varphi(t_2),\\ \psi^3(t_1,t_2) &= \psi(t_1)\,\psi(t_2)\end{aligned}$$

and denote, for $n,k_1,k_2\in\mathbb{Z}$ and $j=1,2,3$,

$$\psi^j_{n,k}(t_1,t_2) = 2^n\,\psi^j\big(2^n t_1-k_1, 2^n t_2-k_2\big).$$

The collection $\big\{\psi^1_{n,k},\psi^2_{n,k},\psi^3_{n,k}\big\}_{k\in\mathbb{Z}^2}$ is an orthonormal basis of $W_n^{(2)}$, $n\in\mathbb{Z}$, and therefore the system $\big\{\psi^1_{n,k},\psi^2_{n,k},\psi^3_{n,k}\big\}_{\substack{n\in\mathbb{Z}\\ k\in\mathbb{Z}^2}}$ is an orthonormal basis of $L^2(\mathbb{R}^2)$.

Proof. By definition $V_{n+1}^{(2)} = V_n^{(2)}\oplus W_n^{(2)}$. Since by definition $V_{n+1}=V_n\oplus W_n$ we also have

$$V_{n+1}^{(2)} = \big(V_n\oplus W_n\big)\otimes\big(V_n\oplus W_n\big) = V_n^{(2)}\oplus\big(W_n\otimes V_n\big)\oplus\big(V_n\otimes W_n\big)\oplus\big(W_n\otimes W_n\big)$$

we deduce that

$$W_n^{(2)} = \big(W_n\otimes V_n\big)\oplus\big(V_n\otimes W_n\big)\oplus\big(W_n\otimes W_n\big).$$

It is clear that $\big\{\psi^1_{n,k}\big\}_{k\in\mathbb{Z}^2}$ is an orthonormal basis of $V_n\otimes W_n$, $\big\{\psi^2_{n,k}\big\}_{k\in\mathbb{Z}^2}$ is an orthonormal basis of $W_n\otimes V_n$, and $\big\{\psi^3_{n,k}\big\}_{k\in\mathbb{Z}^2}$ is an orthonormal basis of $W_n\otimes W_n$, so the statement follows. □

The three wavelets of the previous theorem extract image details at different scales and orientations. Notice also that

$$\begin{aligned}\widehat{\psi^1}(\xi_1,\xi_2) &= \widehat{\varphi}(\xi_1)\,\widehat{\psi}(\xi_2),\\ \widehat{\psi^2}(\xi_1,\xi_2) &= \widehat{\psi}(\xi_1)\,\widehat{\varphi}(\xi_2),\\ \widehat{\psi^3}(\xi_1,\xi_2) &= \widehat{\psi}(\xi_1)\,\widehat{\psi}(\xi_2).\end{aligned}$$

Sometimes ψ^1 is denoted ψ^h and called *horizontal wavelet*, because the corresponding subspaces favour details in the horizontal direction. Similarly $\psi^2=\psi^v$ is called *vertical wavelet* and $\psi^3=\psi^d$ is called *diagonal wavelet*.

Example 2.17. (Shannon wavelet) Let $\{V_n\}_{n\in\mathbb{Z}}$ be the MRA given by the scaling function $\widehat{\varphi}(\xi)=\chi_{[-1/2,1/2]}(\xi)$. Observe that, by Lemma 2.7, such a MRA exists. Actually, by Shannon's theorem, V_n coincides with the subspace of $L^2(\mathbb{R})$ consisting of the functions f such that $\operatorname{supp}(\hat{f})\subset[-2^{n-1},2^{n-1}]$. It can be proved that

the corresponding wavelet ψ is, up to a unimodular constant, given by the identity $\widehat{\psi} = \chi_{[-1,1]\setminus[-1/2,1/2]}$.

In this setting the two-dimensional basis explained above paves the Fourier plane with the dyadic dilations of the rectangles shown below. Here

$$\begin{aligned}
\widehat{\psi^1}(\xi_1,\xi_2) &= \chi_{[-1/2,1/2]}(\xi_1)\,\chi_{[-1,1]\setminus[-1/2,1/2]}(\xi_2)\\
\widehat{\psi^2}(\xi_1,\xi_2) &= \chi_{[-1,1]\setminus[-1/2,1/2]}(\xi_1)\,\chi_{[-1/2,1/2]}(\xi_2)\\
\widehat{\psi^3}(\xi_1,\xi_2) &= \chi_{[-1,1]\setminus[-1/2,1/2]}(\xi_1)\,\chi_{[-1,1]\setminus[-1/2,1/2]}(\xi_1).
\end{aligned}$$

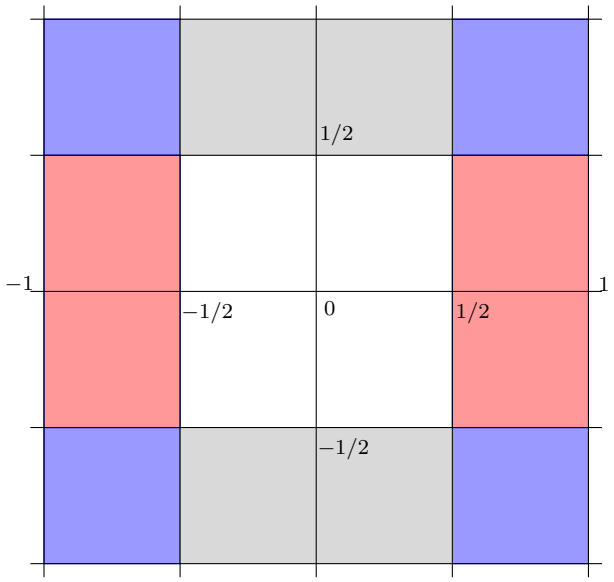

Figure 2.6: Supports of the horizontal (grey), vertical (red) and diagonal (blue) Shannon wavelets.

Chapter 3

Applications

The chapter explores the various practical applications of Wavelet Theory. From signal processing to image analysis, this chapter highlights its practical applications, the profound impact and versatility of wavelets in a variety of fields.

3.1 Signal and audio processing

Wavelets are useful for audio compression, leading to more efficient and high-quality file formats, such as the lossless audio standard FLAC. They are also used for denoising, spectrum analysis and signal filtering.

Denoising is one of the most important aspects of signal processing. It consists in the removal of noise from signals. Our discussion in this section will introduce the fundamental ideas about Haar transform, because, as it has been said in Section 2.1, it is a easy example to understand.

When a signal is received after transmission over some distance, it is frequently contaminated by noise. The term *noise* refers to any undesired change that has altered the values of the original signal. The simplest model for the acquisition of noise by a signal is *additive* noise, which has the form (contaminated signal) = (original signal)+(noise). We represent equation in a more compact way as $f = s+n$, where f is the contaminated signal, s is the original signal and n is the noise.

There are several types of noise, however, in this example we only examine the *random noise*. In this situation the Haar transform is used very effectively for removing the noise. Nevertheless, for real signals, the Haar transform usually performs poorly, and more sophisticated wavelet transforms will be needed, such that Daubechies wavelet transform. However, the essential principles underlying these

more sophisticated wavelet methods are the same principles we describe here for the Haar wavelet.

Threshold Method of Wavelet Denoising

Suppose that the contaminated signal f equals the transmitted signal s plus the noise signal n. Also suppose that these two conditions hold:

1. The energy of s is captured, to a high percentage, by transform values whose magnitudes are all greater than a threshold $T_s > 0$.

2. The noise signal's transform values all have magnitudes which lie below a noise threshold T_n satisfying $T_n < T_s$.

Then the noise in f can be removed by thresholding its transform: all values of its transform whose magnitudes lie below the noise threshold T_n are set to 0 and an inverse transform is performed, providing a good approximation of f.

Let's see how this method applies to signal A shown in Figure 3.2(a). This signal was obtained by adding random noise, whose values oscillate between ± 0.1 with mean of zero to Signal 1 shown in Figure 3.1(a). In this case, Signal 1 is the original signal and Signal A is the contaminated signal. The energy of Signal 1 is captured very effectively by the relativity few transforms values whose magnitudes lie above a threshold of 0.35. So we set T_s equal to 0.35, and condition 1 in the Denoising Method is satisfied.

Now as for condition 2, look at the 10-level Haar transform of Signal A shown in Figure 3.2(b). Comparing this Haar transform with the Haar transform of Signal 1 in Figure 3.1(b), it is clear that the added noise has contributed a large number of small magnitude values to the transform of Signal A, while the high energy transform values of Signal 1 are plainly visible (although slightly altered by the addition of noise). Therefore, we can satisfy condition 2 and eliminate the noise if we choose a noise threshold of, for example, $T_n = 2$. This is indicated by the two horizontal lines shown in Figure 3.2(b); all transform values lying between ± 0.2 are set to 0, producing the thresholded transform shown in Figure 3.2(c). Comparing Figure 3.2(c) with Figure 3.1(b) we see that the thresholded Haar transform of the contaminated signal is a close match to the Haar transform of the original signal. Consequently, after performing an inverse transform on this threshold signal, we obtain a denoised signal that is a close match to the original signal. This denoised signal is shown in Figure 3.2(d), and it is clearly a good approximation to Signal 1, specially considering how much noise was originally present in Signal A.

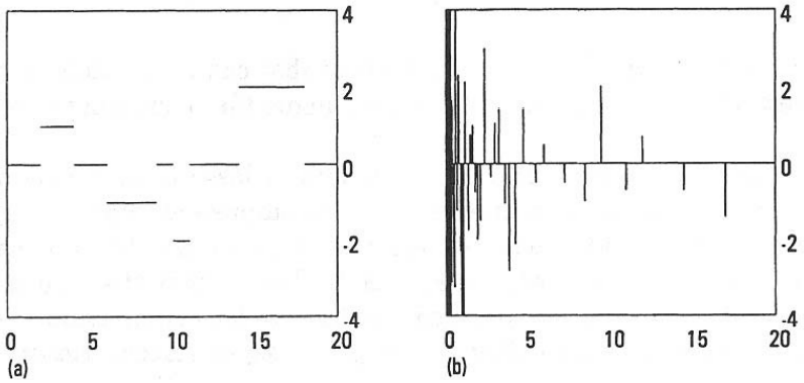

Figure 3.1: (a) Signal A. (b) Haar transform of Signal 1.

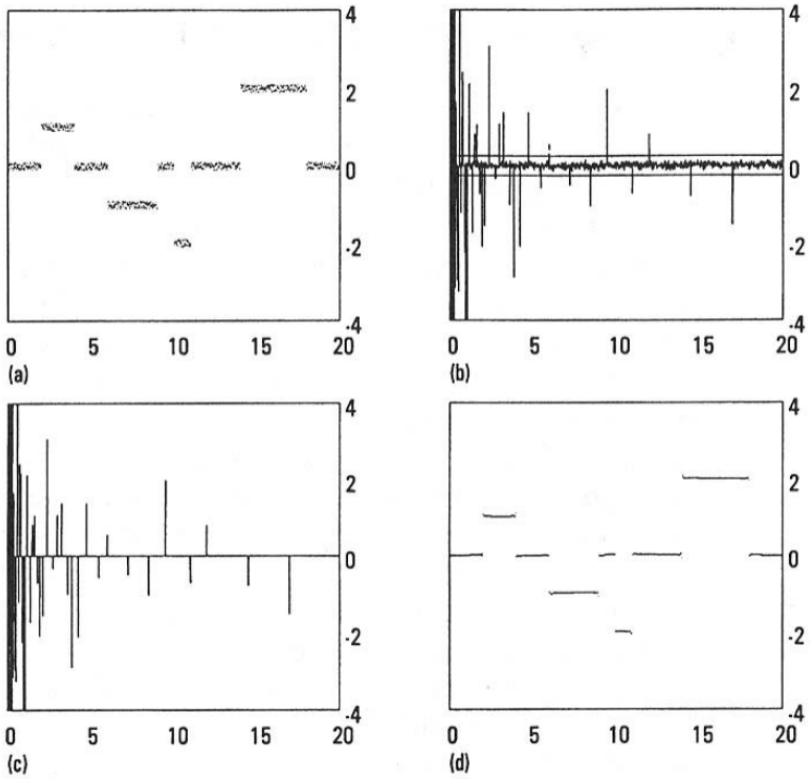

Figure 3.2: (a) Signal A, 2^{10} values. (b) Haar transform of signal A. The two horizontal lines are at ± 0.2 where 0.2 is a denoising threshold. (c) Thresholded transform. (d) Denoised signal.

This example has been borrowed from Chapter 2 of [16].

3.2 Image and video compression

Wavelets allow lossy and lossless compression of images and video, as it is done in JPEG2000, which provides both lossless and lossy compression in a single compression architecture. The ability to preserve fine details means that compressed images maintain better quality.

In addition to JPEG2000, the Wavelet Scalar Quantization algorithm (WSQ) also uses wavelet-based techniques for image compression. It is a compression algorithm used for gray-scale fingerprint images. It has become a standard for the exchange and storage of fingerprint images. WSQ was developed by the FBI, the Los Alamos National Laboratory, and the National Institute of Standards and Technology (NIST).

Although, JPEG2000 offers in general higher compression efficiency, more advanced features and a wider range of applications than WSQ, the latter has its niche in specific applications that require specialised compression, such as fingerprint images.

We will now look at some examples of images compressed with WSQ at different compression ratios. For example, in Figure 3.3 two images can be observed, the original image and its WSQ-Reconstructed image at 15:1, where 15:1 means that we have divided the bytes of the image by 15. At this resolution, it is impossible to distinguish the two images with the human eye.

In order to assess the efficacy of WSQ compared to alternative lossy compression methodologies, a comparative analysis is conducted between WSQ and JPEG, a widely employed standard in lossy image compression. JPEG employs the Discrete Cosine Transform (DCT) to divide images into blocks, allowing for efficient compression suitable for a broad range of photographic images. It offers good compression ratios but operates primarily through lossy compression, causing some loss of image quality. Through the examination of Figures 3.6, 3.7 and 3.8 the identical fingerprint image is reconstructed using both compression systems at compression ratios of 15:1, 30:1, and 45:1, correspondingly.

At the initial 15:1 compression ratio, the visual similarity between the outputs of JPEG and WSQ might be apparent to an untrained observer. However, findings from FBI latent examiners revealed that even at this relatively moderate compression ratio, JPEG-encoded images impeded fingerprint identification. As the compression ratio escalates, especially at 45:1, observable flaws in JPEG compression, notably

blocking artifacts and loss of intricate details, become increasingly conspicuous, as depicted in Figure 3.8.

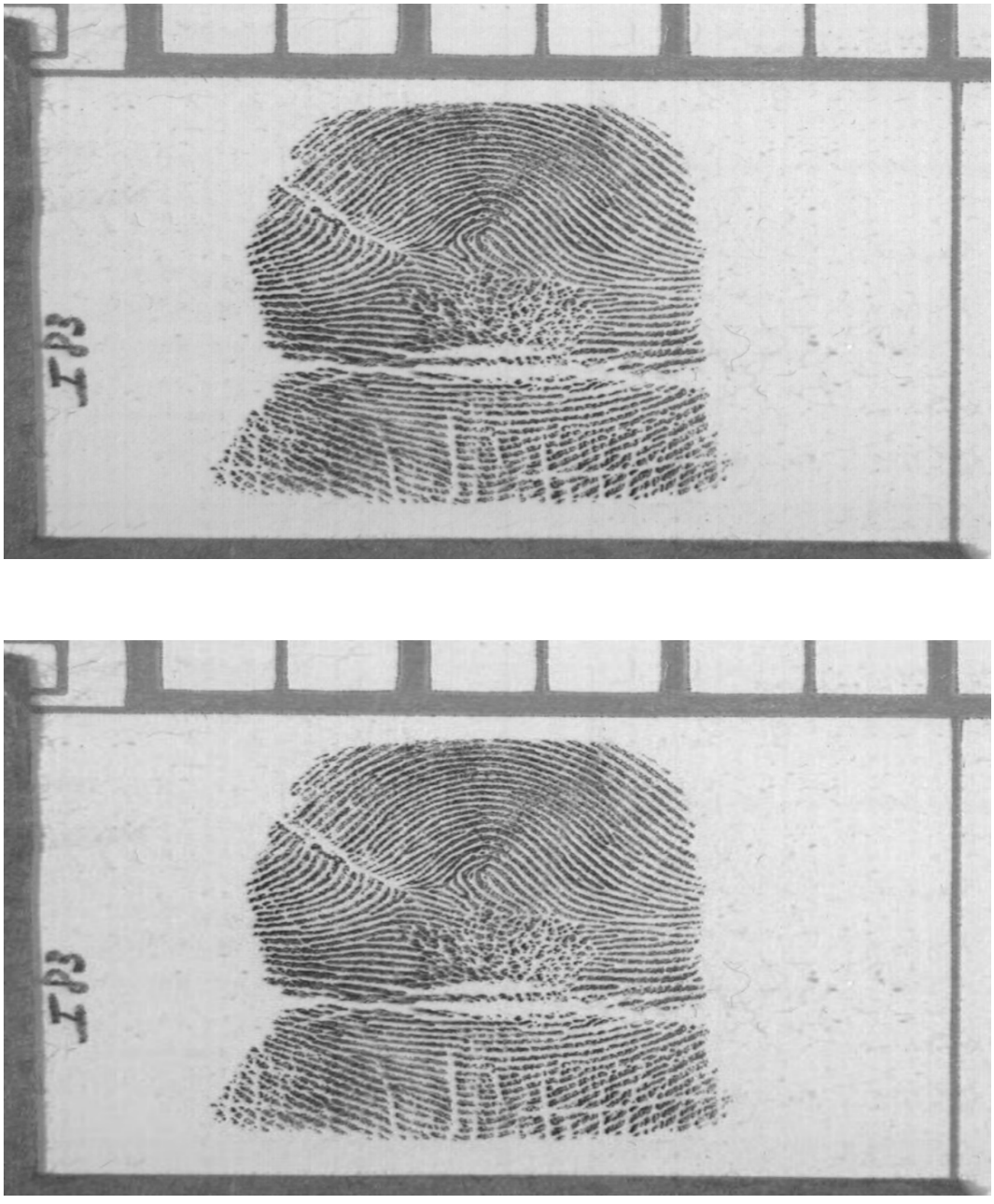

Figure 3.3: An Original Fingerprint Image (above) and its WSQ-Reconstructed at 15:1 (below).

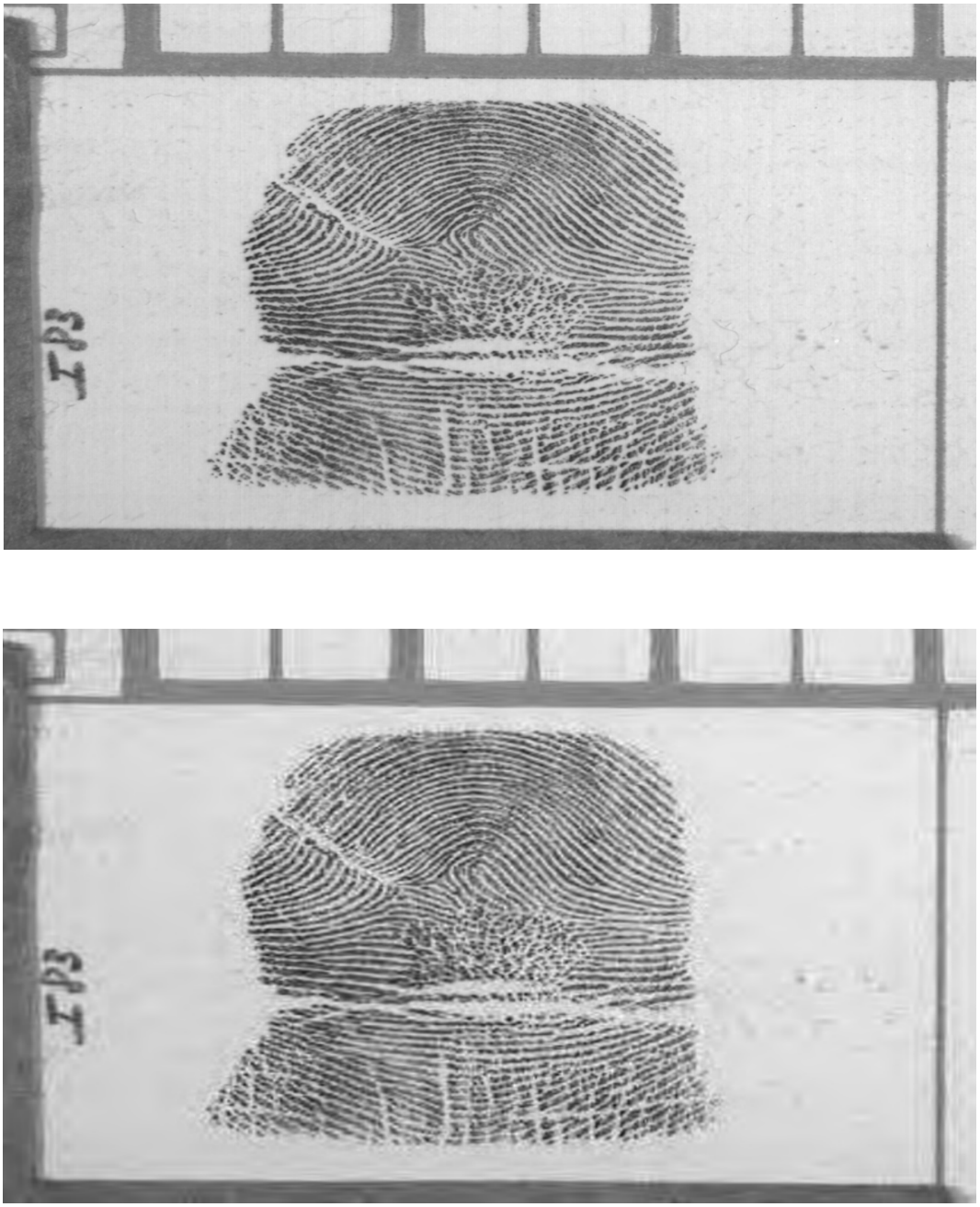

Figure 3.4: An Original Fingerprint Image (above) and its WSQ-Reconstructed at 60:1 (below).

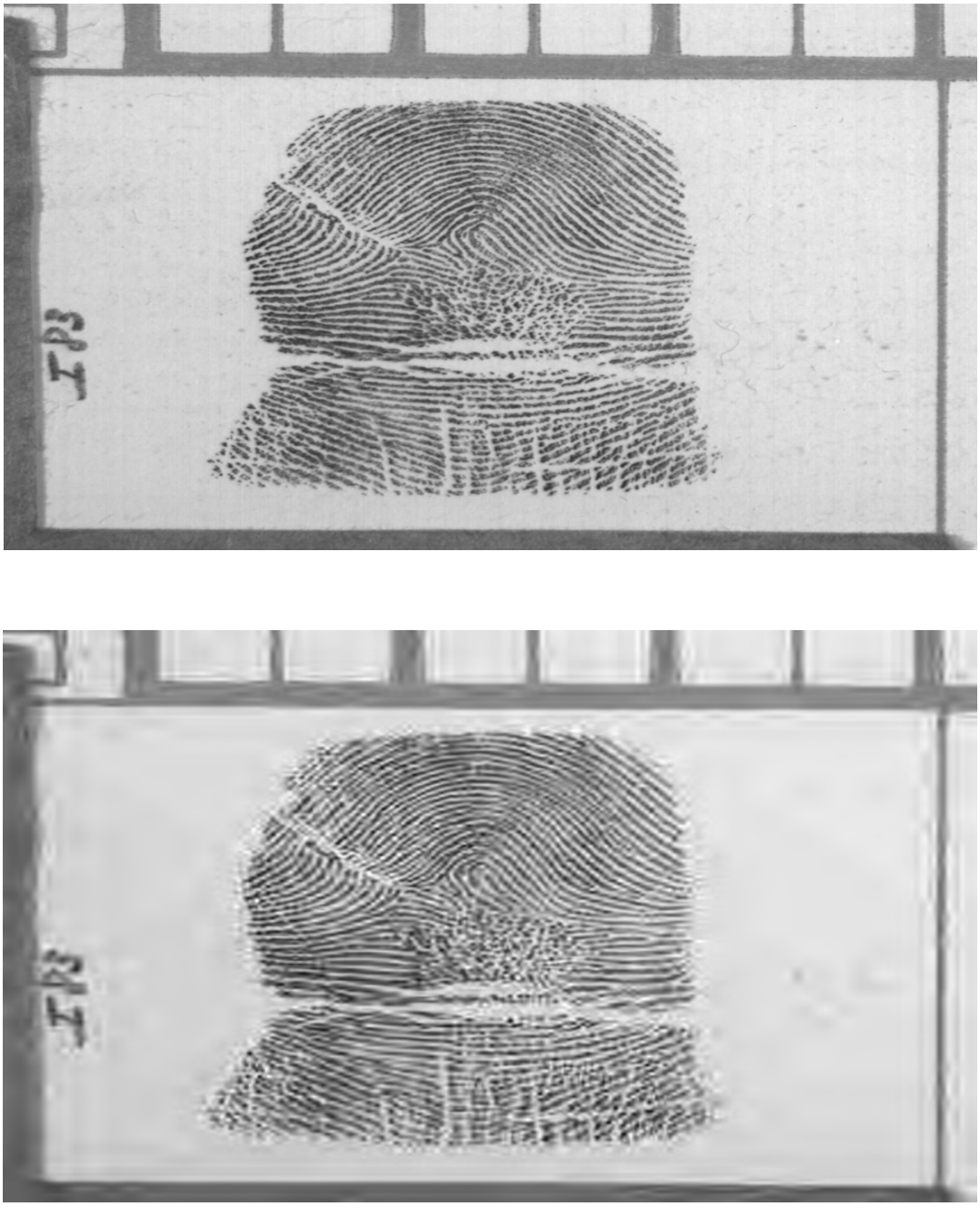

Figure 3.5: An Original Fingerprint Image (above) and its WSQ-Reconstructed at 120:1 (below).

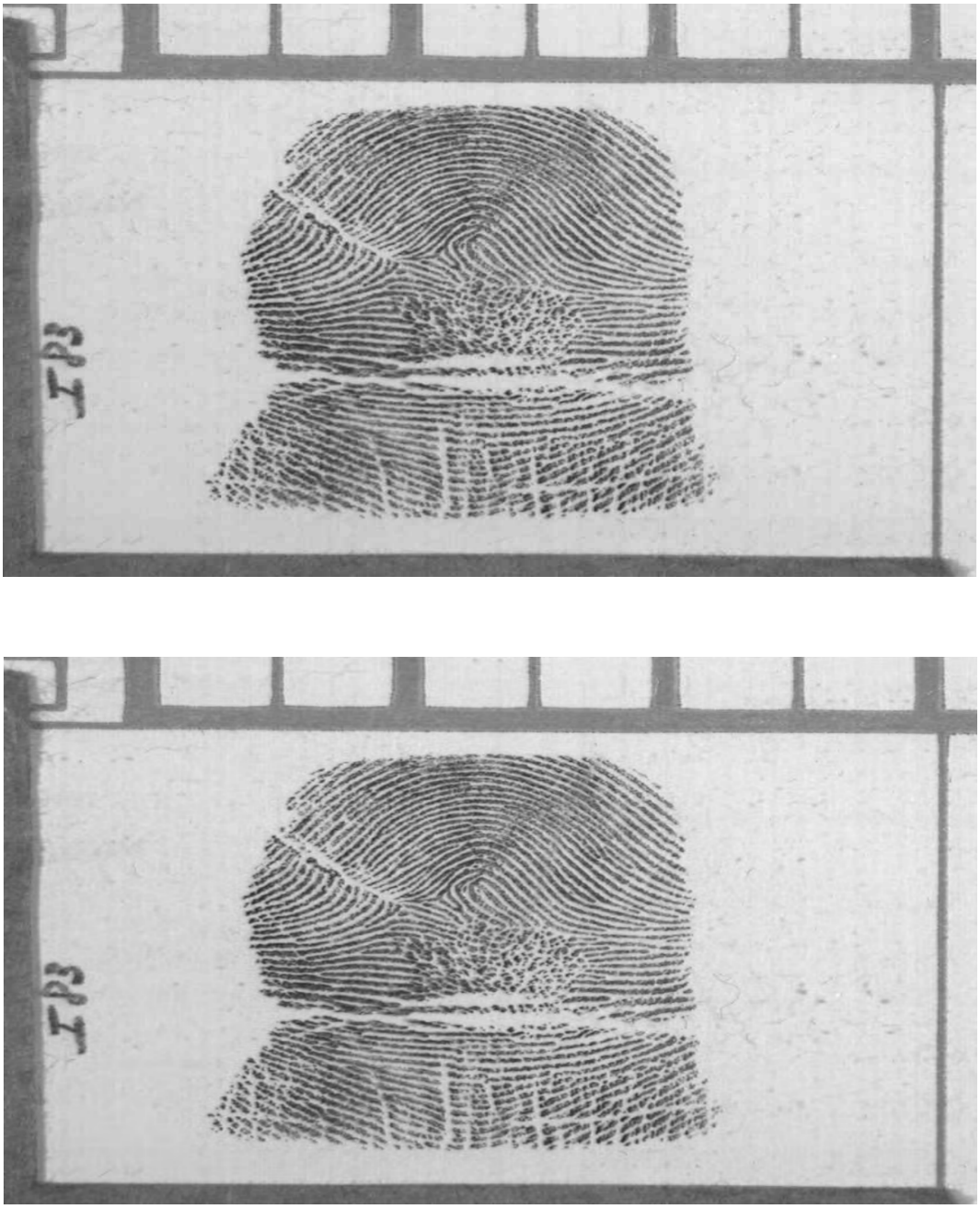

Figure 3.6: JPEG at 15:1 (above) and WSQ at 15:1 (below).

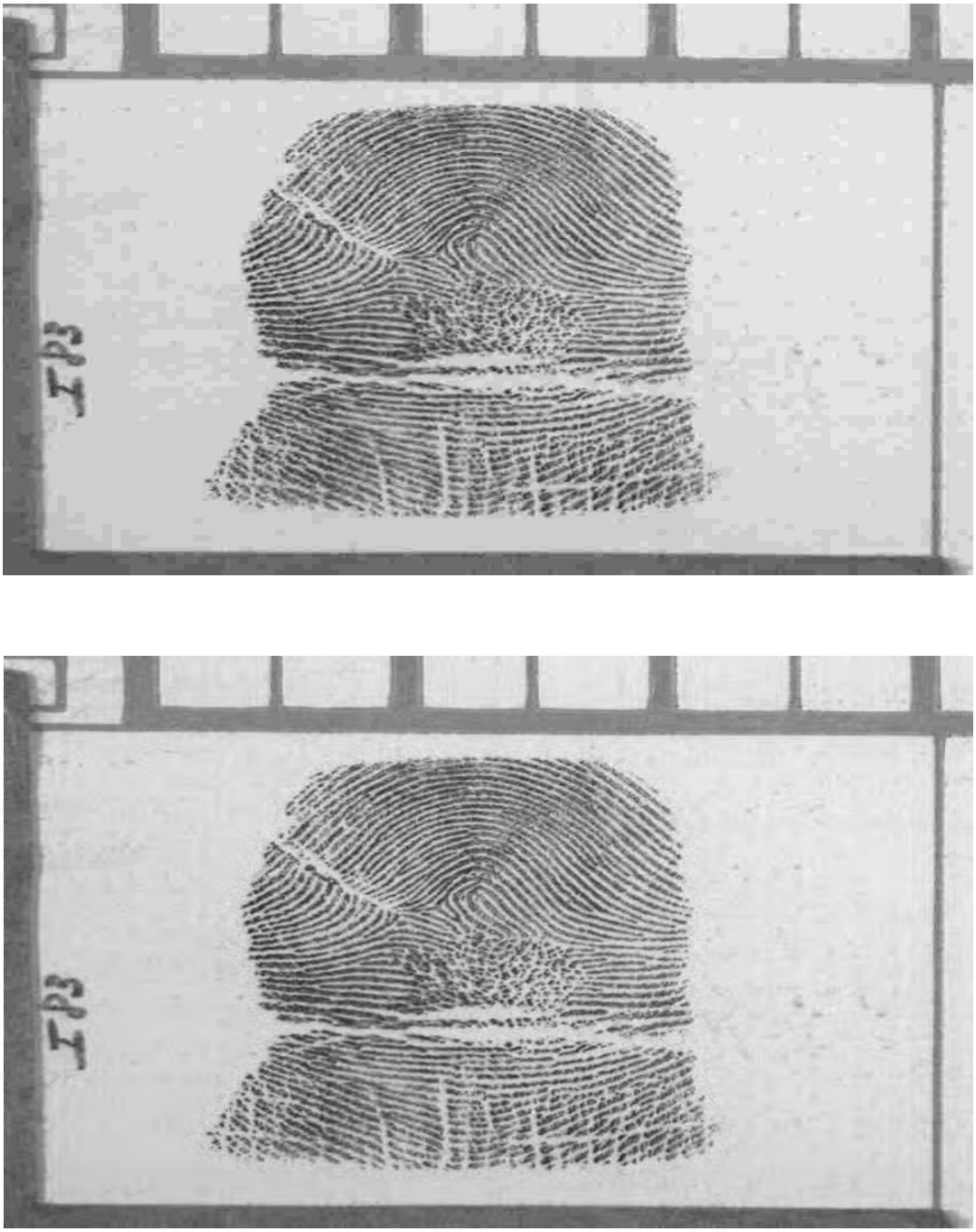

Figure 3.7: JPEG at 30:1 (above) and WSQ at 30:1 (below).

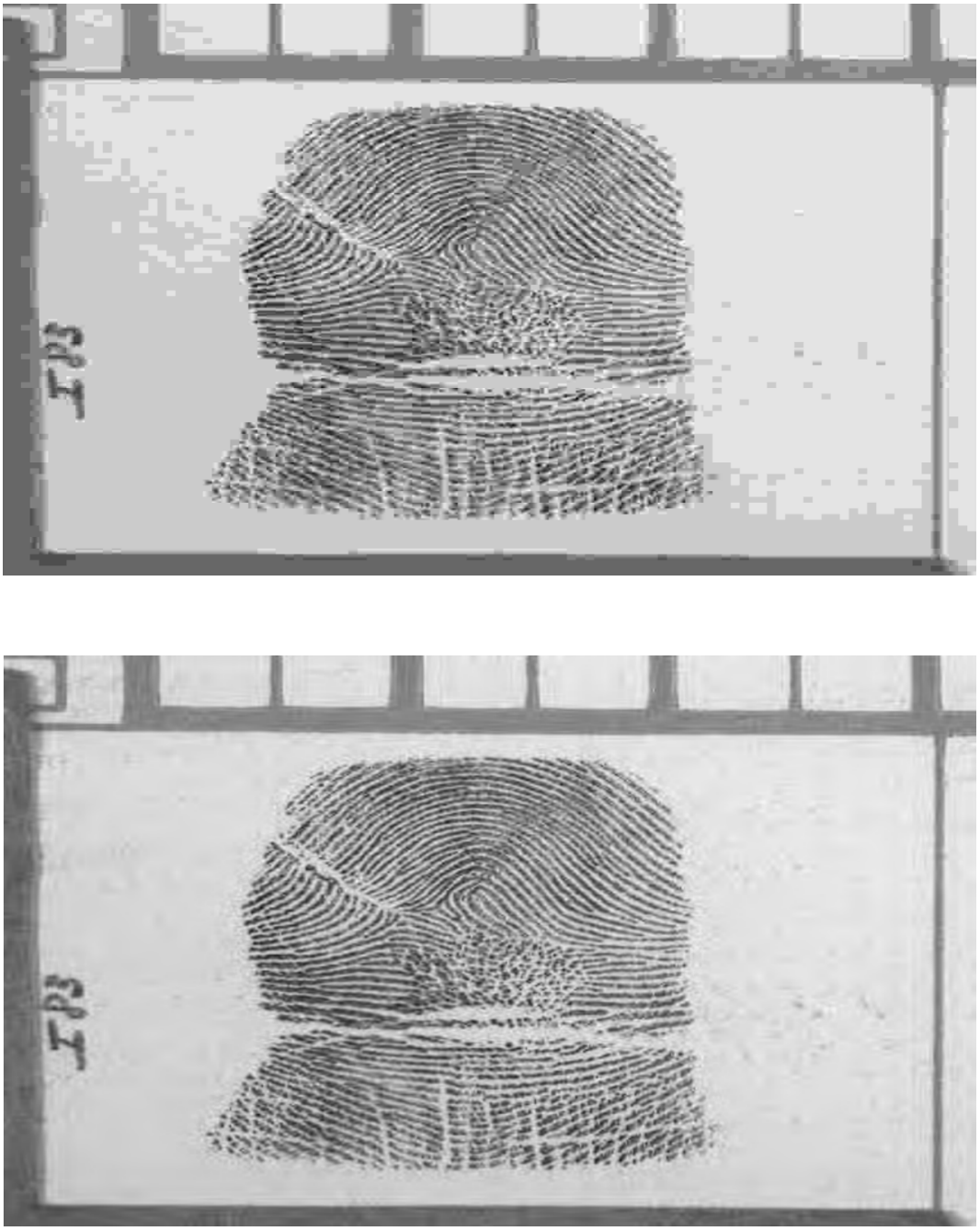

Figure 3.8: JPEG at 45:1 (above) and WSQ at 45:1 (below).

These images have been borrowed from [13].

3.2.1 Medical image processing

In medical applications, such as magnetic resonance imaging (MRI) or computed tomography (CT), wavelets are used for image enhancement, noise reduction and extraction of important features.

3.3 Data and time series analysis

Wavelets can decompose time series into different scales, allowing frequency components to be analysed and patterns in time data to be detected. As we have already seen in Section 1.4, although it is impossible to localise time and frequency simultaneously, based on the Heisenberg uncertainty principle, wavelets allow us to obtain information from both fields, unlike the Fourier transform.

3.4 Geospatial data compression and analysis

Wavelets have been used in GIS (Geographic Information Systems) applications for geospatial data compression and analysis, enabling more efficient management of large datasets.

Geospatial data compression: In the field of geographic information systems (GIS) and digital mapping, geospatial datasets can be huge due to high spatial resolution. Wavelets allow the compression of these data by decomposing them into scales and levels of detail, preserving essential information and eliminating redundancy. This allows for more efficient representation of large geographic datasets without losing crucial information. The use of wavelets in this context can be seen in [8].

Satellite image analysis: In satellite image processing, wavelets are used for noise reduction, edge detection, image segmentation and relevant feature extraction. The ability of wavelets to preserve fine details allows for more detailed analysis of satellite imagery, which is essential in applications such as environmental monitoring, precision agriculture, disaster management, among others. See [3].

Terrain and relief modelling: In topographic analysis and terrain modelling, wavelets can decompose altimetric data at different scales, allowing the identification of topographic patterns, geological features or relief changes at various spatial scales. For example, in [6].

Conclusions

In conclusion, this thesis has provided a comprehensive overview of wavelet theory and its applications. Beginning with fundamental mathematical concepts like Hilbert Spaces, the Fourier transform, and the Haar basis, we progressed to delve deeply into Wavelet Theory, exploring the Haar wavelet's Multiresolution Analysis (MRA) and Mallat's theorem.

The analysis highlighted wavelets' versatile nature, demonstrating their effectiveness in various practical applications, including signal and audio processing, image and video compression, medical image analysis, as well as data and time series analysis. Their ability to efficiently handle complex data while extracting vital information underscores their significant role in real-world problem-solving across diverse domains.

On a personal note, the objectives I had for this work have been more than achieved. The main goal was to study something with some practical applications, and as we have already seen, wavelets have many. Analysing the theoretical basis of wavelets, understanding them and seeing some of their applications has been very enriching for me. Surprisingly, while the primary goal was to introduce Wavelet Theory, I realized I've indirectly employed wavelets throughout this endeavor, like when working with images, downloading, and sharing PDFs, etc. This exploration has illuminated a mathematical domain previously unfamiliar to me, fascinating in its potential for new advancements and interdisciplinary applications.

Appendix A

Python code

In this appendix we present the code created to represent a function $f(x)$ using the Haar wavelet. In Figure 2.1 it has been used to represent the function $f(x) = \frac{1}{2}\cos(x)\sin(x^2+5)$ at different resolutions, but of course it can be used for other functions and resolutions.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


def f(x):
  return 0.5 * np.cos(x) * np.sin(x**2 + 5)


def harr(a, len, x):
  leftmost_point = a + len * ((x - a) // len)
  rightmost_point = a + len * ((x - a) // len) + len
  return (f(leftmost_point) + f(rightmost_point)) / 2


def harr_plotter(a=0, b=5, step=0.001, subdivisions=512):
  interval = np.arange(a, b, step)
  len = (b - a) / subdivisions

  # This plots the function with the harr function
  df6 = pd.DataFrame({'x': interval, 'y': harr(a, len, interval)})
```

```
    df6.plot(kind="scatter", x="x", y="y", s=1)
    plt.gca().set_aspect(1.5)
    # Remove top border and right border
    ax = plt.gca()
    ax.spines['top'].set_visible(False)
    ax.spines['right'].set_visible(False)
    # Remove x and y axis titles
    plt.xlabel('')
    plt.ylabel('')
    # Set y ticks to every 0.5
    plt.yticks(np.arange(-0.5, 0.51, 0.5))

    plt.show()


def function_plotter(interval):
    data = f(interval)
    function = {'x': interval, 'y': list(data)}
    df = pd.DataFrame(function)
    # This plots the function as is
    plot = df.plot(kind="scatter", x="x", y="y", s=1, title='')
    plt.gca().set_aspect(1.5)
    # Remove top border and right border
    ax = plt.gca()
    ax.spines['top'].set_visible(False)
    ax.spines['right'].set_visible(False)
    # Remove x and y axis titles
    plt.xlabel('')
    plt.ylabel('')
    # Set y ticks to every 0.5
    plt.yticks(np.arange(-0.5, 0.51, 0.5))
    return plot


if __name__ == "__main__":
    a = 0
    b = 5
    step = 0.001
    plot = function_plotter(interval=np.arange(a, b, step))
```

```
plt.show()

for i in range(0, 10):
  harr_plotter(a, b, step, subdivisions=2**i)
```

Bibliography

[1] B. Burke Hubbard, *The world according to wavelets. The story of a mathematical technique in the making*, Second Edition, A K Peters, 1998.

[2] I. Daubechies, *Ten lectures on wavelets*, CBMS-NSF Regional Conf. Series in Apl. Math, SIAM, 1992.

[3] H. Demirel and G. Anbarjafari, "Satellite Image Resolution Enhancement Using Complex Wavelet Transform", *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp 123-126, 2010.

[4] J. Duoandikoetxea. *Lecciones sobre las series y transformadas de Fourier*. UNAN-Managua, 2003.

[5] E. Hernández, "Ondículas: historia, teoría y aplicación", *La Gaceta de la RSME*, vol.21, no. 2, 275-299, 2018.

[6] G. Jordan and B. Schott, "Application of wavelet analysis to the study of spatial pattern of morphotectonic lineaments in digital terrain models. A case study", *Remote Sensing of Environment*, vol. 94, no 1, pp 31-38, 2005.

[7] A., Kasnazani and A., AliPanah, "Solving brachistochrone problem via scaling functions of Daubechies wavelets", *Computational Methods for Differential Equations*, vol.9, no 2, pp 511-522, 2020.

[8] T. Lin, R. K. Aggarwal and C. H. Kim, "Wavelet transform and artificial intelligence based condition monitoring for GIS", *2003 IEEE PES Transmission and Distribution Conference and Exposition* (IEEE Cat. No.03CH37495), vol.1, pp 191-195, 2003.

[9] S. Mallat, "Multiresolution approximations and wavelet orthonormal bases for $L^2(\mathbb{R})$", *Trans. Amer. Math. Soc.*, vol. 315, no. 1, pp 69-87, 1989.

[10] S. Mallat, *A wavelet tour of signal processing. The sparse way*, Third Editon, with contributions from G. Peyré, Elsevier/Academic Press, Amsterdam, 2009.

[11] M. J. Mohlenkamp and M. C. Pereyra, *Wavelets, their friends, and what they can do for you*, EMS Series of Lectures in Mathematics, European Math. Soc. Publishing House, Zurich, 2008.

[12] C. V. Nastos and D. A. Saravanos, "Multiresolution Daubechies finite wavelet domain method for transient dynamic wave analysis in elastic solids", *International Journal for Numerical Methods in Engineering*, vol 122, no. 23, pp 7078 - 7100, 2021.

[13] R. J. Onyshczak and A. Youssef, "Fingerprint Image Compression and the Wavelet Scalar Quantization Specification", *Automatic Fingerprint Recognition Systems*, National Institute of Standards and Technology, pp 385-413, 2004.

[14] M. C. Pereyra and L. A. Ward, *Harmonic analysis: From Fourier to wavelets*, Student Math. Library, Vol 63, Amer. Math. Soc., 2012.

[15] N. J. S.R.F, N. S. S.R.F, S. A. Aljasar, Y. Xu and M. Saqib (2020). "Analysis and Detection of Community Acquired Pneumonia Using PSPNET with Complex Daubechies Wavelets". *Indian Journal of Computer Science and Engineering*, vol.11, no. 3, pp 217-225, 2020.

[16] J. S. Walker, *A primer on wavelets and their scientific applications*, Second Edition, Studies in Advanced Mathematics, Chapman & Hall/CRC, Boca Raton, FL, 2008.

[17] S. Zayrit, T. B. Drissi, A. Ammoumou and B. Nsiri, "Daubechies Wavelet Cepstral Coefficients for Parkinson's Disease Detection", *Complex Systems*, vol. 29, no. 3, pp 729-739, 2020.

[18] *File:Wavelet Shan.svg - Wikimedia Commons.* Accessed: Jan. 3, 2024. [Image]. Available: `https://commons.wikimedia.org/wiki/File:Wavelet_Shan.svg`

[19] *File:Daubechies4-functions.svg - Wikimedia Commons.* Accessed: Jan. 3, 2024. [Image]. Available: `https://commons.wikimedia.org/wiki/File:Daubechies4-functions.svg`