



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Wavelets

Autor: Irene Peribáñez Pujolràs

Director: Dr. Francesc Xavier Massaneda Clares

Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 9 de juny de 2026

Contents

Introduction	ii
1 Preliminaries	1
1.1 Hilbert spaces, $L^2([-\pi, \pi])$ and Fourier series	1
1.2 Fourier integral	3
1.3 Inversion formula	5
1.4 Fourier integral in L^2	7
1.5 Kotelnikov-Shannon-Whittaker sampling theorem	8
2 Wavelets	11
2.1 The Haar system	11
2.2 The Haar wavelet	14
2.3 Multi-resolution analysis and Mallat's theorem	17
2.4 Construction of a Multi-resolution analysis	22
2.5 The wavelet from the Fourier side	26
3 Wavelets in $\mathbb{R}^2$	31
3.1 Separable multi-resolutions	31
3.2 The continuous wavelet transform	36
4 Examples and applications	41
4.1 Shannon, Meyer and Daubechies wavelets	42
4.1.1 The Shannon wavelet	42
4.1.2 The Meyer wavelet	42
4.1.3 Daubechies wavelets	43
4.2 Compression and denoising	45
4.3 Signal processing	47
4.4 Image processing	47
4.4.1 JPEG 2000	48
Bibliography	51

Abstract

The main goal of this memoir is to introduce the wavelet theory, with particular emphasis in the construction of a Multi-resolution analysis (MRA). The main theoretical result is Mallat's theorem, to the effect that every MRA has an associated wavelet. Moreover, we show that these methods can be applied to signal and image processing.

L'objectiu principal d'aquesta memòria és introduir la teoria de *wavelets*, amb particular èmfasi en la construcció d'una Anàlisi Multiresolució (MRA). El resultat teòric principal és el teorema de Mallat, que estableix que tot MRA té una *wavelet* associada. A més, mostrem que aquests mètodes poden aplicar-se al processament de senyals i imatges.

Introduction

Wavelets might not be the most famous part of mathematics, but they are important tools that work quietly behind the scenes in many areas. This thesis explores wavelet theory and how wavelets are used in different applications, such as signal or image processing.

To understand why wavelets are helpful, we must look back to the early 19th century and the work of the French mathematician Joseph Fourier, who introduced the Fourier transform. This tool broke down signals in a way that shifted the perspective from the time domain to the frequency domain, changing how signals are studied. However, the Fourier transform has faced limitations when dealing with general signals.

The quest to overcome these limitations led to wavelet theory in the late 20th century. The French geophysicist Jean Morlet was the first to use the term *wavelet* and introduced the concept during his research on seismic data.

However, the theory was formalized and developed by later mathematicians like Yves Meyer, Stéphane Mallat, and Ingrid Daubechies. A wavelet is a localised wave that quickly dies out to zero, unlike an infinite sine wave. This compact shape is defined by a single, original wave called the *mother wavelet*. By shifting or stretching it, generates an entire family of wavelets. This framework forms the basis of the Multi-resolution Analysis (MRA), which allows looking at data at different scales.

The historical relevance of this mathematical concept became clear during a massive data storage crisis at the end of the 20th century. It involved the FBI, which was running out of space to store millions of physical fingerprint cards. Mathematicians developed a system using wavelets that compressed the fingerprint files by a factor of 15-to-1 while keeping the ridges perfectly clear. This success shaped the modern JPEG 2000 image standard.

Motivated by this link between theory and practical application, this thesis goes through the mathematics behind wavelets and explores some of their possible implementations.

The initial chapter introduces the concepts needed to understand wavelet theory. It begins with an introduction on Hilbert spaces, focusing on the concepts of inner products and orthonormality. Following this, the Fourier transform and its inversion formula are stated. The chapter finishes with an analysis of the Fourier transform on $L^2(\mathbb{R})$ and provides an application through the Kotelnikov-Shannon-Whittaker sampling theorem.

The second chapter introduces the concept of a Multi-resolution Analysis (MRA). Instead of presenting definitions and results, an intuitive example is provided first, to illustrate the main ideas. This example is the Haar system and its correspond-

ing wavelet. Following this example, we develop the general framework of an MRA and state Mallat's theorem regarding the existence of an associated wavelet. This is the main theorem proved in this memoir. After discussing the explicit construction of an MRA from a scaling function, the chapter concludes by developing properties of the scaling function on the Fourier side, that is, on the frequency domain.

The objective of the third chapter is to extend wavelet theory to functions in $L^2(\mathbb{R}^2)$. To avoid distortions across two independent dimensions, we introduce separable multi-resolutions, constructed as tensor products of one-dimensional multi-resolution spaces. The chapter concludes with the definition of a continuous wavelet transform, which is independent of the MRA structure.

The final chapter presents several wavelet families alongside their applications in signal and image processing. The exposition begins with classical wavelet families and how their properties suit for specific types of problems. Afterwards, the focus shifts to digital signal processing, illustrating how wavelets are used to compress and denoise. The chapter concludes with an examination of image processing, highlighting the role of wavelets in the JPEG 2000 standard, and providing a comparative analysis with the traditional JPEG format.

Chapter 1

Preliminaries

This chapter introduces the notation and the concepts used throughout the thesis. It begins with a brief review of Hilbert spaces and Fourier series on $L^2([-\pi, \pi])$. The discussion then moves to the Fourier integral, the inversion formula, and the extension of these concepts to $L^2(\mathbb{R})$. The chapter concludes with one important application, the Kotelnikov-Shannon-Whittaker sampling theorem.

The purpose of this chapter is to recall standard results, establish a consistent notation, and present the framework required for the study of wavelets and their applications in the following chapters. Proofs are omitted, since the material presented here is the foundational background rather than the central focus of the thesis. These results have been covered in undergraduate courses. Complete proofs can be found in [7].

1.1 Hilbert spaces, $L^2([-\pi, \pi])$ and Fourier series

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

1. Is sesquilinear: linear in the first variable and conjugate-linear in the second.
2. Is Hermitian: $\langle x, y \rangle = \overline{\langle y, x \rangle}$.
3. Is positive definite: $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

Moreover, H is a complete metric space with the distance induced by the norm $\|x\| = \sqrt{\langle x, x \rangle}$. With the norm $\|x\|$, H has the structure of a Banach space.

Lemma 1.2 (Cauchy-Schwarz inequality). Let $(H, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Then, for all $x, y \in H$, $|\langle x, y \rangle| \leq \|x\| \|y\|$.

The next examples are presented here due to their relevance for the results and constructions introduced in the following sections.

Example 1.3. 1. The space $l^2(\mathbb{Z}) = \{(a_n)_{n \in \mathbb{Z}} \subset \mathbb{C} : \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty\}$ is a Hilbert space with the inner product

$$\langle a, b \rangle = \sum_{n=-\infty}^{\infty} a_n \overline{b_n}, \quad a = (a_n)_{n \in \mathbb{Z}}, \quad b = (b_n)_{n \in \mathbb{Z}}.$$

In this case, the norm is defined as $\|a\| = (\sum_{n \in \mathbb{Z}} |a_n|^2)^{1/2}$.

2. The space $L^2([-\pi, \pi]) = \left\{ f : [-\pi, \pi] \rightarrow \mathbb{C} \mid \|f\|_2^2 = \int_{-\pi}^{\pi} |f(x)|^2 dx < \infty \right\}$ is a Hilbert space with the inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} dt \quad \forall f, g \in L^2([-\pi, \pi]).$$

In this case, the norm is defined as $\|f\|_2 = \sqrt{\langle f, f \rangle} = (\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 dt)^{1/2}$.

3. One can extend the space $L^2([-\pi, \pi])$ to the whole real line, that is

$$L^2(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C} \mid \|f\|_2^2 = \int_{\mathbb{R}} |f(x)|^2 dx < \infty \right\}$$

with the inner product

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

Remark 1.4. The space $L^2([-\pi, \pi])$ can be rescaled to adapt it to any period of the function f . This rescaling requires modifying the inner product accordingly.

At this point, it is natural to seek a representation of elements of H that is both unique and useful. This motivates the introduction of an orthonormal system $\{e_i\}_{i \in I}$ for H . The system completeness ensures that every element of H can be expanded in terms of this system. Then, one has the following identity.

Theorem 1.5. Let $\{e_n\}_{n \in \mathbb{N}}$ be a complete orthonormal system for H . Then, every $x \in H$ satisfies Parseval's identity $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2$.

One can define an explicit complete orthonormal system for the case where we take $L^2([-\pi, \pi])$ as the Hilbert space.

Proposition 1.6. The system $\{e^{int}\}_{n \in \mathbb{Z}}$ is an orthonormal basis of $L^2([-\pi, \pi])$. In particular, every $f \in L^2([-\pi, \pi])$ admits the Fourier series expansion

$$f(t) = \sum_{n \in \mathbb{Z}} c_n e^{int}, \quad (1.1)$$

where the coefficients c_n are defined as

$$c_n = \hat{f}(n) = \langle f, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt.$$

The series in (1.1) is usually denoted by $Sf(t)$ and is referred to as the *Fourier series* of f .

Remark 1.7. 1. The equality (1.1) holds in $L^2([-\pi, \pi])$, which means that

$$\lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} |f(t) - \sum_{|n| \leq N} \hat{f}(n) e^{int}|^2 dt = 0$$

2. As a consequence of Parseval's identity, one has

$$\|f\|_{L^2([-\pi, \pi])}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 dt = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2.$$

1.2 Fourier integral

The Fourier integral can be viewed as the limit of the Fourier series as the period tends to infinity. In this limit, the discrete set of frequencies becomes continuous, and the series representation is replaced by an integral. As a consequence, one considers the space of integrable functions in $\mathbb{R}$ defined as

$$L^1(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C} \mid \|f\|_1 = \int_{\mathbb{R}} |f(x)| dx < \infty \right\}.$$

Definition 1.8. Let $f \in L^1(\mathbb{R})$. Its *Fourier transform* is the function $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$ defined as

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

This expression represents the contribution of the frequency ξ to the signal $f(t)$.

Remark 1.9. A pure frequency component of the form $e^{2\pi i \xi t}$ oscillates with period $1/|\xi|$. Thus, when $|\xi|$ is small, the corresponding exponential varies slowly in time, while large values of $|\xi|$ produce rapid oscillations.

Let us recall some elementary properties.

Proposition 1.10. Let $f, g \in L^1(\mathbb{R})$, $\alpha, \beta \in \mathbb{C}$, and $a \in \mathbb{R}$. Then the Fourier transform satisfies the following properties:

1. Linearity: $\widehat{(\alpha f + \beta g)}(\xi) = \alpha \widehat{f}(\xi) + \beta \widehat{g}(\xi)$, $\xi \in \mathbb{R}$.
2. Conjugation: $\widehat{\overline{f}}(\xi) = \overline{\widehat{f}(-\xi)}$, $\xi \in \mathbb{R}$.
3. Translation: Let $\tau_a f(t) = f(t - a)$. Then $\widehat{(\tau_a f)}(\xi) = \widehat{f}(\xi) e^{-2\pi i a \xi}$, $\xi \in \mathbb{R}$.
4. Modulation: Let $M_a f(t) = f(t) e^{2\pi i a t}$. Then $\widehat{(M_a f)}(\xi) = T_a \widehat{f}(\xi) = \widehat{f}(\xi - a)$, $\xi \in \mathbb{R}$.
5. Dilation: Given $\lambda > 0$, let $D_\lambda f(t) = \frac{1}{\lambda} f(\frac{t}{\lambda})$. Then $\widehat{D_\lambda f}(\xi) = \widehat{f}(\lambda \xi)$, $\xi \in \mathbb{R}$.

A fundamental virtue of the Fourier transform, and indeed the reason Fourier introduced it in the study of the heat equation, is that it converts certain complicated operations, most notably differentiation and convolution, into simpler ones.

Theorem 1.11. Let $f \in L^1(\mathbb{R})$. Then:

1. $\widehat{f}$ is uniformly continuous and $|\widehat{f}(\xi)| \leq \|f\|_1$.
2. If $f, f' \in L^1(\mathbb{R})$, then $\widehat{f'}(\xi) = 2\pi i \xi \widehat{f}(\xi)$, $\xi \in \mathbb{R}$.
3. If $tf(t) \in L^1(\mathbb{R})$, then $\widehat{f}$ is differentiable and $(\widehat{f})'(\xi) = \widehat{(-2\pi i t f)}(\xi)$, $\xi \in \mathbb{R}$.
4. It satisfies the *Riemann-Lebesgue lemma*: $\lim_{|\xi| \rightarrow \infty} \widehat{f}(\xi) = 0$.
5. It satisfies the *multiplication formula*: if $f, g \in L^1(\mathbb{R})$, then $\int_{\mathbb{R}} f(t) \widehat{g}(t) dt = \int_{\mathbb{R}} \widehat{f}(t) g(t) dt$.
6. If $f^{(j)} \in L^1(\mathbb{R})$ for all $0 \leq j \leq k$, then $\widehat{f}(\xi) = o\left(\frac{1}{|\xi|^k}\right)$ as $|\xi| \rightarrow \infty$.

Remark 1.12. Property 2 can be applied iteratively: if $f, f', \dots, f^{(k)} \in L^1(\mathbb{R})$, then $\widehat{f^{(k)}}(\xi) = (2\pi i \xi)^k \widehat{f}(\xi)$.

Example 1.13. 1. Let $f = \chi_{[-1/2, 1/2]}$. Clearly $f \in L^1(\mathbb{R})$ and $\|f\|_1 = 1$. Its Fourier transform is

$$\widehat{f}(\xi) = \int_{-1/2}^{1/2} e^{-2\pi i \xi t} dt = \left[\frac{e^{-2\pi i \xi t}}{-2\pi i \xi} \right]_{t=-1/2}^{t=1/2} = \frac{e^{-\pi i \xi} - e^{\pi i \xi}}{-2\pi i \xi} = \frac{\sin(\pi \xi)}{\pi \xi}.$$

We introduce the cardinal sine function $\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$, so that $\widehat{f}(\xi) = \text{sinc}(\xi)$.

2. Consider the Gaussian function $G(t) = e^{-\pi t^2}$, which is normalised so that $\|G\|_1 = 1$. A direct computation of $\widehat{G}(\xi)$ leads to a complicated integral, so a different approach is used, based on the relation between the Fourier transform and differentiation.

The function G satisfies the differential equation

$$\begin{cases} G'(t) = -2\pi t G(t) \\ G(0) = 1. \end{cases}$$

Taking the Fourier transform of both sides of the equation yields $\widehat{G'}(\xi) = (-2\pi t G(t))^\wedge(\xi)$. By Theorem 1.11, one has $(2\pi i \xi) \widehat{G}(\xi) = -(i \widehat{G})'(\xi)$. Since

$$\widehat{G}(0) = \int_{\mathbb{R}} G(t) dt = 1,$$

it follows that $\widehat{G}$ satisfies the same differential equation and initial condition as G . Therefore, $\widehat{G}(\xi) = e^{-\pi \xi^2}$.

1.3 Inversion formula

The objective here is to determine under which conditions a function $f \in L^1(\mathbb{R})$ can be recovered from the values of its Fourier transform $\widehat{f}(\xi)$. From the perspective of signal processing, this is natural, since the frequency distribution determines the signal. A fundamental tool in proving this result is the convolution of functions.

Definition 1.14. Let $f, g \in L^1(\mathbb{R})$. The *convolution* of f and g is the function $f * g$ defined by

$$(f * g)(t) = \int_{\mathbb{R}} f(s) g(t - s) ds, \quad t \in \mathbb{R}.$$

In many contexts, convolution is used because it provides a way to replace a function by a local average of its values, producing a smoother version of the original one while preserving its fundamental features.

Lemma 1.15. If $f, g \in L^1(\mathbb{R})$, then $f * g = g * f$, the convolution $f * g$ belongs to $L^1(\mathbb{R})$, and $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$. That is, the convolution of two $L^1(\mathbb{R})$ -functions is well-defined.

A fundamental property is the following.

Theorem 1.16. Let $f, g \in L^1(\mathbb{R})$. Then, for $\xi \in \mathbb{R}$ one has $\widehat{(f * g)}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$.

The convolution $f * g$ with a non negative function g satisfying $\|g\|_1 = 1$ may be interpreted as a weighted average of the values of f , where the weights are given by g . When g is concentrated near 0, the function $f * g$ provides a smoothed version of f that retains its essential features (see Proposition 1.19 below).

Definition 1.17. Let $g \in L^1(\mathbb{R})$ be non negative and satisfy $\|g\|_1 = 1$. For $\delta > 0$, define the *dilations*

$$g_\delta(t) = D_\delta g(t) = \frac{1}{\delta} g\left(\frac{t}{\delta}\right).$$

The family $\{g_\delta\}_{\delta>0}$ is called an *approximate identity* if for every $\eta > 0$,

$$\int_{|t|>\eta} g_\delta(t) dt \longrightarrow 0 \quad \text{as } \delta \rightarrow 0. \quad (1.2)$$

There is a more general notion of approximate identity, but only this specific form will be used here.

Remark 1.18. The condition (1.2) implies that g_δ becomes increasingly concentrated around 0 as δ tends to 0.

Proposition 1.19. Let $f \in L^1(\mathbb{R})$ and let $\{g_\delta\}_{\delta>0}$ be an approximate identity. Then

$$\lim_{\delta \rightarrow 0} \|f * g_\delta - f\|_1 = 0.$$

In particular, $\lim_{\delta \rightarrow 0} (f * g_\delta)(t) = f(t) \quad \text{for a.e. } t \in \mathbb{R}.$

The following results show that the Fourier transform completely determines the original function and allows it to be recovered from its frequency representation.

Theorem 1.20 (The inversion formula for L^1 functions). Let $f \in L^1(\mathbb{R})$ be such that $\widehat{f} \in L^1(\mathbb{R})$. Then

$$f(t) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i t \xi} d\xi \quad \text{a.e. } t \in \mathbb{R}.$$

Moreover, the identity holds at every point t where f is continuous.

Notice that if $f, \hat{f} \in L^1(\mathbb{R})$ then, by applying this result twice, we have $f(t) = (\hat{\hat{f}})(-t)$ a.e. $t \in \mathbb{R}$. Moreover, the right-hand side defines a continuous function of t . Thus, in general, the equality cannot hold for all $t \in \mathbb{R}$. Furthermore, if $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$, then $|f(t)| \leq \|\hat{f}\|_1$ holds for a.e. $t \in \mathbb{R}$.

Corollary 1.21 (Uniqueness theorem). If $f \in L^1(\mathbb{R})$ is such that $\hat{f} = 0$ a.e. $\xi \in \mathbb{R}$, then $f = 0$ a.e. $t \in \mathbb{R}$.

1.4 Fourier integral in L^2

It is natural to make use of the Hilbert space structure of $L^2(\mathbb{R})$ in Fourier analysis. An initial obstacle is that functions in $L^2(\mathbb{R})$ do not necessarily belong to $L^1(\mathbb{R})$, which requires additional care.

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

Theorem 1.22 (Plancherel's theorem). Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then $\hat{f} \in L^2(\mathbb{R})$ and $\|f\|_2 = \|\hat{f}\|_2$. In particular, if $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then

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

Thus, the Fourier transform acts as an isometry on $L^2(\mathbb{R})$.

At this point, one wants to get rid of the assumption that $f \in L^1(\mathbb{R})$. Given $f \in L^2(\mathbb{R})$, one can construct a sequence $\{f_n\}_n \subset L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\|f_n - f\|_2 \rightarrow 0$. By Plancherel's theorem, the sequence $\{\hat{f}_n\}_n$ is Cauchy in $L^2(\mathbb{R})$. Thus, one can define the Fourier transform of f as $\hat{f}(\xi) = \lim_{n \rightarrow \infty} \hat{f}_n(\xi)$, with convergence in the L^2 sense. This definition is independent of the particular approximating sequence. If $\{g_n\}_n \subset L^1 \cap L^2$ is any other such sequence then, by Plancherel,

$$\|\hat{f}_n - \hat{g}_n\|_2 = \|f_n - g_n\|_2 \leq \|f_n - f\|_2 + \|f - g_n\|_2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

One may summarise these facts in the following statement.

Theorem 1.23. Let $f \in L^2(\mathbb{R})$.

1. The Fourier transform of f is given by

$$\hat{f}(\xi) = \lim_{n \rightarrow \infty} \int_{-n}^n f(t) e^{-2\pi i t \xi} dt,$$

with convergence in the L^2 sense.

2. If $g \in L^2(\mathbb{R})$, then

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

and

$$\int_{\mathbb{R}} f(t) \widehat{g}(t) dt = \int_{\mathbb{R}} \widehat{f}(t) g(t) dt.$$

With this interpretation, the inversion formula also holds in $L^2(\mathbb{R})$. Let us first state the following property of the convolution.

Lemma 1.24. Let $f \in L^2(\mathbb{R})$ and $g \in L^1(\mathbb{R})$. Then $f * g \in L^2(\mathbb{R})$, $\|f * g\|_2 \leq \|f\|_2 \|g\|_1$, and $\widehat{f * g}(\xi) = \widehat{f}(\xi) \widehat{g}(\xi)$.

Theorem 1.25. Let $f \in L^2(\mathbb{R})$. Then

$$f(t) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i \xi t} d\xi = \lim_{n \rightarrow \infty} \int_{-n}^n \widehat{f}(\xi) e^{2\pi i \xi t} d\xi,$$

with convergence in the L^2 sense.

Moreover, the identity holds pointwise at every point where f is continuous.

1.5 Kotelnikov-Shannon-Whittaker sampling theorem

The following fundamental application of Fourier analysis plays a central role in digital signal processing, as it provides a condition on the sampling rate required to reconstruct a continuous-time signal of finite bandwidth from its discrete samples.

Assume that $f(t)$ is a continuous signal of finite energy, that is, $f \in L^2(\mathbb{R})$. Suppose also that f has finite bandwidth, that is, there exists $\tau > 0$ such that $\text{supp } \widehat{f} \subset [-\tau, \tau]$. This assumption is natural in many practical situations. First, the range of frequencies perceptible by the human ear is limited (approximately from 20 Hz to 20,000 Hz). Second, physical transmission media tend to attenuate extreme frequencies, restricting the usable bandwidth of the signal.

Theorem 1.26 (Kotelnikov-Shannon-Whittaker sampling theorem). Let $f \in L^2(\mathbb{R})$ with $\text{supp } \widehat{f} \subset [-\tau, \tau]$. Then f can be completely reconstructed from its samples $\{f(k/2\tau)\}_{k \in \mathbb{Z}}$ via the cardinal series

$$f(t) = \sum_{k \in \mathbb{Z}} f\left(\frac{k}{2\tau}\right) \text{sinc}(2\tau t - k),$$

where $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$. Moreover,

$$\|f\|_2^2 = \int_{\mathbb{R}} |f(t)|^2 dt = \frac{1}{2\tau} \sum_{k \in \mathbb{Z}} \left| f\left(\frac{k}{2\tau}\right) \right|^2.$$

This statement is sometimes referred to as the *Fundamental theorem of information theory*. It allows to encode a signal through a sequence of numbers from which it can be completely recovered.

Chapter 2

Wavelets

A large class of signals, such as sound, exhibit slowly varying behaviour. In contrast, images concentrate most of their relevant information along discontinuities and regions of non-smooth behaviour. Fourier methods handle abrupt variations inefficiently, because exponentials are not spatially localised. The aim of this chapter is to represent functions in $L^2(\mathbb{R})$ by means of a family of well-localised functions called *wavelets*.

The first section introduces the Haar system, which provides an alternative orthonormal basis for $L^2([0, 1])$. The second section builds on this construction to define the Haar wavelet and to illustrate how dyadic scaling produces different levels of resolution.

The third section develops the general concept of a multi-resolution analysis (MRA) and explains Mallat's theorem, which guarantees that every MRA admits an associated wavelet. The fourth section explains how to construct an MRA starting from a candidate scaling function.

The final section turns to the Fourier side of wavelet theory and explains how Fourier analysis can be used to describe and construct both scaling functions and wavelets.

2.1 The Haar system

Proposition 1.6 provided an orthonormal basis for $L^2([0, 1])$, once properly rescaled. Let us now turn to an orthonormal basis for $L^2([0, 1])$ of a different nature.

We start by defining the constant function $\chi \equiv \mathbf{1}$, which obviously belongs to

$L^2([0, 1])$. Let also

$$\psi_0(t) = \begin{cases} -1 & \text{if } 0 \leq t < \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} \leq t < 1. \end{cases}$$

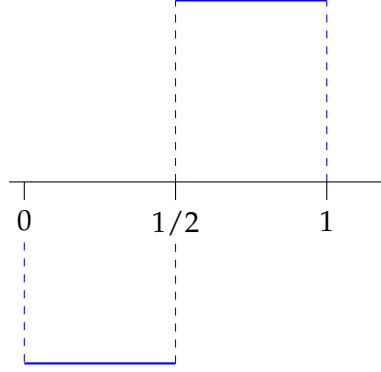


Figure 2.1: Plot of $\psi_0(t)$

Functions that remain constant on each half of the interval $[0, 1]$ can be written as linear combinations of χ and ψ_0 , since

$$\chi_{[0, \frac{1}{2})} = \frac{1}{2}(\chi - \psi_0), \quad \text{and} \quad \chi_{[\frac{1}{2}, 1)} = \frac{1}{2}(\chi + \psi_0).$$

We also consider the dyadic rescalings and translations of this function.

Definition 2.1. For every $n \geq 0$ and $k = 0, \dots, 2^n - 1$, we denote by $I_{n,k} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$ the *dyadic intervals of the n^{th} generation*.

Now take for every $n \geq 0$ and $k = 0, \dots, 2^n - 1$ the function

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

The same idea can be extended to dyadic rescalings and translations of this function. Functions that are constant on each $I_{n,k}$ are obtained as linear combinations of dilated and translated copies of χ and ψ_0 . This observation will be used in the proof of Theorem 2.2, where we show that the system formed by χ and $\{\psi_{n,k}\}_{n,k}$ forms a complete orthonormal basis of $L^2([0, 1])$.

The function χ captures the global behaviour of a function on $[0, 1]$, while the functions $\psi_{n,k}$ describe the local variations inside each interval $I_{n,k}$. Increasing n provides finer details of the function.

Theorem 2.2. With the same notation above, we have:

1. Each $\psi_{n,k}$ is supported in $I_{n,k}$.
2. $\int_0^1 \psi_{n,k}(t) dt = 0 \quad \forall n \geq 0 \text{ and } k = 0, \dots, 2^n - 1.$
3. The system $\{\chi, \psi_{n,k}\}_{n,k}$ is a complete orthonormal basis of $L^2([0, 1])$.

Proof. Items 1 and 2 follow directly from the definition.

To prove item 3, first notice that $\|\psi_{n,k}\|_{L^2([0,1])} = 1$ for every $n, k \in \mathbb{Z}$ and $\langle \chi, \psi_{n,k} \rangle = 0$ for every $n, k \in \mathbb{Z}$. Hence, it is enough to prove that $\langle \psi_{n,k}, \psi_{m,j} \rangle = 0$ for $(n, k) \neq (m, j)$ and that $\{\chi, \psi_{n,k}\}_{n,k}$ spans $L^2([0, 1])$.

First, let us prove that $\langle \psi_{n,k}, \psi_{m,j} \rangle = 0$ for $(n, k) \neq (m, j)$. This is direct in the cases where $I_{n,k} \cap I_{m,j} = \emptyset$, since their respective supports are disjoint. By construction of the dyadic intervals, if $I_{n,k} \cap I_{m,j} \neq \emptyset$ then either $I_{n,k} \subset I_{m,j}$ or $I_{m,j} \subset I_{n,k}$.

Suppose that $I_{n,k} \subset I_{m,j}$, so $n > m$ (the other case would be analogous). In this situation, $I_{n,k}$ is contained in one of the two halves of $I_{m,j}$, hence $\psi_{m,j}$ is constant on the support of $\psi_{n,k}$. It would be $-2^{m/2}$ or $2^{m/2}$, depending on which of the two halves of $I_{m,j}$ is $I_{n,k}$ contained. Then, by 1 and 2 it follows that

$$\langle \psi_{n,k}, \psi_{m,j} \rangle = \pm 2^{m/2} \int_0^1 \psi_{n,k}(t) dt = 0.$$

It remains to prove that $\{\chi, \psi_{n,k}\}_{n,k}$ is complete. Since continuous functions on $[0, 1]$ are dense in $L^2([0, 1])$, it suffices to show that the system is complete for this subspace. Moreover, it is enough to approximate functions which are constant on dyadic intervals of the same generation, as the following lemma proves.

Lemma 2.3. Let $f \in \mathcal{C}([0, 1])$. Then for all $\varepsilon > 0$ there exist $n \geq 1$ and $g_n = \sum_{k=0}^{2^n-1} \alpha_k \chi_{I_{n,k}}$ with $\alpha_k \in \mathbb{C}$, such that

$$\|f - g_n\|_\infty = \sup_{x \in [0,1]} |f(x) - g_n(x)| < \varepsilon.$$

In particular, $\|f - g_n\|_{L^2([0,1])} < \varepsilon$. In other words, g_n can be interpreted as a piecewise constant approximation of f at a scale fine enough to guarantee an error smaller than ε .

Proof. Since $f \in \mathcal{C}([0, 1])$, it is uniformly continuous on $[0, 1]$. Hence, for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $t, t' \in [0, 1]$

$$|t - t'| < \delta \Rightarrow |f(t) - f(t')| < \varepsilon.$$

Take $n \geq 1$ such that $2^{-n} < \delta$ and let $\alpha_k = f(t_k)$ for an arbitrary choice of $t_k \in I_{n,k}$. Let g_n be the function defined in the statement.

Now fix $t \in [0, 1]$. Since the intervals $\{I_{n,k}\}_{k=0}^{2^n-1}$ form a partition of $[0, 1]$, there exists a unique index k such that $t \in I_{n,k}$. For this index, one has $g_n(t) = \alpha_k = f(t_k)$. Moreover, $|t - t_k| < 2^{-n} < \delta$, and therefore

$$|f(t) - g_n(t)| = |f(t) - f(t_k)| < \varepsilon.$$

Since t was arbitrary, it follows that $\|f - g_n\|_\infty < \varepsilon$, as required. $\square$

It remains to prove that the system $\{\psi_{n,k}\}_{n,k}$ generates the functions which are constant on dyadic intervals of the same generation. This is an immediate consequence of Lemma 2.3, because the functions which are constant on each $I_{n,k}, k = 0, \dots, 2^n - 1$, are precisely the linear combinations of $\{\chi, \psi_{n,k}\}_k$.

To see this, fix $n \geq 1$ and let $g(t) = \sum_{k=0}^{2^n-1} \alpha_k \chi_{I_{n,k}}(t)$ be any function that is constant on dyadic intervals of level n . For each interval $I_{n-1,j}$ of level $n-1$, the function g takes two constant values on the two level- n intervals contained in $I_{n-1,j}$, namely $I_{n,k}$ and $I_{n,k+1}$. Any such function, is a linear combination of the constant function $\chi_{I_{n-1,j}}$ and the Haar function $\psi_{n-1,j}$. Summing over all j shows that g lies in the span of χ and the Haar functions up to level $n-1$.

Thus, the Haar system generates all dyadic step functions, and therefore is complete in $L^2([0, 1])$. $\square$

2.2 The Haar wavelet

A central idea in wavelet theory is to describe functions through a sequence of approximation spaces that reflect different levels of resolution.

In our particular example, we begin with the space V_0 , consisting of the functions which are constant on each interval $[k, k+1)$:

$$V_0 = \left\{ f \in L^2(\mathbb{R}) \mid f|_{I_{0,k}} = c_k \text{ constant } \forall k \in \mathbb{Z} \right\}.$$

Observe that the function $\varphi(t) := \mathbf{1}_{[0,1)}(t)$, together with its translates $\varphi_{0,k}(t) = \varphi(t - k) = \mathbf{1}_{[k, k+1)}(t), k \in \mathbb{Z}$, form an orthonormal basis of V_0 . Consequently, a function f belongs to V_0 if and only if it can be written as

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

Moreover, for a general function $f \in L^2(\mathbb{R})$, its orthogonal projection onto V_0 is

$$P_0 f = \sum_{k \in \mathbb{Z}} \langle f, \varphi_{0,k} \rangle \varphi_{0,k} = \sum_{k \in \mathbb{Z}} c_k \chi_{I_{0,k}},$$

where

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

is the average value of f on the interval $I_{0,k}$.

Thus, V_0 represents f at a coarse level of resolution, since $P_0 f$ retains only the mean value of f on each unit interval, and therefore provides only a rough approximation of the original function.

To obtain finer resolutions, we consider dyadic dilations of the scaling function φ , which will lead us to the next approximation space V_1 .

Consider the intervals $I_{1,k} = [\frac{k}{2}, \frac{k+1}{2})$. Then, the V_1 space is given by:

$$V_1 = \left\{ f \in L^2(\mathbb{R}) \mid f|_{I_{1,k}} = c_k \text{ constant } \forall k \in \mathbb{Z} \right\}.$$

The functions $\varphi_{1,k}(t) = \sqrt{2} \varphi(2t - k) = \sqrt{2} \chi_{I_{1,k}}(t)$, $k \in \mathbb{Z}$ form an orthonormal basis of V_1 . Observe that this orthonormal basis is obtained by rescaling the basis of V_0 introduced above.

As before, a function f is in V_1 if and only if it can be represented as

$$f = \sum_{k \in \mathbb{Z}} c_k \varphi_{1,k} \quad \text{with} \quad \|f\|^2 = \sum_{k \in \mathbb{Z}} |c_k|^2 < \infty.$$

Furthermore, a general $f \in L^2(\mathbb{R})$ 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

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

Hence

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

is the function that assigns to each interval $I_{1,k}$ the average of f on that interval.

Therefore, the space V_1 provides a finer approximation, since its functions are constant on intervals of length $1/2$, thereby doubling the resolution.

Remark 2.4. The detail added in $[0, 1)$ to P_0f when increasing the resolution to obtain P_1f is a multiple of the function $\psi := \chi_{I_{1,1}} - \chi_{I_{1,0}}$. Indeed,

$$\begin{aligned} (P_1f - P_0f)\chi_{[0,1)} &= \left(\int_{I_{1,0}} f\right) \chi_{I_{1,0}} + \left(\int_{I_{1,1}} f\right) \chi_{I_{1,1}} - \left(\int_{I_{0,0}} f\right) \chi_{I_{0,0}} \\ &= \left(\int_{I_{1,0}} f - \int_{I_{0,0}} f\right) \chi_{I_{1,0}} + \left(\int_{I_{1,1}} f - \int_{I_{0,0}} f\right) \chi_{I_{1,1}}. \end{aligned}$$

and is a multiple of ψ since

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

The same argument applies to every interval $[k, k+1)$, where the detail added when going from V_0 to V_1 is a multiple of $\psi_{0,k}(t) := \psi(t-k)$.

Let W_0 be the orthogonal complement of V_0 in V_1 , denoted $W_0 = V_1 \ominus V_0$. This can be seen as the space of details that need to be added to P_0f in order to get P_1f for all $f \in L^2(\mathbb{R})$. Notice that $V_1 = V_0 \oplus W_0$ and $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of W_0 .

This procedure can be replicated at all scales. Consider the intervals $I_{n,k}$ together with the subset

$$V_n = \left\{ f \in L^2(\mathbb{R}) \mid f|_{I_{n,k}} = c_k \text{ constant } \forall k \in \mathbb{Z} \right\}.$$

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 general function $f \in L^2(\mathbb{R})$ is in V_n if is of the form

$$f = \sum_{k \in \mathbb{Z}} c_k \chi_{I_{n,k}} \quad \text{with} \quad \|f\|^2 = \sum_{k \in \mathbb{Z}} |c_k|^2 < \infty$$

Observe that the coefficients c_k depend on n . However, to simplify the notation, we will not write this dependence explicitly.

The orthogonal projection $P_n f$ is the best approximation of f within the space V_n . Thus, $P_n f$ is the unique element of V_n that minimizes the L^2 -distance to f , that is, the unique piecewise constant function on the intervals $I_{n,k}$ closest to f .

The detail added when going from resolution V_n to V_{n+1} forms a closed space of $L^2(\mathbb{R})$ denoted by W_n . Therefore, $V_{n+1} = V_n \oplus W_n$. The functions $\{\psi_{n,k}\}_{k \in \mathbb{Z}}$ form an orthonormal basis of W_n .

Given these definitions, it follows that

$$\overline{\bigcup_{n \in \mathbb{Z}} V_n} = L^2(\mathbb{R}) \quad \text{and} \quad \bigcap_{n \in \mathbb{Z}} V_n = \{0\}.$$

From the iteration on the V_n spaces, it also follows that

$$V_n = \bigoplus_{j=-\infty}^{n-1} W_j \quad \text{and} \quad L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

This last identity means that any function can be viewed as the superposition of the details at every resolution. Moreover, the system $\{\psi_{n,k}\}_{n,k}$ is an orthonormal basis of $L^2(\mathbb{R})$.

This is an example of a *multi-resolution analysis (MRA)*, specifically the Haar MRA. The function φ is the *scaling function*, and ψ is the (*mother*) *wavelet*. The family $\{\psi_{n,k}\}_{n,k}$ forms the wavelet basis of $L^2(\mathbb{R})$.

2.3 Multi-resolution analysis and Mallat's theorem

This section introduces multi-resolution analysis (MRA) and its role in constructing wavelet bases for $L^2(\mathbb{R})$. Mallat's theorem ensures that every MRA admits a wavelet and provides an explicit construction based on the scaling function.

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

Definition 2.5. A *multi-resolution analysis (MRA)* is an increasing sequence $\dots V_n \subset V_{n+1} \subset \dots$ of closed subspaces of $L^2(\mathbb{R})$ such that:

1. There exists a $\varphi \in V_0$ such that its translates $\varphi_{0,k}(t) = \varphi(t - k)$, $k \in \mathbb{Z}$ form an orthonormal basis of V_0 . The function φ is called the *scaling function* of the MRA.
2. $f(t) \in V_n$ if and only if $f(2t) \in V_{n+1}$. Equivalently, $V_{n+1} = D_1(V_n)$ for all $n \in \mathbb{Z}$. This means that $V_n = D_n(V_0)$. Furthermore,

$$\varphi_{n,k} = D_n(\varphi_{0,k}) = 2^{n/2} \varphi(2^n t - k), \quad k \in \mathbb{Z} \quad (2.1)$$

form an orthonormal basis of V_n .

3. $\overline{\bigcup_{n \in \mathbb{Z}} V_n} = L^2(\mathbb{R})$.
4. $\bigcap_{n \in \mathbb{Z}} V_n = \{0\}$.

Remark 2.6. 1. It follows from the definition that the scaling function determines the MRA.

2. It turns out that condition 4 is actually a consequence of the others. The proof is not included here. A complete argument can be found in [4, Section 2.1].

As in the Haar example, we can define W_n as the orthogonal complement of V_n in V_{n+1} , that is $W_n = V_{n+1} \ominus V_n$. Notice that $W_n = D_n(W_0)$. Likewise, one can see that

$$V_{n+1} = \bigoplus_{j \leq n} W_j \quad \text{and} \quad L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

Moreover, we can consider the orthogonal projections $P_n : L^2(\mathbb{R}) \rightarrow V_n$, given by

$$P_n f = \sum_{k \in \mathbb{Z}} \langle f, \varphi_{n,k} \rangle \varphi_{n,k}, \quad (2.2)$$

which represent f at resolution n . Similarly, we define $Q_n : L^2(\mathbb{R}) \rightarrow W_n$ as

$$Q_n f = P_{n+1} f - P_n f,$$

which is the detail added to $P_n f$ in order to obtain $P_{n+1} f$.

Following the same idea, a (*mother*) *wavelet of the MRA* $\{V_n\}_{n \in \mathbb{Z}}$ is a function $\psi \in W_0$ whose translates $\psi_{0,k}(t) = \psi(t - k)$, $k \in \mathbb{Z}$, form an orthonormal basis of W_0 . Since $L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$, a wavelet automatically generates an orthonormal basis of $L^2(\mathbb{R})$.

Proposition 2.7. Assume that ψ is a wavelet of the MRA $\{V_n\}_{n \in \mathbb{Z}}$ and $\psi_{n,k} = D_n(\psi_{0,k})$. Then $\{\psi_{n,k}\}_{n,k}$ is an orthonormal basis of $L^2(\mathbb{R})$.

The family $\{\psi_{n,k}\}_{n,k}$ is therefore called the *wavelet basis of the MRA*.

The natural question now is whether every MRA has a wavelet. The answer is yes, and the wavelet can be obtained from the scaling function. This is by no means a trivial result, which is the object of next theorem.

Mallat's theorem Let $\{V_n\}_{n \in \mathbb{Z}}$ be an MRA as above and let $\{\varphi_{n,k}\}_{n,k}$ be defined as in (2.1). Then the function

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

is a wavelet of the MRA.

In order to prove the theorem, the following lemma will be required.

Lemma 2.8. Using the same notation as in Mallat's theorem, the following equalities are satisfied:

1. $\sum_{k \in \mathbb{Z}} |\langle \varphi, \varphi_{1,k} \rangle|^2 = 1$
2. $\sum_{k \in \mathbb{Z}} \langle \varphi, \varphi_{1,k} \rangle \overline{\langle \varphi, \varphi_{1,k-2l} \rangle} = 0 \quad \forall l \in \mathbb{Z} \setminus \{0\}.$

Proof. 1. Since $\varphi \in V_0$ and $V_0 \subset V_1$, it follows that $\varphi \in V_1$. Because $\{\varphi_{1,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of the Hilbert space V_1 , the function φ admits the expansion

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

By Parseval's identity 1.5 in V_1 , and since φ is normalized,

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

2. Let $l \neq 0$. The goal is to show that

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

Since $\{\varphi_{0,l}\}_{l \in \mathbb{Z}}$ is an orthonormal basis of V_0 and $l \neq 0$, one has $\langle \varphi, \varphi_{0,l} \rangle = 0$.

The inner product $\langle \varphi, \varphi_{0,l} \rangle$ is obtained by expanding $\varphi_{0,l}$ in the orthonormal basis $\{\varphi_{1,k}\}_{k \in \mathbb{Z}}$ of V_1 . 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

$$\varphi_{1,k}(t-l) = \sqrt{2}\varphi(2(t-l)-k) = \sqrt{2}\varphi(2t-(k+2l)) = \varphi_{1,k+2l}, \quad (2.3)$$

one obtains

$$\varphi_{0,l}(t) = \sum_{k \in \mathbb{Z}} \langle \varphi, \varphi_{1,k} \rangle \varphi_{1,k+2l}(t) = \sum_{k \in \mathbb{Z}} \langle \varphi, \varphi_{1,k-2l} \rangle \varphi_{1,k}. \quad (2.4)$$

The second equality follows from the change of variable $m = k + 2l$ in the summation, after which the index m is renamed k .

Then, using the expression for φ , the claim follows. $\square$

Proof of Mallat's theorem. To show that the function ψ defined in the statement is a wavelet of the MRA, it is necessary to verify that $\psi \in V_1 \ominus V_0$ and that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ forms an orthonormal basis of W_0 .

By definition, is clear that $\psi \in V_1$. To prove that $\psi \in V_0^\perp$, it is enough to see that for all $l \in \mathbb{Z}$ $\langle \varphi_{0,-l}, \psi \rangle = \langle \varphi, \psi_{0,l} \rangle = 0$, since $\{\varphi_{0,l}\}_{l \in \mathbb{Z}}$ is an orthonormal basis of V_0 .

Using (2.3), denoting $c_k := \langle \varphi, \varphi_{1,k} \rangle$ and re-indexing the sum that defines ψ , one obtains

$$\begin{aligned} \psi_{0,l}(t) &= \psi(t-l) = \sum_{k \in \mathbb{Z}} (-1)^k \overline{c_{1-k}} \varphi_{1,k}(t-l) = \sum_{k \in \mathbb{Z}} (-1)^k \overline{c_{1-k}} \varphi_{1,k+2l}(t) \\ &= \sum_{m \in \mathbb{Z}} (-1)^m \overline{c_{1-m+2l}} \varphi_{1,m}(t). \end{aligned} \quad (2.5)$$

The last equality follows from the change of index $m = k + 2l$ in the summation. One also has

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

Consequently,

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

Since $(-1)^m$ and $(-1)^{1-m+2l}$ have opposite signs, each term appears twice with opposite signs, and therefore the last summation vanishes.

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

It remains to see that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of W_0 . To check orthogonality, notice that, since $\langle \psi_{0,m}, \psi_{0,j} \rangle = \langle \psi, \psi_{0,j-m} \rangle$, it is enough to prove that $\langle \psi, \psi_{0,l} \rangle = \delta_{0,l}$ for every $l \in \mathbb{Z}$.

Using (2.5), the definition of ψ and the identities provided by Lemma 2.8, one obtains

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

To complete the proof, one has to show that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ generates the whole W_0 , which forms the hardest part of the argument. This reduces to proving 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 $f \in V_1$, it is completely determined by its coordinates in the basis $\{\varphi_{1,k}\}_{k \in \mathbb{Z}}$. Consider thus the sequence $F := (f_k)_{k \in \mathbb{Z}}$ with $f_k := \langle f, \varphi_{1,k} \rangle$.

In the same way, using (2.4), the functions $\{\psi_{0,l}\}_{l \in \mathbb{Z}}$ can be identified with the sequences of coefficients $\Phi_l := (c_{k-2l})_{k \in \mathbb{Z}}$.

Similarly, one can also identify $\{\psi_{0,l}\}_{l \in \mathbb{Z}}$ with the sequence $\Psi_l := ((-1)^k \overline{c_{1-k+2l}})_{k \in \mathbb{Z}}$.

Then, the orthogonality assumptions can be written equivalently as

$$\langle F, \Phi_l \rangle_{l^2(\mathbb{Z})} = 0, \text{ for all } l \in \mathbb{Z}, \text{ which is } f \perp V_0. \quad (2.6)$$

and

$$\langle F, \Psi_l \rangle_{l^2(\mathbb{Z})} = 0, \text{ for all } l \in \mathbb{Z}, \text{ which is } \langle f, \psi_{0,l} \rangle = 0. \quad (2.7)$$

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

$$\begin{array}{ccccccccc}
\cdots & \Phi_1 & & \Psi_1 & & \Phi_0 & & \Psi_0 & & \Phi_{-1} & & \Psi_{-1} & \cdots \\
& \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
& \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\
& c_{-1-2} & & (-1)^{-1}\overline{c_{1+1+2}} & & c_{-1} & & (-1)^{-1}\overline{c_{1+1}} & & c_{-1+2} & & (-1)^{-1}\overline{c_{1+1-2}} & \\
& c_{0-2} & & (-1)^0\overline{c_{1-0+2}} & & c_0 & & (-1)^0\overline{c_{1-0}} & & c_{0+2} & & (-1)^0\overline{c_{1-0-2}} & \\
\cdots & c_{1-2} & & (-1)^1\overline{c_{1-1+2}} & & c_1 & & (-1)^1\overline{c_{1-1}} & & c_{1+2} & & (-1)^1\overline{c_{1-1-2}} & \cdots \\
& \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

Since the columns of the matrix are sequences in $l^2(\mathbb{Z})$, their inner products can be expressed in terms of the corresponding functions in $L^2(\mathbb{R})$. Using that $\{\varphi_{1,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of V_1 and $\varphi_{0,l}, \psi_{0,l} \in V_1$, one obtains

$$\begin{aligned}
\langle \Phi_l, \Phi_m \rangle_{\ell^2} &= \langle \varphi_{0,l}, \varphi_{0,m} \rangle_{L^2}, & \langle \Psi_l, \Psi_m \rangle_{\ell^2} &= \langle \psi_{0,l}, \psi_{0,m} \rangle_{L^2}, \\
\langle \Phi_l, \Psi_m \rangle_{\ell^2} &= \langle \varphi_{0,l}, \psi_{0,m} \rangle_{L^2}.
\end{aligned}$$

Moreover, by the orthogonality of $\{\varphi_{0,l}\}_{l \in \mathbb{Z}}$ and $\{\psi_{0,l}\}_{l \in \mathbb{Z}}$, one has

$$\langle \varphi_{0,l}, \varphi_{0,m} \rangle = \delta_{lm}, \quad \langle \psi_{0,l}, \psi_{0,m} \rangle = \delta_{lm}, \quad \langle \varphi_{0,l}, \psi_{0,m} \rangle = 0.$$

The identities (2.6) and (2.7) above imply the orthogonality of the columns of M .

Claim 2.9. $MM^* = I$, where M^* denotes the conjugate transpose of M .

Assuming the preceding claim has been verified, relations (2.6) and (2.7) yield $M^*F = 0$. Consequently, $F = MM^*F = 0$, which completes the proof.

The claim itself follows from an explicit computation. To this end, decompose M into 2×2 blocks of the form

$$M_{m,l} = \begin{pmatrix} c_{m-2l} & (-1)^m \overline{c_{1-m+2l}} \\ c_{m-2l+1} & -(-1)^m \overline{c_{-m+2l}} \end{pmatrix}, \quad m, l \in \mathbb{Z},$$

where the first column correspond to the entries of Φ_l and the second to those of Ψ_l . Then, considering that $(M^*)_{j,l} = (M_{l,j})^*$ the matrix MM^* is constructed from the blocks

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

Using the definition of c_k and the orthogonality relations (2.6), (2.7), one sees that this is

$$\begin{pmatrix} \delta_{m,l} & 0 \\ 0 & \delta_{m,l} \end{pmatrix}.$$

This shows that the family $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ forms an orthonormal basis of the space W_0 . By applying the same argument at an arbitrary scale n , one obtains that $\{\psi_{n,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of W_0 . Since $\bigcap_{n \in \mathbb{Z}} V_n = \{0\}$ and $\overline{\bigcup_{n \in \mathbb{Z}} V_n} = L^2(\mathbb{R})$, it follows that $L^2(\mathbb{R}) = \overline{\bigoplus_{n \in \mathbb{Z}} W_n}$. Consequently, the collection $\{\psi_{n,k}\}_{n,k \in \mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{R})$. □

Example 2.10. We now obtain the Haar wavelet from the scaling function, as given by Mallat's theorem. Let $\varphi = \chi_{[0,1]}$ be the scaling function. Then

$$\varphi = \chi_{[0,1/2)} + \chi_{[1/2,1)} = \frac{1}{\sqrt{2}}\varphi_{1,0} + \frac{1}{\sqrt{2}}\varphi_{1,1}$$

so that

$$c_k = \langle \varphi, \varphi_{1,k} \rangle = \begin{cases} 1/\sqrt{2} & \text{if } k = 0, 1 \\ 0 & \text{if } k \neq 0, 1. \end{cases}$$

Then, by Mallat's theorem

$$\begin{aligned} \psi &= (-1)^0 \overline{c_1} \varphi_{1,0} + (-1)^1 \overline{c_0} \varphi_{1,1} \\ &= \frac{1}{\sqrt{2}} \sqrt{2} \chi_{[0,1/2)} - \frac{1}{\sqrt{2}} \sqrt{2} \chi_{[1/2,1)} = \chi_{[0,1/2)} - \chi_{[1/2,1)}. \end{aligned}$$

Remark 2.11. The wavelet function ψ is uniquely determined up to a sign factor ± 1 . The choice of sign convention does not alter the definition of W_0 .

2.4 Construction of a Multi-resolution analysis

The idea behind the theorem is that a MRA separates a signal into what is already visible at a coarse level and what becomes visible only when the resolution is refined. Mallat's result formalizes this idea and provides a method to construct a wavelet basis from the scaling function:

1. Take a scaling function $\varphi \in L^2(\mathbb{R})$ such that $\{\varphi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal system. Define $V_0 = \overline{\langle \varphi_{0,k} \rangle_{k \in \mathbb{Z}}}$.

2. Prove that $V_n := \overline{D_n(V_0)}$ is an increasing sequence of closed subspaces of $L^2(\mathbb{R})$ and that $\bigcup_{n \in \mathbb{Z}} V_n = L^2(\mathbb{R})$.
3. Using Mallat's theorem, find the associated wavelet such that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of W_0 , where $W_0 = V_1 \ominus V_0$.

Fourier analysis can be used to formulate conditions ensuring that the requirements above are satisfied. The discussion begins by expressing item 1 in terms of the Fourier transform, leading to an equivalent formulation.

Theorem 2.12. Let $\varphi \in L^2(\mathbb{R})$. Then $\{\varphi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal system if and only if

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

Proof. Using Plancherel's formula 1.22 and Fubini's theorem, since $\widehat{\varphi_{0,k}}(\xi) = e^{2\pi i k \xi} \widehat{\varphi}(\xi)$, which is in 1.10, one has

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

Let $F(\xi) = \sum_{n \in \mathbb{Z}} |\widehat{\varphi}(\xi + n)|^2$ be the periodization of $|\widehat{\varphi}|^2$. By its definition, one sees that F is 1-periodic and in $L^1([0, 1])$, since $\widehat{\varphi} \in L^2(\mathbb{R})$. Hence, by the last formula, $\langle \varphi, \varphi_{0,k} \rangle = \widehat{F}(k)$.

Since the system $\{\varphi_{0,k}\}_{k \in \mathbb{Z}}$ is orthonormal, $\widehat{F}(k) = 0$ for all $k \neq 0$ and $\widehat{F}(0) = 1$. By the uniqueness theorem for L^1 functions 1.21, one concludes that $F(\xi) = 1$. On the other hand, if $F(\xi) = 1$ a.e. $\xi \in \mathbb{R}$, then $\widehat{F}(k) = 0$ for all $k \neq 0$ and $\widehat{F}(0) = 1$. $\square$

An equivalent reformulation of item 2, expressed in terms of the Fourier transform of the scaling function, is given in the next theorem.

Theorem 2.13. Let $\{V_n\}_{n \in \mathbb{Z}}$ be an increasing sequence of closed subspaces of $L^2(\mathbb{R})$ such that $D_1(V_n) = V_{n+1}$, $\forall n \in \mathbb{Z}$. Let φ be such that $\{\varphi_{0,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of V_0 and $|\widehat{\varphi}|$ is continuous at 0. Then $\bigcup_{n \in \mathbb{Z}} V_n = L^2(\mathbb{R})$ if and only if $\widehat{\varphi}(0) \neq 0$. In that case, $|\widehat{\varphi}(0)| = 1$.

Once the scaling function φ has been obtained, the associated wavelet ψ and the corresponding wavelet basis $\{\psi_{n,k}\}_{n,k \in \mathbb{Z}}$ can be constructed using Mallat's theorem.

Proof. Assume first that $\widehat{\varphi}(0) \neq 0$ holds. We want to prove that

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

To show that $\mathcal{V}^\perp = \{0\}$ we will prove first that $\mathcal{V}$ is invariant under translations. Let $f \in \mathcal{V}$ and consider a dyadic translation of f by $m/2^l$, $m, l \in \mathbb{Z}$. That is,

$$\tau_{m/2^l} f(x) = f(x - \frac{m}{2^l}), \quad m, l \in \mathbb{Z}.$$

By the definition of $\mathcal{V}$, given $\varepsilon > 0$ there exists $n_0 \in \mathbb{Z}$ and $h \in V_{n_0}$ such that $\|f - h\|_2 < \varepsilon$. Since $\{V_n\}_{n \in \mathbb{Z}}$ is an increasing sequence, $h \in V_n$ for all $n \geq n_0$ and therefore it has an expression of the form

$$h(t) = \sum_{k \in \mathbb{Z}} a_{n,k} \varphi(2^n t - k).$$

Then

$$\begin{aligned} \tau_{m/2^l} h(t) &= h(t - \frac{m}{2^l}) = \sum_{k \in \mathbb{Z}} a_{n,k} \varphi(2^n (t - \frac{m}{2^l}) - k) \\ &= \sum_{k \in \mathbb{Z}} a_{n,k} \varphi(2^n t - (2^{n-l} m + k)). \end{aligned}$$

Because $2^{n-l} m$ is an integer for every $n \geq l$, the function $\tau_{m/2^l} h(t)$ belongs to V_n as well. Then $\tau_{m/2^l} f \in \mathcal{V}$, since $\|\tau_{m/2^l} f - \tau_{m/2^l} h\|_2 = \|f - h\|_2 < \varepsilon$.

For a general translation $\tau_x f$, $x \in \mathbb{R}$, take a dyadic number $m/2^l$ close enough to x so that $\|\tau_{m/2^l} f - \tau_x f\|_2 < \varepsilon$. This is possible because translations are continuous in $L^2(\mathbb{R})$.

It remains to verify $\mathcal{V}^\perp = \{0\}$. Let $g \in \mathcal{V}^\perp$. By the invariance of translations

$$\langle \tau_{-x} f, g \rangle = \int_{\mathbb{R}} f(t+x) \overline{g(t)} dt = 0 \quad \text{for all } x \in \mathbb{R} \text{ and for all } f \in \mathcal{V}.$$

By Plancherel's formula 1.22, this is

$$\int_{\mathbb{R}} e^{2\pi i x \xi} \widehat{f}(\xi) \overline{\widehat{g}(\xi)} d\xi = 0 \quad \text{for all } x \in \mathbb{R} \text{ and for all } f \in \mathcal{V}.$$

Using the Cauchy-Schwartz inequality 1.2, one has that $\widehat{f} \widehat{g} \in L^1(\mathbb{R})$. Therefore, the inversion formula yields $\widehat{f}(\xi) \widehat{g}(\xi) = 0$ a.e. $\xi \in \mathbb{R}$.

In particular, for $f(t) = 2^n \varphi(2^n t)$, one has $f \in V_n \subset \mathcal{V}$ and $\widehat{f}(\xi) = \widehat{\varphi}(\xi/2^n)$, by the properties of the Fourier integral. Hence

$$\widehat{\varphi}(\xi/2^n) \widehat{g}(\xi) = 0 \quad \text{a.e. } \xi \in \mathbb{R}.$$

Since $\widehat{\varphi}(0) \neq 0$ and $|\widehat{\varphi}|$ is continuous at 0, there exists an interval $I = (-\eta, \eta)$ where $\widehat{\varphi}(\xi) \neq 0$ for all $\xi \in I$. For any fixed $\xi \in \mathbb{R}$, choose n large enough so that $\xi/2^n \in I$, which implies $\widehat{g}(\xi) = 0$. Since this holds for all $\xi \in \mathbb{R}$, it follows that $\widehat{g} \equiv 0$, and hence $g = 0$, as desired.

Conversely, suppose that $\bigcup_{n \in \mathbb{Z}} V_n = L^2(\mathbb{R})$. One must prove that $\widehat{\varphi}(0) \neq 0$.

Let $f(t) = \text{sinc}(2t) = \sin(2\pi t)/(2\pi t)$, that is $\widehat{f} = \chi_{[-1,1]}$, and let $P_n : L^2(\mathbb{R}) \rightarrow V_n$ be the orthogonal projections. Then

$$||f||_2 - ||P_n f||_2 \leq ||f - P_n f||_2 \xrightarrow{n \rightarrow \infty} 0.$$

Since $\{\varphi_{n,k}\}_{k \in \mathbb{Z}}$ is an orthonormal basis of V_n , using Parseval's identity and Plancherel's formula, one has

$$||P_n f||_2^2 = \sum_{k \in \mathbb{Z}} |\langle f, \varphi_{n,k} \rangle|^2 \xrightarrow{n \rightarrow \infty} ||f||_2^2 = ||\widehat{f}||_2^2 = 2. \quad (2.8)$$

Since

$$\varphi_{n,k}(t) = 2^{n/2} \varphi(2^n t - k) = 2^{n/2} \varphi(2^n(t - k/2^n)) = 2^{-n/2} \tau_{k/2^n} D_{1/2^n} \varphi(t),$$

we deduce that

$$\widehat{\varphi}_{n,k}(\xi) = 2^{-n/2} e^{2\pi i(k/2^n)\xi} \widehat{\varphi}(\xi/2^n).$$

Then, Plancherel and the substitution $2^{-n}\xi = \omega$ yield

$$\begin{aligned} ||P_n f||_2^2 &= \sum_{k \in \mathbb{Z}} |\langle \widehat{f}, \widehat{\varphi}_{n,k} \rangle|^2 = \sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i(k/2^n)\xi} 2^{-n/2} \overline{\widehat{\varphi}(2^{-n}\xi)} d\xi \right|^2 \\ &= \sum_{k \in \mathbb{Z}} \left| \int_{-1}^1 e^{-2\pi i(k/2^n)\xi} 2^{-n/2} \overline{\widehat{\varphi}(2^{-n}\xi)} d\xi \right|^2 \\ &= \sum_{k \in \mathbb{Z}} \left| \int_{-2^{-n}}^{2^{-n}} e^{-2\pi i k \omega} 2^{n/2} \overline{\widehat{\varphi}(\omega)} 2^n d\omega \right|^2 = 2^n \sum_{k \in \mathbb{Z}} \left| \int_{-2^{-n}}^{2^{-n}} e^{-2\pi i k \omega} \overline{\widehat{\varphi}(\omega)} d\omega \right|^2. \end{aligned}$$

For $n \geq 0$ the integral appearing in the last sum is the k^{th} Fourier coefficient of the function $\chi_{[-2^{-n}, 2^{-n}]} \overline{\widehat{\varphi}}$, which is in $L^2([-1, 1])$. Hence, by Plancherel

$$||P_n f||_2^2 = 2^{n+1} ||\chi_{[-2^{-n}, 2^{-n}]} \overline{\widehat{\varphi}}||_{L^2([-1, 1])}^2 = 2^n \int_{-2^{-n}}^{2^{-n}} |\widehat{\varphi}(\omega)|^2 d\omega = 2 \int_{-2^{-n}}^{2^{-n}} |\widehat{\varphi}(\omega)|^2 d\omega.$$

By the continuity of $|\widehat{\varphi}|$ at 0, one has

$$\int_{-2^{-n}}^{2^{-n}} |\widehat{\varphi}(\omega)|^2 d\omega \xrightarrow{n \rightarrow \infty} |\widehat{\varphi}(0)|^2.$$

This together with (2.8) shows that $\widehat{\varphi}(0) \neq 0$ and $|\widehat{\varphi}(0)| = 1$. □

Example 2.14. To exemplify the theorems above, consider the Haar system. Let $\varphi = \chi_{[0,1]}$. Its Fourier transform is given by

$$\begin{aligned} \widehat{\varphi}(\xi) &= \int_{\mathbb{R}} \varphi(t) e^{-2\pi i t \xi} dt = \int_0^1 e^{-2\pi i t \xi} dt = \frac{1 - e^{-2\pi i \xi}}{2\pi i \xi} \\ &= e^{-\pi i \xi} \frac{\sin(\pi \xi)}{\pi \xi} = e^{-\pi i \xi} \operatorname{sinc}(\xi). \end{aligned}$$

Consequently, a.e $\xi \in \mathbb{R}$ one has

$$\sum_{n \in \mathbb{Z}} |\widehat{\varphi}(\xi + n)|^2 = \sum_{n \in \mathbb{Z}} \left| e^{-\pi i (\xi + n)} \operatorname{sinc}(\xi + n) \right|^2 = \sum_{n \in \mathbb{Z}} \operatorname{sinc}(\xi + n)^2 = 1.$$

Moreover, the function $|\widehat{\varphi}(\xi)| = |\operatorname{sinc}(\xi)|$ is continuous at 0, with $|\widehat{\varphi}(0)| = 1$.

Equivalently, the result is obtained from the fact that $\widehat{\varphi} = \chi_{[-1/2, 1/2]}$.

2.5 The wavelet from the Fourier side

In the previous section, we have given conditions so that φ is the scaling function of an MRA in terms of its Fourier transform $\widehat{\varphi}$. Fourier analysis also provides a way to construct the associated wavelet ψ . The procedure can be outlined as follows.

Since $\varphi \in V_0$ and $V_0 \subset V_1$ in any MRA, the scaling function admits an expansion in the orthonormal basis $\{\varphi_{1,k}\}_{k \in \mathbb{Z}}$ of V_1 :

$$\varphi(t) = \sum_{k \in \mathbb{Z}} c_k \varphi_{1,k}(t), \quad c_k = \langle \varphi, \varphi_{1,k} \rangle, \quad \varphi_{1,k}(t) = \sqrt{2} \varphi(2t - k).$$

Taking Fourier transforms on each side and using the identity 1.10

$$\widehat{\varphi(2t - k)}(\xi) = \frac{1}{2} e^{-2\pi i k \xi / 2} \widehat{\varphi}(\xi / 2),$$

one obtains

$$\widehat{\varphi}(\xi) = \left(\frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} c_k e^{-2\pi i k \xi / 2} \right) \widehat{\varphi}(\xi / 2).$$

Define now the function

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

Since H is a Fourier series, it is 1-periodic. Therefore, it satisfies $H(\xi + 1) = H(\xi)$. With this notation, the previous identity becomes

$$\widehat{\varphi}(\xi) = H(\xi/2) \widehat{\varphi}(\xi/2). \quad (2.9)$$

Remark 2.15. Following the same idea, one can see that f belongs to V_1 if and only if $\widehat{f}(\xi) = m(\xi/2) \widehat{\varphi}(\xi/2)$, with m a 1-periodic function.

The function H is called the *refinement mask* or *low pass filter* of the MRA and satisfies the *quadratic mirror filter (QMF)* identity

$$|H(\xi)|^2 + |H(\xi + 1/2)|^2 = 1, \quad \xi \in \mathbb{R}. \quad (2.10)$$

Let us see this. By Theorem 2.12, one has

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

Substituting the identity (2.9) into this expression yields

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

Now, split this sum over all integers $n \in \mathbb{Z}$ into even and odd integers:

$$\sum_{m \in \mathbb{Z}} \left| H\left(\frac{\xi}{2} + m\right) \right|^2 \left| \widehat{\varphi}\left(\frac{\xi}{2} + m\right) \right|^2 + \sum_{m \in \mathbb{Z}} \left| H\left(\frac{\xi}{2} + m + \frac{1}{2}\right) \right|^2 \left| \widehat{\varphi}\left(\frac{\xi}{2} + m + \frac{1}{2}\right) \right|^2 = 1.$$

Since H is 1-periodic, $|H(\eta + m)|^2 = |H(\eta)|^2$ for every $m \in \mathbb{Z}$. Thus, one can pull the filter terms out of their respective summations:

$$\left| H\left(\frac{\xi}{2}\right) \right|^2 \sum_{m \in \mathbb{Z}} \left| \widehat{\varphi}\left(\frac{\xi}{2} + m\right) \right|^2 + \left| H\left(\frac{\xi}{2} + \frac{1}{2}\right) \right|^2 \sum_{m \in \mathbb{Z}} \left| \widehat{\varphi}\left(\frac{\xi}{2} + \frac{1}{2} + m\right) \right|^2 = 1.$$

Applying Theorem 2.12 once again to the remaining sums, observe that

$$\sum_{m \in \mathbb{Z}} \left| \widehat{\varphi}\left(\frac{\xi}{2} + m\right) \right|^2 = 1 \quad \text{and} \quad \sum_{m \in \mathbb{Z}} \left| \widehat{\varphi}\left(\left(\frac{\xi}{2} + \frac{1}{2}\right) + m\right) \right|^2 = 1.$$

Therefore, the equation simplifies directly to

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

Changing variables by setting $\eta = \xi/2$ and renaming the variable η as ξ yields the QMF identity (2.10).

With this, one gets a description of the functions in W_0 .

Lemma 2.16. A function $f \in L^2(\mathbb{R})$ belongs to the detail space W_0 if and only if there exists a 1-periodic function v such that

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

Proof. Since $W_0 = V_1 \ominus V_0$, a function f belongs to W_0 if and only if

$$f \in V_1 \quad \text{and} \quad \langle f, \varphi(\cdot - k) \rangle = 0 \quad \text{for all } k \in \mathbb{Z}.$$

As we have remarked above, $f \in V_1$ if and only if $\widehat{f}(\xi) = m_f(\xi/2) \widehat{\varphi}(\xi/2)$ for some 1-periodic function m_f . Consequently, one can reduce the problem to showing that

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

where $\sigma(\xi)$ is a $(1/2)$ -periodic function. Then $v(\xi) = \sigma(\xi/2)$ will have period one.

The condition $\langle f, \varphi(\cdot - k) \rangle = 0$ for all k shows that

$$0 = \langle \widehat{f}, \widehat{\varphi_{0,k}} \rangle = \int_{\mathbb{R}} \widehat{f}(\xi) \overline{\widehat{\varphi}(\xi)} e^{2\pi i k \xi} d\xi, \quad k \in \mathbb{Z},$$

by Plancherel. Substituting the expression for $\widehat{f}$ and using the identity (2.9) one has

$$0 = \int_{\mathbb{R}} m_f(\xi/2) \widehat{\varphi}(\xi/2) \overline{\widehat{\varphi}(\xi)} e^{2\pi i k \xi} d\xi = \int_{\mathbb{R}} m_f(\xi/2) |\widehat{\varphi}(\xi/2)|^2 \overline{H(\xi/2)} e^{2\pi i k \xi} d\xi.$$

One can decompose the integral as a sum of integrals over the intervals $[n, n+1)$. After this decomposition, a change of variables and the periodicity of the exponential functions yields

$$\begin{aligned} 0 &= \sum_{n \in \mathbb{Z}} \int_n^{n+1} m_f(\xi/2) |\widehat{\varphi}(\xi/2)|^2 \overline{H(\xi/2)} e^{2\pi i k \xi} d\xi \\ &= \int_0^1 e^{2\pi i k \xi} \sum_{n \in \mathbb{Z}} m_f((\xi+n)/2) |\widehat{\varphi}((\xi+n)/2)|^2 \overline{H((\xi+n)/2)} d\xi. \end{aligned}$$

Now we separate the last sum into the sums over the odd and even integers, so for all $k \in \mathbb{Z}$ we obtain

$$\begin{aligned} 0 &= \int_0^1 e^{2\pi i k \xi} m_f(\xi/2) \overline{H(\xi/2)} \sum_{m \in \mathbb{Z}} |\widehat{\varphi}(\xi/2 + m)|^2 \\ &\quad + m_f(\xi/2 + 1/2) \overline{H(\xi/2 + 1/2)} \sum_{m \in \mathbb{Z}} |\widehat{\varphi}((\xi + 1)/2 + m)|^2 d\xi \\ &= \int_0^1 e^{2\pi i k \xi} [m_f(\xi/2) \overline{H(\xi/2)} + m_f(\xi/2 + 1/2) \overline{H(\xi/2 + 1/2)}] d\xi, \end{aligned}$$

where theorem 2.12 has been used. As a consequence, the function

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

has the Fourier coefficients equal to zero. Thus, $F(\xi) = 0$ a.e.

This equation states that the vector $\vec{v} \in \mathbb{C}^2$ given by $\vec{v} = (m_f(\xi), m_f(\xi + 1/2))$ must be orthogonal to the vector $\vec{w} \in \mathbb{C}^2$ given by $\vec{w} = (H(\xi), H(\xi + 1/2))$.

The identity (2.10) ensures that $\vec{w}$ is not zero, in fact $|\vec{w}| = 1$. Hence, one needs to find a non-zero vector $\vec{u} \in \mathbb{C}^2$ orthogonal to $\vec{w}$ to completely characterize all the vectors $\vec{v}$ that are orthogonal to $\vec{w}$. In fact, $\vec{v} = \lambda \vec{u}$ for $\lambda \in \mathbb{C}$. The vector

$$\vec{u} = (-\overline{H(\xi + 1/2)}, \overline{H(\xi)})$$

is orthogonal to $\vec{w}$. Therefore, $\vec{v} = \lambda(-\overline{H(\xi + 1/2)}, \overline{H(\xi)})$ so one concludes that m_f has to satisfy

$$m_f(\xi) = -\lambda(\xi) \overline{H(\xi + 1/2)}, \quad m_f(\xi + 1/2) = \lambda(\xi) \overline{H(\xi)}.$$

Therefore, $\lambda(\xi)$ is a 1-periodic function such that $-\lambda(\xi + 1/2) = \lambda(\xi)$. Equivalently, $\lambda(\xi) = e^{2\pi i \xi} \sigma(\xi)$, where $\sigma(\xi)$ is a $(1/2)$ -periodic function. This concludes the proof, since then

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

as required. □

Since ψ is in W_0 , it must have the form given by 2.16. One can use the notation

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

where

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

and σ is $(1/2)$ -periodic.

Since the construction requires $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ to be an orthonormal system, using again Theorem 2.12, one can see that m_ψ also satisfies the QMF identity (2.10):

$$\begin{aligned} |m_\psi(\xi)|^2 + |m_\psi\left(\xi + \frac{1}{2}\right)|^2 &= |\sigma(\xi)|^2 |H\left(\xi + \frac{1}{2}\right)|^2 + |\sigma\left(\xi + \frac{1}{2}\right)|^2 |H(\xi + 1)|^2 \\ &= |\sigma(\xi)|^2 \left(|H\left(\xi + \frac{1}{2}\right)|^2 + |H(\xi)|^2 \right) = |\sigma(\xi)|^2 = 1. \end{aligned}$$

Mallat's approach consist on taking the simplest $\sigma(\xi)$ with $|\sigma(\xi)| = 1$, that is $\sigma(\xi) \equiv 1$. With this choice, one has

$$\hat{\psi}(\xi) = G(\xi/2) \hat{\varphi}(\xi/2),$$

with

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

which is 1-periodic, defined as the *high pass filter*. This formula determines the wavelet ψ entirely in terms of φ .

It remains to verify that the function ψ indeed belongs to W_0 and that the family $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ form an orthonormal basis of W_0 .

Lemma 2.16 implies directly that ψ is in W_0 .

Furthermore, since G satisfies the QMF identity (2.10), Theorem 2.12 ensures that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ is orthonormal.

To see that $\{\psi_{0,k}\}_{k \in \mathbb{Z}}$ spans W_0 , consider $f \in W_0$. By Lemma 2.16 and the identities above, its Fourier transform can be expressed as

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

for some 1-periodic function v . In this case, v has the form $v(\xi) = \sum_{k \in \mathbb{Z}} c_k e^{2\pi i k \xi}$, with $\{c_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$. Consequently,

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

Hence, $f = \sum_{k \in \mathbb{Z}} c_k \psi_{0,k}$, and therefore $W_0 = \overline{\langle \psi_{0,k} \rangle_{k \in \mathbb{Z}}}$.

Mallat's theorem provides an algorithm for constructing the wavelet from the MRA and the scaling function via the Fourier coefficients of the 1-periodic refinement mask. This can be implemented numerically through the *cascade algorithm*, which schematically goes as follows:

$$\begin{array}{ccccccc} \varphi & \longrightarrow & H & \longrightarrow & G & \longrightarrow & \psi \\ \text{scaling function} & \longrightarrow & \text{low-pass filter} & \longrightarrow & \text{high-pass filter} & \longrightarrow & \text{wavelet} \end{array}$$

Chapter 3

Wavelets in $\mathbb{R}^2$

The aim of this chapter is to extend the notion of wavelets to functions in $\mathbb{R}^2$, motivated by applications in image processing. As noted in the previous chapter, the edges and contours that characterize images are naturally captured by wavelets.

Images vary along two independent directions, and meaningful features arise from the interaction between these directions. In order to analyze such features, it is important to scale consistently across both axes. This preserves the meaning of scale and ensures that the wavelet transform behaves in a way that aligns with how images are interpreted and processed.

The first section introduces separable multi-resolutions, obtained as products of one-dimensional multi-resolutions. This construction provides a natural and computationally efficient extension of wavelet methods to the two-dimensional domain. It also preserves the structure that underlies many algorithms used in image analysis.

The second section turns to the concept of the continuous wavelet transform, where there is no need of an MRA. The focus lies on the required conditions on the wavelet ψ .

3.1 Separable multi-resolutions

Let $\{\psi_{n,k}\}_{n,k \in \mathbb{Z}}$ be an orthonormal wavelet basis of $L^2(\mathbb{R})$. The objective is to construct a separable system in $L^2(\mathbb{R}^2)$ by considering the products of functions of the form

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

with the extra condition that the same resolution level is used in both coordinate directions. If not, one could have different scales along the t_1 and t_2 axes. This

would break the idea that each scale should represent a uniform level of detail. It complicates interpretation and makes the implementation less natural.

Therefore, we will be restricted to products for which $n_1 = n_2$. This assumption ensures that all elements correspond to the same scale at all directions.

Definition 3.1. A *multi-resolution analysis (MRA)* in $\mathbb{R}^2$ is an increasing sequence $\dots V_n^{(2)} \subset V_{n+1}^{(2)} \subset \dots$ of closed subspaces of $L^2(\mathbb{R}^2)$ satisfying the properties of an MRA as stated in Definition 2.5, with all objects now defined on $\mathbb{R}^2$.

As in the one-dimensional approach, the notion of resolution is expressed through the orthogonal projections onto the spaces $V_n^{(2)}$. The approximation of a function $f(t_1, t_2) \in L^2(\mathbb{R}^2)$ at resolution n is given by its orthogonal projection $P_n f$ onto $V_n^{(2)}$. Moreover, the space $V_n^{(2)}$ is the range of P_n , and therefore consists of all functions of the form $P_n f$ with $f \in L^2(\mathbb{R}^2)$.

In the remainder of this section, only separable MRA's will be considered. This class is sufficient for most practical purposes, since separable constructions lead to simpler and more efficient computations.

Given a multi-resolution $\{V_n\}_{n \in \mathbb{Z}}$ in $L^2(\mathbb{R})$, its associated *separable two-dimensional MRA* is $\{V_n^{(2)}\}_{n \in \mathbb{Z}}$, where $V_n^{(2)} = V_n \otimes V_n$. Hence, $F \in V_n^{(2)}$ is of 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 $\sum_{m \in \mathbb{Z}} |c_m|^2 = \|F\|_{L^2(\mathbb{R}^2)}^2 < \infty$.

One verifies that $\{V_n^{(2)}\}_{n \in \mathbb{Z}}$ is an MRA 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

$$\varphi_{n,k}^{(2)}(t_1, t_2) = \varphi_{n,k_1}(t_1) \varphi_{n,k_2}(t_2) = 2^n \varphi(2^n t_1 - k_1) \varphi(2^n t_2 - k_2),$$

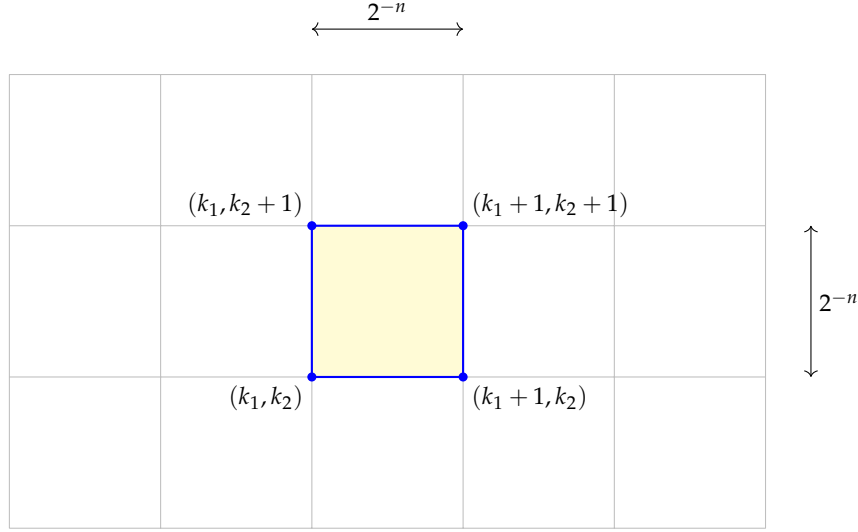
where $n \in \mathbb{Z}$ and $k = (k_1, k_2) \in \mathbb{Z}^2$, forms an orthonormal basis of $V_n^{(2)}$. In particular, the associated scaling function is

$$\varphi^{(2)}(t_1, t_2) = \varphi(t_1) \varphi(t_2).$$

Let us illustrate this with the Haar wavelet.

Example 3.2. Let $\{V_n\}_{n \in \mathbb{Z}}$ be the Haar MRA given in Section 2.2. Then, $V_n^{(2)}$ is the approximation space consisting on the functions in $L^2(\mathbb{R}^2)$ which are constant on dyadic squares

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

Figure 3.1: $Q_{n,k}$ defined in the Haar wavelet

Therefore, the two-dimensional scaling function is

$$\varphi^{(2)}(t_1, t_2) = \chi_{[0,1)}(t_1)\chi_{[0,1)}(t_2) = \chi_{[0,1) \times [0,1)}(t_1, t_2).$$

Given an MRA $\{V_n^{(2)}\}_{n \in \mathbb{Z}}$ of $L^2(\mathbb{R}^2)$, let $W_n^{(2)} = V_{n+1}^{(2)} \ominus V_n^{(2)}$ denote the corresponding detail space at level n .

Theorem 3.3. Let φ be the scaling function of an MRA $\{V_n\}_{n \in \mathbb{Z}}$ in $L^2(\mathbb{R})$ and let ψ denote the associated wavelet. Define

$$\psi^1(t_1, t_2) = \varphi(t_1)\psi(t_2),$$

$$\psi^2(t_1, t_2) = \psi(t_1)\varphi(t_2),$$

$$\psi^3(t_1, t_2) = \psi(t_1)\psi(t_2)$$

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

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

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

The functions ψ^1, ψ^2, ψ^3 are called the horizontal, vertical and diagonal wavelets, respectively. Every function is built from a factor φ and a oscillatory factor ψ , or from two oscillatory factors. The naming is a description of which coordinate carries the oscillation.

In the case where only one coordinate oscillates, the wavelet is more sensitive to changes perpendicular to that coordinate. On the other hand, when both coordinates oscillate, the wavelet responds to variations in both directions simultaneously. In this way, the three labels indicate the direction of strongest variation.

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

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

Therefore,

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

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

Example 3.4. Let φ and ψ be the one-dimensional Haar scaling function and Haar wavelet, respectively:

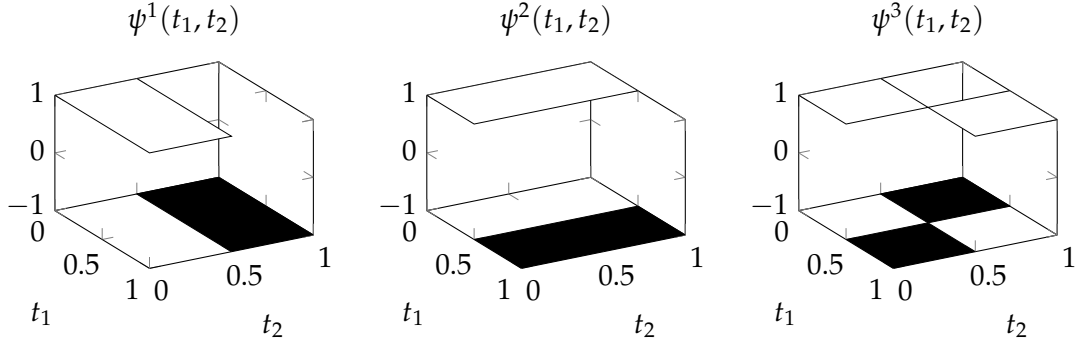
$$\varphi(t) = \begin{cases} -1 & \text{if } 0 \leq t < 1, \\ 0 & \text{otherwise,} \end{cases} \quad \psi(t) = \begin{cases} -1 & \text{if } 0 \leq t < \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} \leq t < 1. \end{cases}$$

These generate the Haar MRA $\{V_n\}_{n \in \mathbb{Z}}$ of $L^2(\mathbb{R})$, with

$$\varphi_{n,k}(t) = 2^{n/2} \varphi(2^n t - k), \quad \psi_{n,k}(t) = 2^{n/2} \psi(2^n t - k), \quad k \in \mathbb{Z}.$$

Then, the Haar wavelets in $\mathbb{R}^2$ given by the theorem are

$$\begin{aligned} \psi^1(t_1, t_2) &= \varphi(t_1) \psi(t_2) = \begin{cases} 1 & \text{if } (t_1, t_2) \in [0, 1) \times [0, \frac{1}{2}), \\ -1 & \text{if } (t_1, t_2) \in [0, 1) \times [\frac{1}{2}, 1), \end{cases} \\ \psi^2(t_1, t_2) &= \psi(t_1) \varphi(t_2) = \begin{cases} 1 & \text{if } (t_1, t_2) \in [0, \frac{1}{2}) \times [0, 1), \\ -1 & \text{if } (t_1, t_2) \in [\frac{1}{2}, 1) \times [0, 1), \end{cases} \\ \psi^3(t_1, t_2) &= \psi(t_1) \psi(t_2) = \begin{cases} 1 & \text{if } (t_1, t_2) \in [0, \frac{1}{2}) \times [0, \frac{1}{2}), \\ -1 & \text{if } (t_1, t_2) \in [0, \frac{1}{2}) \times [\frac{1}{2}, 1), \\ -1 & \text{if } (t_1, t_2) \in [\frac{1}{2}, 1) \times [0, \frac{1}{2}), \\ 1 & \text{if } (t_1, t_2) \in [\frac{1}{2}, 1) \times [\frac{1}{2}, 1). \end{cases} \end{aligned}$$

Figure 3.2: 3D plots of the functions ψ^1, ψ^2, ψ^3 .

Their dilates and translates are

$$\psi_{n,k_1,k_2}^j(t_1, t_2) = 2^n \psi^j(2^n t_1 - k_1, 2^n t_2 - k_2), \quad j = 1, 2, 3, k_1, k_2 \in \mathbb{Z}.$$

By the theorem, for every $n \in \mathbb{Z}$ the family $\{\psi_{n,k}^1, \psi_{n,k}^2, \psi_{n,k}^3\}_{k \in \mathbb{Z}^2}$ is an orthonormal basis of $W_n^{(2)}$. As a consequence, the system $\{\psi_{n,k}^1, \psi_{n,k}^2, \psi_{n,k}^3\}_{n \in \mathbb{Z}, k \in \mathbb{Z}^2}$ forms an orthonormal basis of $L^2(\mathbb{R}^2)$.

The separable construction extends naturally to $\mathbb{R}^p$, $p > 2$. Given the same notation as above, the associated *separable p -dimensional MRA* is defined by

$$V_n^{(p)} = \underbrace{V_n \otimes \cdots \otimes V_n}_{p \text{ times}}.$$

The corresponding p -dimensional scaling function is

$$\varphi^{(p)}(x_1, \dots, x_p) = \varphi(x_1) \cdots \varphi(x_p).$$

To describe the wavelets, let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_p) \in \{0, 1\}^p$ and set $f_0 = \varphi$, $f_1 = \psi$. Define

$$\Psi_\varepsilon(x_1, \dots, x_p) = f_{\varepsilon_1}(x_1) \cdots f_{\varepsilon_p}(x_p).$$

The case $\varepsilon = (0, \dots, 0)$ yields the scaling function $\varphi^{(p)}$, while each nonzero ε produces one of the $2^p - 1$ separable p -dimensional wavelets. Their dilated and translated versions are given by

$$\Psi_{\varepsilon,n,k}(x) = 2^{np/2} \Psi_\varepsilon(2^n x - k), \quad n \in \mathbb{Z}, k \in \mathbb{Z}^p.$$

Then, the collection

$$\{\Psi_{\varepsilon,n,k} : \varepsilon \neq 0, n \in \mathbb{Z}, k \in \mathbb{Z}^p\}$$

forms an orthonormal basis of $L^2(\mathbb{R}^p)$.

Remark 3.5. Although non-separable MRA's exist, they are rarely used in applications. Their construction is mathematically difficult, and their implementation offers little practical advantage. For this reason, multi-dimensional wavelet systems are almost always built from separable MRA's.

3.2 The continuous wavelet transform

The constructions developed in the preceding sections can be interpreted as discretised versions of the continuous wavelet transform. This section provides a brief overview of this transform.

The central idea is to analyse a signal through both translations and dilations of a single function. In contrast with the MRA structure, the continuous setting allows translations and dilations to vary over the real line, not only dyadic scales and integer shifts. Consequently, the transform is no longer based on an orthonormal basis generated by an MRA, but instead on a continuous family of wavelets indexed by scale and position.

Let $\psi \in L^2(\mathbb{R})$ be a wavelet function such that

1. $\|\psi\|_{L^2} = 1$,
2. ψ is either compactly supported or very rapidly decaying,
3. $\int_{\mathbb{R}} \psi(t) dt = 0$.

In many applications it is also important that ψ has some regularity, such as continuity or continuous differentiability.

Example 3.6. The Haar wavelet provides a function satisfying the above requirements. As one has seen before, it is defined by

$$\psi(t) = \begin{cases} 1 & \text{if } 0 \leq t < \frac{1}{2}, \\ -1 & \text{if } \frac{1}{2} \leq t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

It is immediate that $\psi \in L^2(\mathbb{R})$ and that it is compactly supported. Moreover, a direct computation shows that

$$\|\psi\|_{L^2}^2 = \int_0^{1/2} 1^2 dt + \int_{1/2}^1 (-1)^2 dt = \frac{1}{2} + \frac{1}{2} = 1.$$

Furthermore,

$$\int_{\mathbb{R}} \psi(t) dt = \int_0^{1/2} 1 dt + \int_{1/2}^1 (-1) dt = \frac{1}{2} - \frac{1}{2} = 0,$$

and therefore it satisfies the conditions described above.

In the next section, additional examples of wavelets of this type will be presented, including smoother and more regular families that are commonly used in practical applications.

Given $a > 0$ and $b \in \mathbb{R}$, define

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right).$$

as the associated family of dilated and translated wavelets.

For any $\psi \in L^2(\mathbb{R})$ one can also define the *involution* $\psi^*(t) := \overline{\psi(-t)}$.

Definition 3.7. The *continuous wavelet transform* of $f \in L^2(\mathbb{R})$ associated to a wavelet ψ is

$$Wf(a, b) = (f * \psi_a^*)(b) = \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt.$$

A fundamental property is that, under mild assumptions on ψ , the wavelet transform $Wf(a, b)$ contains enough information to reconstruct f . One has already seen this for the discrete case. Let us see the case for the continuous setting.

Theorem 3.8 (Inverse wavelet transform). Let $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a real-valued function satisfying the admissibility condition

$$c_\psi := \int_0^\infty \frac{|\widehat{\psi}(\xi)|^2}{\xi} d\xi < \infty.$$

Then, for every $f \in L^2(\mathbb{R})$,

$$f(t) = \frac{1}{c_\psi} \int_{\mathbb{R}} \int_0^\infty Wf(a, b) \psi_{a,b}(t) \frac{da}{a^2} db, \quad (3.1)$$

and the following Plancherel-type identity holds:

$$\|f\|_{L^2}^2 = \frac{1}{c_\psi} \int_{\mathbb{R}} \int_0^\infty |Wf(a, b)|^2 \frac{da}{a^2} db. \quad (3.2)$$

Proof. Let us first prove the reconstruction formula (3.1). The right side integral $s(t)$ of (3.1) can be rewritten as a sum of convolutions. Using $Wf(a, b) = (f * \overline{\psi_a})(b)$ with $\psi_a(t) = a^{-1/2} \psi(t/a)$, then

$$s(t) = \frac{1}{c_\psi} \int_0^\infty (Wf(a, \cdot) * \psi_a)(t) \frac{da}{a^2} = \frac{1}{c_\psi} \int_0^\infty (f * \overline{\psi_a} * \psi_a)(t) \frac{da}{a^2}.$$

Let us show that $s = f$ by comparing their Fourier transforms. Using Theorem 1.16 and Proposition 1.10, one sees that the Fourier transform of s is

$$\widehat{s}(\xi) = \frac{1}{c_\psi} \int_0^\infty \widehat{f}(\xi) \sqrt{a} \widehat{\psi}^*(a\xi) \sqrt{a} \widehat{\psi}(a\xi) \frac{da}{a^2} = \frac{\widehat{f}(\xi)}{c_\psi} \int_0^\infty |\widehat{\psi}(a\xi)|^2 \frac{da}{a}.$$

Since ψ is real-valued, $|\widehat{\psi}(-\xi)|^2 = |\widehat{\psi}(\xi)|^2$. Hence, the change of variables $\omega = a\xi$ yields

$$\widehat{s}(\xi) = \widehat{f}(\xi) \frac{1}{c_\psi} \int_0^\infty \frac{|\widehat{\psi}(\omega)|^2}{\omega} d\omega = \widehat{f}(\xi).$$

To prove the Plancherel identity (3.2), one can again use the Fourier representation of W_f . For fixed $a > 0$, Plancherel in the b -variable and Proposition 1.10 give

$$\int_{\mathbb{R}} |W_f(a, b)|^2 db = \int_{\mathbb{R}} |\widehat{f}(\xi)|^2 |\widehat{\psi}_a(\xi)|^2 d\xi = \int_{\mathbb{R}} |\widehat{f}(\xi)|^2 a |\widehat{\psi}(a\xi)|^2 d\xi.$$

Integrating with respect to $\frac{da}{a^2}$ yields

$$\int_0^\infty \int_{\mathbb{R}} |W_f(a, b)|^2 \frac{da}{a^2} db = \int_{\mathbb{R}} |\widehat{f}(\xi)|^2 \left(\int_0^\infty |\widehat{\psi}(a\xi)|^2 \frac{da}{a} \right) d\xi.$$

The inner integral equals c_ψ by the admissibility condition, so

$$\int_0^\infty \int_{\mathbb{R}} |W_f(a, b)|^2 \frac{da}{a^2} db = c_\psi \|f\|_{L^2}^2.$$

Dividing by c_ψ gives (3.2). □

The condition

$$c_\psi = \int_0^\infty \frac{|\widehat{\psi}(\xi)|^2}{\xi} d\xi < \infty$$

is called the *wavelet admissibility condition*.

Example 3.9. For the Haar wavelet,

$$\begin{aligned} \widehat{\psi}(\xi) &= \int_0^{1/2} e^{-2\pi i t \xi} dt - \int_{1/2}^1 e^{-2\pi i t \xi} dt = \frac{1}{2\pi i \xi} (1 - e^{-\pi i \xi} + e^{-2\pi i \xi} - e^{-\pi i \xi}) \\ &= \frac{(1 - e^{\pi i \xi})^2}{2\pi i \xi} = \frac{e^{-\pi i \xi} (e^{i \frac{\pi}{2} \xi} - e^{-i \frac{\pi}{2} \xi})^2}{2\pi i \xi} = -e^{-\pi i \xi} \frac{\sin^2(\frac{\pi \xi}{2})}{\frac{\pi}{2} \xi}. \end{aligned}$$

Therefore,

$$c_\psi = \int_0^\infty \frac{\sin^4(\frac{\pi}{2} \xi)}{\frac{\pi^2}{4} \xi^3} d\xi = \int_0^1 \frac{\sin^4(\frac{\pi}{2} \xi)}{\frac{\pi^2}{4} \xi^3} d\xi + \int_1^\infty \frac{\sin^4(\frac{\pi}{2} \xi)}{\frac{\pi^2}{4} \xi^3} d\xi$$

For the first integral, the only possible singularity is at $\xi = 0$. Near 0, the integrand behaves like ξ , a continuous function. Consequently, the first integral is finite.

For the second integral, bounding the sine function by 1,

$$\int_1^\infty \frac{\sin^4(\frac{\pi}{2}\xi)}{\frac{\pi^2}{4}\xi^3} d\xi \leq \frac{4}{\pi^2} \int_1^\infty \frac{d\xi}{\xi^3} = \frac{4}{\pi^2} \frac{1}{2} < \infty.$$

Thus, $c_\psi < \infty$ and therefore one has the wavelet admissibility condition.

Chapter 4

Examples and applications

The preceding chapters have focused on the foundations of wavelet analysis, with the Haar system serving as the main example. Although the Haar wavelet is fundamental from a conceptual point of view, practical applications typically use smoother wavelet families. The purpose of this final chapter is to present examples of wavelets commonly used in signal and image processing.

The first section introduces several classical families of wavelets, including the Shannon, Meyer, and Daubechies wavelets. These examples are among the most used in practice. They demonstrate how the general framework developed earlier is not restricted to the Haar construction.

The second section discusses how wavelets can be used for compression and denoising. The main idea is to represent a signal with as few significant coefficients as possible while suppressing noise.

The third section introduces applications to signal processing. Wavelets are useful in this setting because they break a signal into components that reveal how it behaves at different times and scales. This makes it possible to process signals without damaging important features.

The final section turns to specific applications in image processing. In this context, wavelets form the basis of several compression standards, such as the JPEG 2000 format. We compare it to the first JPEG method and highlight the properties that make wavelet compression more effective.

4.1 Shannon, Meyer and Daubechies wavelets

4.1.1 The Shannon wavelet

The Shannon scaling function is defined in the Fourier domain by

$$\hat{\varphi}(\xi) = \chi_{[-\frac{1}{2}, \frac{1}{2}]}(\xi).$$

Therefore, the Shannon scaling function may be viewed as the Fourier-domain counterpart of the Haar scaling function. The Haar scaling function is $\varphi_H(t) = \mathbf{1}_{[0,1)}(t)$, which is compactly supported in time, while the Shannon scaling function is compactly supported in the frequency domain.

The fact that $\hat{\varphi}$ is supported in the interval $[-\frac{1}{2}, \frac{1}{2}]$ means that only the frequency components lying in this range are preserved when a function is projected onto the approximation space V_0 . Thus

$$V_0 = \left\{ f \in L^2(\mathbb{R}) : \text{supp}(\hat{f}) \subseteq [-\frac{1}{2}, \frac{1}{2}] \right\},$$

since this choice already satisfies the MRA properties. This shows that V_0 contains precisely those functions whose Fourier transforms are supported in the low-frequency region. Using the construction of the spaces V_n in the MRA, one obtains

$$V_n = \left\{ f \in L^2(\mathbb{R}) : \text{supp}(\hat{f}) \subseteq [-2^{n-1}, 2^{n-1}] \right\}.$$

Thus, finer detail spaces provide functions with larger frequency support. The associated wavelet satisfies

$$\hat{\psi}(\xi) = \chi_{[-1, \frac{1}{2}) \cup (\frac{1}{2}, 1]}(\xi),$$

hence, by a computation with the inversion formula one obtains

$$\psi(t) = \frac{\sin(2\pi t) - \sin(\pi t)}{\pi t}.$$

Then, the Shannon MRA can be interpreted as an ideal low-pass filtering procedure. The projection onto V_0 eliminates all high-frequency components of a signal and what remains is the coarse approximation. On the other hand, the associated Shannon wavelet isolates the complementary frequency band. However, ψ has slow asymptotic decay and is not smooth in the frequency domain.

4.1.2 The Meyer wavelet

The Meyer wavelet is a frequency band-limited function that has a Fourier transform that is smooth, unlike the Fourier transform of the Shannon wavelet. This smoothness provides a much faster asymptotic decay in time.

Start by setting a smooth function $\nu(\xi)$ such that

- $\nu(\xi)$ is even, that is, symmetric around 0.
- $\nu(\xi) = 1$ for $\xi \in [-\frac{2\pi}{3}, \frac{2\pi}{3}]$.
- $\nu(\xi) = 0$ outside $[-\frac{4\pi}{3}, \frac{4\pi}{3}]$.
- For $0 \leq \xi \leq 2\pi$, the relation $\nu^2(\xi) + \nu^2(2\pi - \xi) = 1$ holds, ensuring a smooth transition between the low and high frequency regions.

Then define a function $g(t)$ such that $\widehat{g}(\xi) = \nu(\xi)$. The translated functions $g(x - k)$, for $k \in \mathbb{Z}$, form an orthonormal basis of a closed subspace of $L^2(\mathbb{R})$, which becomes the approximation space V_0 in the Meyer MRA.

The corresponding wavelet is obtained by assigning the complementary frequency band to the detail space. The Fourier transform of the Meyer wavelet is defined by

$$\widehat{\psi}_M(\xi) = \begin{cases} \sin\left(\frac{\pi}{2} \nu\left(\frac{3|\xi|}{2\pi} - 1\right)\right) & \text{if } \frac{2\pi}{3} \leq |\xi| \leq \frac{4\pi}{3}, \\ 0 & \text{otherwise,} \end{cases}$$

This formula shows that the function ν controls the transition between the low and high-frequency regions, now being smooth.

The associated scaling function is defined on the low-frequency region by

$$\widehat{\varphi}_M(\xi) = \begin{cases} 1 & \text{if } |\xi| \leq \frac{2\pi}{3}, \\ \cos\left(\frac{\pi}{2} \nu\left(\frac{3|\xi|}{2\pi} - 1\right)\right) & \text{if } \frac{2\pi}{3} \leq |\xi| \leq \frac{4\pi}{3}, \\ 0 & \text{otherwise.} \end{cases}$$

In contrast with the Shannon construction, the Meyer functions are $\mathcal{C}^\infty$ in the frequency domain. This smoothness leads to rapid decay in the time domain. Therefore, the Meyer wavelet is more suitable in applications where regularity is important. Moreover, it still preserves frequency localization. However, it is no longer compactly supported in time.

4.1.3 Daubechies wavelets

Up to this point, all the wavelets considered have been infinitely supported. This is a consequence of their construction, since their Fourier transforms are compactly supported.

The lack of compact support is a limitation in applications. It reduces time localization and increases the computational cost. Daubechies wavelets overcome these problems. They provide the first family of compactly supported orthonormal wavelets, while still allowing a certain number of vanishing moments.

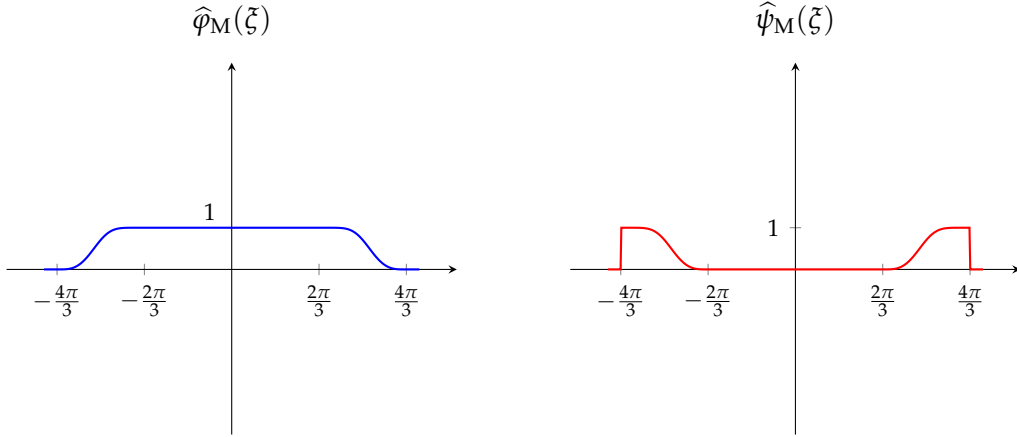


Figure 4.1: Fourier transforms of the Meyer scaling function and the Meyer wavelet using the auxiliary function $\nu(x) = \frac{f(x)}{f(x)+f(1-x)}$ defined by the exponentials $f(x) = \exp(-1/x^2)$ for $x > 0$.

As shown by I. Daubechies, for any $n \in \mathbb{N}$, one can construct a wavelet ψ with compact support such that all derivatives of ψ up to order n exist and are bounded. The construction relies on the notion of MRA introduced in the previous sections.

Start by setting a 2π -periodic function m_0 such that

- $m_0 \in \mathcal{C}^n$.
- $m_0(0) = 1$.
- $|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1$.

A convenient way to construct such filters is through the Daubechies polynomials. For a fixed integer $N \geq 1$, consider the polynomial

$$P_N(x) = \sum_{k=0}^{N-1} \binom{N-1+k}{k} x^k, \quad x \in [0, 1].$$

This polynomial is strictly positive on $(0, 1)$ and satisfies $P_N(1) = \binom{2N-2}{N-1}$. Daubechies showed that the function

$$|m_0(\xi)|^2 = \cos^{2N} \left(\frac{\xi}{2} \right) P_N \left(\sin^2 \left(\frac{\xi}{2} \right) \right)$$

defines a 2π -periodic trigonometric polynomial that satisfies the orthonormality condition

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1.$$

These conditions are necessary for the refinement equation

$$\widehat{\varphi}(\xi) = m_0(\xi/2) \widehat{\varphi}(\xi/2) \quad (4.1)$$

to admit a non-trivial solution. This function m_0 plays the same role as the refinement mask H introduced in Section 2.5. The two functions are related by the identity $m_0(\xi) = H(\xi/2\pi)$.

Iterating (4.1), yields the product

$$\widehat{\varphi}(\xi) = \widehat{\varphi}(2^{-n}\xi) \prod_{j=1}^n m_0(2^{-j}\xi) \equiv \widehat{\varphi}(2^{-n}\xi) \Pi_n(\xi),$$

where $\Pi_n(\xi) = \prod_{j=1}^n m_0(2^{-j}\xi)$.

The sequence $\{\Pi_n(\xi) \mid n \in \mathbb{N}\}$ converges uniformly on bounded sets if m_0 is $\mathcal{C}^1$. Then, it defines a function $\varphi \in L^2(\mathbb{R})$ by letting

$$\widehat{\varphi}(\xi) = \prod_{j=1}^{\infty} m_0(2^{-j}\xi) \quad \text{for } \xi \in \mathbb{R}.$$

The refinement equation implies that φ has compact support, and the associated wavelet ψ is compactly supported as well. The Haar wavelet corresponds to the simplest choice of such a filter.

Compact support imposes restrictions on smoothness. A compactly supported wavelet cannot be infinitely differentiable. The smoothness that can be achieved is limited by the number of vanishing moments imposed on the wavelet and by the structure of the filter m_0 .

Daubechies wavelets are practical because their associated transforms involve only a finite number of non-zero coefficients. At the same time, the presence of vanishing moments allows the wavelets to capture smooth trends.

4.2 Compression and denoising

One of the main goals in signal and image processing is to be able to code the information with as few data as possible. Having fewer data points allows faster transmission and easier storage. In the presence of noise, one also wants to separate the noise from the signal. Moreover, one would like to have a basis that concentrates the signal in a few large coefficients and the noise in very small coefficients. Then, one can proceed with the following steps.

1. Transform the data. Find the coefficients with respect to a given basis.

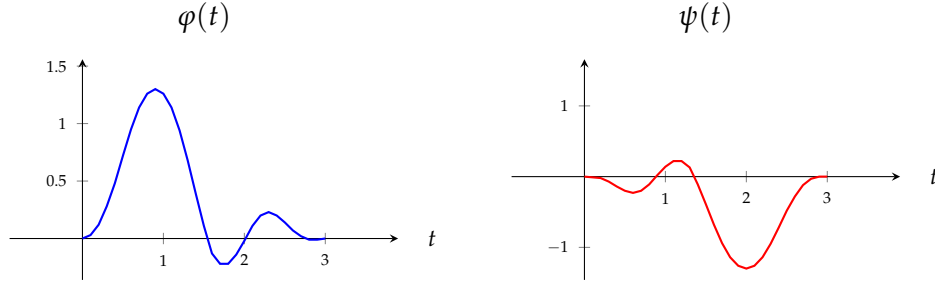


Figure 4.2: The Daubechies scaling function and the Daubechies wavelet function for order $N = 2$. These plots correspond to the DB2 wavelet, generated via numerical approximations. The graphs show the property of compact support, since they vanish outside $[0, 3]$.

2. Threshold the coefficients. Keep the large ones and discard the small ones. Information is lost in this step, so perfect reconstruction is no longer possible.
3. Reconstruct with the thresholded coefficients. One uses only the coefficients kept in the previous step.

The goal is that the resulting compressed signal approximates the original signal well enough.

When the basis in step 1 is a wavelet basis, one simple approach for step 2 is to use the projection onto an approximation space as your compressed signal, as defined in (2.2). Then, discard all the details after a certain scale k .

In terms of how to approximate the original signal, there are two possible methods:

- *Linear approximation.* Select N elements in the basis and project onto the subspace generated by those elements, regardless of the function that is being approximated. One will have

$$P_N^{\text{linear}} f := \sum_{n=1}^N \langle f, \psi_n \rangle \psi_n.$$

- *Nonlinear approximation.* Choose the basis elements depending on the function. For instance, the N basis elements could be chosen so that the coefficients are the largest in size for the particular function. This time, the chosen basis elements depend on the function being approximated. One will have

$$P_N^{\text{non-linear}} f := \sum_{n=1}^N \langle f, \psi_{n,f} \rangle \psi_{n,f}.$$

The non-linear approach has proven successful.

4.3 Signal processing

Since most real signals are locally smooth, only a small number of coefficients have large magnitude. Thus, most of their wavelet coefficients are small. Since wavelets with vanishing moments cancel out smooth behaviour, only coefficients located near singularities have large values. Consequently, the representation is sparse and useful for transform coding.

Keeping the M largest coefficients yields the optimal M approximation. If f_M denotes this approximation and Λ_M its index set, then the approximation error is

$$\|f - f_M\|_2^2 = \sum_{(j,k) \notin \Lambda_M} |f_{j,k}|^2.$$

Quantizing a coefficient means replacing its exact real value by one element of a finite set of allowed values. In practice, the real line is divided into intervals named bins, and all coefficients falling inside the same interval are mapped to the same representative value. This operation reduces the precision but is more suitable for digital encoding.

Information is represented using bits, that is binary units that take values in $\{0, 1\}$. A coefficient encoded with b bits can take 2^b distinct quantized values. The total number of bits required to encode a signal depends on its value and position. When the wavelet representation is sparse, the number of significant coefficients is small. However, their locations among all possible indices becomes the dominant contribution.

To reduce the total bit cost, one can use that quantized values do not appear with equal frequency. Hence, one can represent the frequent values using fewer bits than the rare ones. This leads to a reduction in the total number of used bits.

Wavelet methods are also widely used in denoising. Consider a signal f that is polluted by additive noise. Then, if one sets a cutoff value for the wavelet coefficients, f has an estimator

$$\tilde{f}(t) = \sum_{j,k} T_\lambda(f_{j,k}) \psi_{j,k}(t),$$

where T_λ is a thresholding operator. This procedure removes noise coefficients while preserving the significant ones.

4.4 Image processing

Wavelets can also be applied to two-dimensional settings, such as image processing. The idea is the same as in the one-dimensional case. The most used

wavelets are the Daubechies wavelets, because of their compact support, regularity and high order zero moments. We will focus on the JPEG 2000 algorithm of image compression.

4.4.1 JPEG 2000

The JPEG 2000 standard is a coding system operating in a wavelet basis. In the classical JPEG method, the image was divided into blocks and each block was represented as a weighted sum of cosine functions. This method presented blocking parts at high compression rates. JPEG 2000 avoids this issue. It uses an algorithm that allows the image to load progressively. One can also define regions of interest to improve the resolution on specific image parts.

The selection of an optimal wavelet basis aims to maximize the number of zero-quantized coefficients. This requires balancing the number of vanishing moments, the support size, and the regularity of the wavelet ψ . Analogously to the case of signals, high coefficients detect irregularities, such as contours and edges.

JPEG 2000 improves compression efficiency by analyzing pixel relationships and taking advantage of similar patterns in neighboring areas. It restricts the implementation to two principal wavelet families.

1. The 9/7 Symmetric Biorthogonal Wavelet Family.

The 9/7 wavelet family is named as the primary mechanism for lossy compression configurations. The low-pass filter H has 9 coefficients, while the high-pass filter G has 7 coefficients. The numerical values of its coefficients are irrational numbers derived from the factorization of Daubechies polynomials, mentioned in Section 4.1.3. The following table shows an approximation of this coefficients.

Analysis low pass filter	Analysis high pass filter	Synthesis low pass filter	Synthesis high pass filter
0.026748757411	0	0	0.026748757411
-0.016864118443	0.091271763114	-0.091271763114	0.016864118443
-0.078223266529	-0.057543526229	-0.057543526229	-0.078223266529
0.266864118443	-0.591271763114	0.591271763114	-0.266864118443
0.602949018236	1.11508705	1.11508705	0.602949018236
0.266864118443	-0.591271763114	0.591271763114	-0.266864118443
-0.078223266529	-0.057543526229	-0.057543526229	-0.078223266529
-0.016864118443	0.091271763114	-0.091271763114	0.016864118443
0.026748757411	0	0	0.026748757411

Table 4.1: Filter coefficients for the 9/7 Biorthogonal Wavelet family

The operational efficacy of this family is determined by the fact that its wavelet exhibits 4 vanishing moments. Consequently, it produces zero or near-zero values across regions where changes are smoothly enough to be approximated by a cubic polynomial.

2. The 5/3 Symmetric Biorthogonal Wavelet Family.

For situations where perfectly lossless compression is required, the 5/3 wavelet family is used. It uses a low-pass filter H with 5 coefficients and a high-pass filter G with 3 coefficients, both defined using rational numbers.

The 5/3 wavelet transform has the advantage that all of its steps can be carried out using exact integer operations. This makes the transform perfectly reversible, since the original signal can be recovered exactly from its detail and approximation coefficients.

Since the human visual system is less sensitive to compression errors in fine details than in coarse structures, JPEG 2000 applies a quantization step size $\Delta_j = \Delta \cdot w_j$ that depends on the scale j . This approach allows specific regions of interest (ROI) to be prioritized and encoded first.

After applying the wavelet transform, coefficients are grouped into independent square blocks called code-blocks, denoted S_k . As each code-block is encoded, the encoder produces a sequence of progressively more accurate descriptions of that block. Any point where this sequence can be stopped while still yielding a valid reconstruction is called a truncation point R_j^n .

The distortion in block S_k when stopping at truncation point R_k^n is

$$d_k^n = \sum_{m \in S_k} |Q(f_B[m]) - f_B[m]|^2,$$

where Q is a uniform quantizer with step size 2^n .

To construct the final compressed file, the encoder must decide how much information to keep from each code-block. Choosing a truncation point means deciding how many bits are allocated to block S_k , and therefore how much distortion d_k^n remains.

The goal is to select one truncation point for each block so that the total distortion is as small as possible while keeping the total number of bits below a fixed budget R :

$$\min_{\{n_k\}_k} \sum_k d_k^{n_k} \quad \text{subject to} \quad \sum_k R_k^{n_k} \leq R.$$

This is a constrained optimization problem. To make it easier to solve, one can introduce a Lagrange multiplier λ and convert it into an unconstrained objective

$$L(\lambda) = \sum_k d_k^{n_k} + \lambda \sum_k R_k^{n_k}.$$

For a fixed value of λ , each block independently chooses the truncation point n_k that minimizes its local cost

$$d_k^{n_k} + \lambda R_k^{n_k}.$$

By going over different values of λ , the encoder ranks all truncation points according to how efficiently they reduce distortion per bit. The final compressed stream is then connected by selecting the most efficient points until the total rate reaches the limit R .

Bibliography

- [1] C. K. Chui, *An Introduction to Wavelets*, Academic Press, 1992.
- [2] A. Cohen, I. Daubechies and J.-C. Feauveau, *Biorthogonal Bases of Compactly Supported Wavelets*, *Communications on Pure and Applied Mathematics*, **45**, no. 5 (1992), 485-560.
- [3] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.
- [4] E. Hernández and G. Weiss, *A First Course on Wavelets*, CRC Press, 1996.
- [5] S. G. Mallat, *Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$* , *Transactions of the American Mathematical Society*, **315**, no. 1 (1989), 69-87.
- [6] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd edition, Academic Press, 2008.
- [7] M. C. Pereyra and L. A. Ward, *Harmonic Analysis: From Fourier to Wavelets*, Student Mathematical Library, IAS/Park City Mathematical Subseries, Vol. 63, American Mathematical Society, 2012.
- [8] J. S. Walker, *A Primer on Wavelets and Their Scientific Applications*, CRC Press, 1999.
- [9] D. F. Walnut, *An Introduction to Wavelet Analysis*, Birkhäuser, 2002.