ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Chair for Mathematical Information Science

Verner Vlačić
Sternwartstrasse 7
CH-8092 Zürich

Mathematics of Information
Orthonormal Wavelets

These notes are based on [1], [2], and [3].

1 Introduction

There are many structures, such as Gabor frames, frames of wavelets, curvelets, shearlets, countourlets, bandlets, etc., useful for obtaining (sparse) approximations of various types of signals, e.g. audio signals, natural images, seismic measurements. In this chapter we will cover a broad class of orthonormal wavelets, including a general method for their construction (MRA), the associated algorithms for their implementation (FWT), as well as results explaining their suitability as a sparsity basis.

1.1 A Historical Overview of Wavelets

The first wavelet basis was introduced in 1909 by Haar, who constructed an orthonormal basis by dilating and translating a piecewise constant functions. In the late 1970s, the analysis of geophysical signals led to the formalization of the continuous wavelet transform by Morlet and Grossmann. In 1980, Strömberg extended Haar's work and found a piecewise linear function, different from the one used by Haar, that generates an orthonormal basis and gives better approximations of smooth functions. Meyer did not know this result, and motivated by the work of Morlet and Grossman on the continuous wavelet transform, he tried to prove that there exists no regular wavelet generating an orthonormal basis. This attempt was a failure since he ended up constructing a whole family of orthonormal wavelet bases, with functions that are infinitely continuously differentiable and possess very good time and frequency localization properties. This last result marked the beginning of a widespread search for new orthonormal wavelet bases, leading in particular to Daubechies wavelets of compact support. A systematic theory for constructing orthonormal wavelet bases was established by Meyer and Mallat through the elaboration of multiresolution signal approximations, which we will study in this discussion session. Mallat made the important discovery that there is a fast algorithm for the wavelet transform. His proposed pyramidal algorithm based on convolutions brought the breakthrough for the applicability of the wavelet transform theory.

Example 1 (Haar Wavelet Basis). Define the functions φ and ψ as follows:

$$\varphi(t) = \begin{cases} 1, & 0 \leqslant t \leqslant 1 \\ 0, & \text{elsewhere} \end{cases} \qquad \psi(t) = \begin{cases} 1, & 0 \leqslant t < 1/2 \\ -1, & 1/2 < t \leqslant 1 \\ 0, & \text{elsewhere.} \end{cases}$$

and set

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

We show that the set $\mathcal{B} = \{\varphi_{0,k}\}_{k\in\mathbb{Z}} \cup \{\psi_{j,k}\}_{j\in\mathbb{N}_0,k\in\mathbb{Z}}$ forms an orthonormal basis for $L^2(\mathbb{R})$. To show the orthonormality, see Homework 3, Problem 7 (Haar wavelets). In order to prove the completeness of $\mathcal{B}$, let

$$\mathcal{V}_j = \operatorname{span}\{\varphi_{j,k}\}_{k\in\mathbb{Z}}$$
$$\mathcal{W}_j = \operatorname{span}\{\psi_{j,k}\}_{k\in\mathbb{Z}}.$$

Then, the set $\mathcal{V}_j$ consists of all step functions on $\mathbb{R}$ that are constant on the dyadic intervals $I_{j,k} \triangleq [2^{-j}k, 2^{-j}(k+1)]$, $k \in \mathbb{Z}$ for fixed j. It is not too difficult to show that this implies

$$\mathcal{V}_j \subseteq \mathcal{V}_{j+1}$$
$$\bigcap_{j\in\mathbb{Z}} \mathcal{V}_j = \{0\}$$
$$\overline{\bigcup_{j\in\mathbb{N}_0} \mathcal{V}_j} = L^2(\mathbb{R}).$$

Furthermore, we can decompose $\mathcal{V}_{j+1}$ as orthogonal direct sum $\mathcal{V}_{j+1} = \mathcal{V}_j \oplus \mathcal{W}_j$ (see Figure 1):

Therefore, we have

$$\begin{aligned}
\mathcal{V}_1 &= \mathcal{V}_0 \oplus \mathcal{W}_0 \\
\mathcal{V}_2 &= \mathcal{V}_1 \oplus \mathcal{W}_1 = \mathcal{V}_0 \oplus \mathcal{W}_0 \oplus \mathcal{W}_1, \\
\mathcal{V}_3 &= \mathcal{V}_2 \oplus \mathcal{W}_2 = \mathcal{V}_0 \oplus \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \mathcal{W}_2, \\
&\vdots \\
\mathcal{V}_n &= \mathcal{V}_0 \oplus \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \cdots \oplus \mathcal{W}_n.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\operatorname{span}(\mathcal{B}) &= \operatorname{span}\left(\mathcal{V}_0 \cup \bigcup_{j=0}^{\infty} \mathcal{W}_j\right) \\
&= \operatorname{span}\left(\bigcup_{n=0}^{\infty} (\mathcal{V}_0 \oplus \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \cdots \oplus \mathcal{W}_n)\right) \\
&= \operatorname{span}\left(\bigcup_{n=0}^{\infty} \mathcal{V}_n\right).
\end{aligned}$$

Taking the closure of this space thus gives

$$\overline{\operatorname{span}(\mathcal{B})} = \overline{\operatorname{span}\left(\bigcup_{n=0}^{\infty} \mathcal{V}_n\right)} = L^2(\mathbb{R}),$$

meaning that $\mathcal{B}$ is complete.

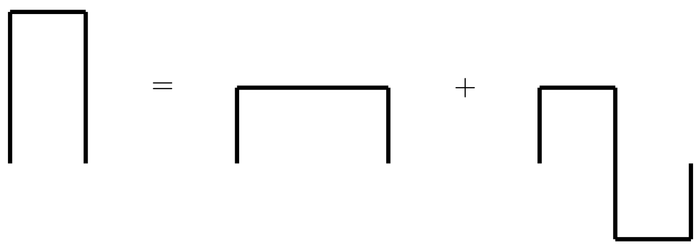

Figure 1: $\sqrt{2}\varphi_{j+1,k} = \varphi_{j,k} + \psi_{j,k}$.

We can see that each Haar wavelet $\psi_{j,k}$ has a zero average over its support $[2^{-j}k, 2^{-j}(k+1)]$. Therefore, if a function $f \in L^2(\mathbb{R})$ is locally regular and j is big, then it is nearly constant over this interval and thus, the wavelet coefficient $\langle f, \psi_{j,k} \rangle$ is nearly zero. This means that large wavelet coefficients are located at sharp signal transitions only and signals that are piecewise regular have an approximately sparse wavelet representation. This property is used in the JPEG-2000 compression standard.

2 Multiresolution Approximation

Multiresolution signal approximations (MRA) were inspired by original ideas developed in computer vision by Burt and Adelson to analyze images at several resolutions. MRAs provide a method of constructing orthonormal bases like the Haar basis.

Definition 1 (Multiresolution approximation). A multiresolution approximation of $L^2(\mathbb{R})$ is a sequence $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ of closed linear subspaces of $L^2(\mathbb{R})$ with the following properties[1]

(i) $\mathcal{V}_j \subset \mathcal{V}_{j+1}$, for all $j \in \mathbb{Z}$,

(ii) for all $f \in L^2(\mathbb{R})$ and all $j \in \mathbb{Z}$, $f \in \mathcal{V}_j \iff f(2\,\cdot) \in \mathcal{V}_{j+1}$,

(iii) $\bigcap_{j\in\mathbb{Z}} \mathcal{V}_j = \{0\}$,

(iv) $\bigcup_{j\in\mathbb{Z}} \mathcal{V}_j$ is dense in $L^2(\mathbb{R})$, and

(v) there exists a function $\varphi \in \mathcal{V}_0$, known as the scaling function, such that $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of the space $\mathcal{V}_0$.

Examples.

1. Piecewise constant approximations
   A simple multiresolution approximation of $L^2(\mathbb{R})$ is composed of piecewise constant functions. The space $\mathcal{V}_j$ then contains all the functions which are constant on the intervals $[2^{-j}k, 2^{-j}(k+1)))$ for $k \in \mathbb{Z}$. The approximation at a resolution 2^{-j} of a function f is thus the closest piecewise constant function on intervals of size 2^{-j}.

[1] By abuse of notation, we denote by $f(t)$ the function f.

2. Shannon approximations
   Frequency-limited functions also yield a multiresolution approximation of $L^2(\mathbb{R})$. The space $\mathcal{V}_j$ can be defined as

$$\mathcal{V}_j = \left\{ f \in L^2(\mathbb{R})\colon\ \operatorname{supp}(\mathcal{F}f) \subseteq [-2^j, 2^j] \right\}.$$

We often interpret $\mathcal{V}_j$ as the space of approximation at resolution 2^{-j}. The projection $P_{\mathcal{V}_j} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ can then be interpreted as a map taking in a function in $L^2(\mathbb{R})$ and outputting its approximation obtained by averaging over characteristic lengthscales 2^{-j}. For instance, if our MRA is the piecewise contant approximation in Example 1., then $P_{\mathcal{V}_j} f$ is the function obtained by averaging f over intervals of length 2^{-j}. For other MRAs, this averaging analogy is not as straightforward, but the intuition and the terminology remains the same. The following lemma tells us that Condition (iii) of an MRA corresponds to $P_{\mathcal{V}_j}$ losing all detail in a vector x, as $j \to -\infty$, whereas Condition (iv) corresponds to $P_{\mathcal{V}_j}$ recovering x entirely, as $j \to +\infty$.

Lemma 1. Let $\mathcal{X}$ be a Hilbert space and $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ a sequence of closed linear subspaces of $\mathcal{X}$ such that $\mathcal{V}_j \subset \mathcal{V}_{j+1}$, for all $j \in \mathbb{Z}$. Let $P_{\mathcal{V}_j} : \mathcal{X} \to \mathcal{X}$ denote the orthonormal projection onto $\mathcal{V}_j$, for all $j \in \mathbb{Z}$. Then,

(i) $\bigcap_{j\in\mathbb{Z}} \mathcal{V}_j = \{0\}$ if and only if $\lim_{j\to-\infty} \|P_{\mathcal{V}_j} x\| \to 0$, for all $x \in \mathcal{X}$.

(ii) $\overline{\bigcup_{j\in\mathbb{Z}} \mathcal{V}_j} = L^2(\mathbb{R})$ if and only if $\lim_{j\to\infty} \|x - P_{\mathcal{V}_j} x\| \to 0$, for all $x \in \mathcal{X}$.

The following two theorems tell us that, when constructing an MRA, it suffices to ensure that we have properties (i), (ii), and (v) with a φ such that $\hat{\varphi}$ is continuous at zero and $\hat{\varphi}(0) = 1$. The properties (iii) and (iv) follow automatically.

Theorem 1 ([2, Thm.1.6])**.** Let $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ be a sequence of closed linear subspaces of $L^2(\mathbb{R})$ satisfying conditions (i), (ii), and (v) of Definition 1. Then (iii) is satisfied as well.

Theorem 2 ([2, Thm.1.7])**.** Let $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ be a sequence of closed linear subspaces of $L^2(\mathbb{R})$ satisfying conditions (i), (ii), and (v) of Definition 1. Assume $\varphi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and $\hat{\varphi}$ is continuous at 0. Then

$$\hat{\varphi}(0) \neq 0 \iff \overline{\bigcup_{j\in\mathbb{Z}} \mathcal{V}_j} = L^2(\mathbb{R}).$$

Moreover, if either of the two equivalent statements holds, then $\hat{\varphi}(0) = 1$.

With this, we are ready to see how an MRA generates an orthonormal wavelet basis. Let $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ be an MRA. Let $\mathcal{W}_0$ be the orthogonal complement of $\mathcal{V}_0$ in $\mathcal{V}_1$, so that $\mathcal{V}_1 = \mathcal{V}_0 \oplus \mathcal{W}_0$. If we dilate the elements of $\mathcal{W}_0$ by 2^j (i.e., for $f \in \mathcal{W}_0$ consider $f(2^j\cdot)$), we get $\mathcal{W}_j$ such that

$$\mathcal{V}_{j+1} = \mathcal{V}_j \oplus \mathcal{W}_j, \quad \text{for all } j \in \mathbb{Z},$$

simply because $\mathcal{V}_j$ is obtained by dilating the elements of $\mathcal{V}_0$ by 2^j. Therefore, for all integers j, l such that $l \leqslant j$, we have the following orthogonal decomposition:

$$\mathcal{V}_{j+1} = \mathcal{V}_l \oplus \mathcal{W}_l \oplus \mathcal{W}_{l+1} \oplus \cdots \oplus \mathcal{W}_j.$$

Thus, by Pythagoras' identity, we have

$$\|P_{\mathcal{V}_{j+1}}f\|^2 = \|P_{\mathcal{V}_l}f\|^2 + \|P_{\mathcal{W}_l}f\|^2 + \|P_{\mathcal{W}_{l+1}}f\|^2 + \cdots + \|P_{\mathcal{W}_j}f\|^2$$

for all j, l such that $l \leqslant j$, and all $f \in L^2(\mathbb{R})$. Now, using Lemma 1 to first let $l \to -\infty$, and then $j \to \infty$, we obtain

$$\|f\|^2 = \sum_{j\in\mathbb{Z}} \|P_{\mathcal{W}_j}f\|^2, \tag{1}$$

for all $f \in L^2(\mathbb{R})$. If we can find a $\psi \in \mathcal{W}_0$ such that $\{\psi_{0,k} = \psi(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal basis of $\mathcal{W}_0$, then, for each $j \in \mathbb{Z}$, $\{\psi_{j,k} : k \in \mathbb{Z}\}$ is an ONB for $\mathcal{W}_j$. This implies

$$P_{\mathcal{W}_j}f = \sum_{k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle \psi_{j,k},$$

and thus

$$\|P_{\mathcal{W}_j}f\|^2 = \sum_{k\in\mathbb{Z}} |\langle f, \psi_{j,k}\rangle|^2, \tag{2}$$

for each $j \in Z$. Combining (1) with (2), we obtain

$$\|f\|^2 = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} |\langle f, \psi_{j,k}\rangle|^2.$$

As $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ is an orthonormal system obeying energy conservation, it is an ONB for $L^2(\mathbb{R})$ (formally, by Theorem 11 in the chapter on Hilbert spaces).

Definition 2. Let $\psi \in L^2(\mathbb{R})$. We say that ψ is an orthonormal wavelet for $L^2(\mathbb{R})$ if $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ is an orthonormal basis for the Hilbert space $L^2(\mathbb{R})$, where we define $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$, $\forall t \in \mathbb{R}$.

We have shown that, given an MRA, we can construct a wavelet. However, it is *not* the case that every wavelet ψ comes from an MRA in the way described above.

Proposition 2.1 ([2, Prop.1.11])**.** Let $g \in L^2(\mathbb{R})$. Then $\{g(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal system if and only if

$$\sum_{n\in\mathbb{Z}} |(\mathcal{F}g)(\omega + n)|^2 = 1, \quad \text{for all } \omega \in \mathbb{R}. \tag{3}$$

Proof. Note that, since both sides of (3) are 1-periodic functions, it suffices to consider $\omega \in [0, 1]$. Let $\delta_{k,l} = 1$ if $k = l$, and $\delta_{k,l} = 0$ if $k \neq l$. The statement that $\{g(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal system is equivalent to

$$\int_{\mathbb{R}} g(x - l)\overline{g(x - k)}\mathrm{d}x = \delta_{k,l} \quad \text{for all } k, l \in \mathbb{Z}.$$

By Plancherel's identity we have

$$\int_{\mathbb{R}} g(x - l)\overline{g(x - k)}\mathrm{d}x = \int_{\mathbb{R}} (\mathcal{F}g(\cdot - l))(\omega)\overline{(\mathcal{F}g(\cdot - k))(\omega)}\mathrm{d}\omega = \int_{\mathbb{R}} (\mathcal{F}g)(\omega)e^{-2\pi i l\omega}\overline{(\mathcal{F}g)(\omega)}e^{2\pi i k\omega}\mathrm{d}\omega$$

for all $k, l \in \mathbb{Z}$. We now calculate

$$\begin{aligned}
\int_{\mathbb{R}} g(x-l)\overline{g(x-k)}\mathrm{d}x &= \int_{\mathbb{R}} |(\mathcal{F}g)(\omega)|^2 e^{-2\pi i(l-k)\omega}\mathrm{d}\omega \\
&= \sum_{n\in\mathbb{Z}} \int_n^{n+1} |(\mathcal{F}g)(\omega)|^2 e^{-2\pi i(l-k)\omega}\mathrm{d}\omega \\
&= \sum_{n\in\mathbb{Z}} \int_0^1 |(\mathcal{F}g)(\omega+n)|^2 e^{-2\pi i(l-k)(\omega+n)}\mathrm{d}\omega \\
&= \int_0^1 \left(\sum_{n\in\mathbb{Z}} |(\mathcal{F}g)(\omega+n)|^2\right) e^{-2\pi i(l-k)\omega}\mathrm{d}\omega
\end{aligned}$$

Therefore, $\{g(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal system if and only if

$$\int_0^1 \left(\sum_{n\in\mathbb{Z}} |(\mathcal{F}g)(\omega+n)|^2\right) e^{-2\pi i(l-k)\omega}\mathrm{d}\omega = \delta_{k,l} = \delta_{l-k}, \text{ for all } k, l \in \mathbb{Z}.$$

This is the case if and only if the zeroth Fourier coefficient of the $L^2[0,1]$ function $\omega \mapsto \sum_{n\in\mathbb{Z}} |(\mathcal{F}g)(\omega+n)|^2$ is equal to 1, and all its other Fourier coefficients are equal to 0. But, by the injectivity of the map

$$L^2[0,1] \to \ell^2(\mathbb{Z}), \qquad f \mapsto \{\hat{f}_m\}_{m\in\mathbb{Z}},$$

this is the case if and only if $\omega \mapsto \sum_{n\in\mathbb{Z}} |(\mathcal{F}g)(\omega+n)|^2$ is the identically 1 function, i.e., if $\sum_{n\in\mathbb{Z}} |(\mathcal{F}g)(\omega+n)|^2 = 1$, for all $\omega \in [0,1]$. □

The previous result is very useful for checking the orthonormality translation invariant systems such as $\{\varphi(\cdot - k) : k \in Z\}$ and $\{\psi_{0,k}(\cdot - k) : k \in Z\}$. Also, the technique of rewriting $\int_{\mathbb{R}} = \sum_{n\in\mathbb{Z}} \int_n^{n+1}$ is useful in proofs of many results regarding MRAs.

2.1 Constructing a wavelet from the conjugate mirror filter of an MRA

Let $\{\mathcal{V}_j : j \in \mathbb{Z}\}$ be an MRA with scaling function φ, so that $\mathcal{B}_0^\varphi = \{\varphi(\cdot - k) : k \in \mathbb{Z}\}$ is an ONB for $\mathcal{V}_0$. For $j \in \mathbb{Z}$, let $\mathcal{W}_j$ be the orthogonal complement of $\mathcal{V}_j$ in $\mathcal{V}_{j+1}$, and furthermore suppose $\psi \in \mathcal{W}_0$ is such that $\mathcal{B}_0^\psi = \{\psi_{0,k} = \psi(\cdot - k) : k \in \mathbb{Z}\}$ is ONB for $\mathcal{W}_0$. Note that then $\mathcal{B}_{-1}^\varphi = \left\{\frac{1}{\sqrt{2}}\varphi\left(\frac{\cdot}{2} - k\right) : k \in \mathbb{Z}\right\}$ and $\mathcal{B}_{-1}^\psi = \left\{\frac{1}{\sqrt{2}}\psi\left(\frac{\cdot}{2} - k\right) : k \in \mathbb{Z}\right\}$ are ONBs for $\mathcal{V}_{-1}$ and $\mathcal{W}_{-1}$, both being subspaces of $\mathcal{V}_0$. Therefore the elements of $\mathcal{V}_{-1}$ can be expressed in terms of the $\mathcal{B}_0^\varphi$ by first expressing the elements of $\mathcal{B}_{-1}^\varphi$ in terms of $\mathcal{B}_0^\varphi$. Similarly, the elements of $\mathcal{W}_{-1}$ can be expressed in terms of $\mathcal{B}_0^\varphi$ by first expressing the elements of $\mathcal{B}_{-1}^\psi$ in terms of $\mathcal{B}_0^\varphi$.
In fact, due to the translationally invariant structure of $\mathcal{B}_{-1}^\varphi$ and $\mathcal{B}_{-1}^\psi$, it suffices to expand

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{\cdot}{2}\right) = \sum_{k\in\mathbb{Z}} h[k]\varphi(\cdot - k), \tag{4}$$

$$\frac{1}{\sqrt{2}}\psi\left(\frac{\cdot}{2}\right) = \sum_{k\in\mathbb{Z}} g[k]\varphi(\cdot - k), \tag{5}$$

where

$$h[k] = \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \varphi\left(\frac{t}{2}\right) \overline{\varphi(t-k)} \mathrm{d}t, \qquad \sum_{k \in \mathbb{Z}} |h[k]|^2 < \infty, \text{ and}$$

$$g[k] = \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \psi\left(\frac{t}{2}\right) \overline{\varphi(t-k)} \mathrm{d}t, \qquad \sum_{k \in \mathbb{Z}} |g[k]|^2 < \infty.$$

The sequence $\{h[k]\}_{k\in\mathbb{Z}}$ is called the *conjugate mirror filter*, and the sequence $\{g[k]\}_{k\in\mathbb{Z}}$ is called the *mirror filter*. By taking the L^2-Fourier transform of both sides of (4) and (5), we obtain

$$\sqrt{2}(\mathcal{F}\varphi)(2\omega) = \sum_{k \in \mathbb{Z}} h[k](\mathcal{F}\varphi)(\omega)e^{-2\pi i k\omega} =: \widehat{h}(\omega)(\mathcal{F}\varphi)(\omega), , \text{ and} \tag{6}$$

$$\sqrt{2}(\mathcal{F}\psi)(2\omega) = \sum_{k \in \mathbb{Z}} g[k](\mathcal{F}\varphi)(\omega)e^{-2\pi i k\omega} =: \widehat{g}(\omega)(\mathcal{F}\varphi)(\omega), \tag{7}$$

where $\widehat{h}$ and $\widehat{g}$ are the discrete-time Fourier transforms of h and g, given by

$$\widehat{h}(\omega) = \sum_{k \in \mathbb{Z}} h[k]e^{-2\pi i k\omega} \tag{8}$$

$$\widehat{g}(\omega) = \sum_{k \in \mathbb{Z}} g[k]e^{-2\pi i k\omega}. \tag{9}$$

As $\sum_{k\in\mathbb{Z}} |h[k]|^2 < \infty$ and $\sum_{k\in\mathbb{Z}} |g[k]|^2 < \infty$, the 1-periodic functions $\widehat{h}$ and $\widehat{g}$ can be considered functions in $L^2[0,1]$. Now, suppose we are able to construct an MRA $\{\mathcal{V}_j : j \in \mathbb{Z}\}$ together with its scaling function φ, but we do not know the corresponding ψ. Then, with just the knowledge of φ, we can define h according to (4), and then $\widehat{h}$ according to (8). The following theorem tells us how to obtain a wavelet ψ (and thus the corresponding mirror filter g, according to (5)), if we are given just a scaling function φ.

Theorem 3 (Constructing a wavelet from an MRA, [2, Prop.2.13])**.** Let $\{\mathcal{V}_j : j \in \mathbb{Z}\}$ be an MRA with scaling function φ and corresponding conjugate mirror filter h. Let ν be an arbitrary complex number such that $|\nu| = 1$. Define $\psi \in L^2(\mathbb{R})$ via its L^2-Fourier transform according to

$$(\mathcal{F}\psi)(\omega) = \nu e^{i\pi\omega} \overline{\widehat{h}\left(\frac{1+\omega}{2}\right)} \cdot \frac{1}{\sqrt{2}} (\mathcal{F}\varphi)(\omega/2), \tag{10}$$

for all $\omega \in \mathbb{R}$. Let $\mathcal{W}_0 = \overline{\operatorname{span}\{\psi_{0,k} : k \in \mathbb{Z}\}}$. Then,

(i) $\psi \in V_1$,

(ii) $\{\psi_{0,k} : k \in \mathbb{Z}\}$ is an orthonormal basis for $\mathcal{W}_0$,

(iii) The spaces $\mathcal{V}_0$ and $\mathcal{W}_0$ are orthogonal, and

(iv) $\mathcal{V}_1 = \mathcal{V}_0 \oplus \mathcal{W}_0$.

The role of the constant ν in Theorem 3 is to "fine tune" the wavelet ψ so that it is a real-valued function (recall that elements of $L^2(\mathbb{R})$ are generally *complex-valued* functions). Having a real-valued orthonormal wavelet ψ is important in practical applications, such as the Fast Wavelet Transform (FWT).

By plugging 2ω into (10) and multiplying both sides by $\sqrt{2}$, we obtain

$$\sqrt{2}(\mathcal{F}\psi)(2\omega) = \nu e^{2i\pi\omega} \overline{\widehat{h}\left(\frac{1}{2}+\omega\right)} \cdot (\mathcal{F}\varphi)(\omega)$$

Comparing this with (7), we see that the mirror filter g corresponding to the wavelet obtained according to (10) satisfies

$$\widehat{g}(\omega) = \nu e^{2i\pi\omega} \overline{\widehat{h}\left(\frac{1}{2}+\omega\right)}. \tag{11}$$

Examples.

1. (Haar wavelet) For each $j \in \mathbb{Z}$, let $\mathcal{V}_j$ be the space of $L^2(\mathbb{R})$ functions constant on intervals $\left[2^{-j}k, 2^{-j}(k+1)\right)$. Then $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ is an MRA with scaling function $\varphi = \mathbb{1}_{[0,1)}$. We then have

$$(\mathcal{F}\varphi)(\omega) = e^{-i\pi\omega} \frac{\sin(\pi\omega)}{\pi\omega}.$$

Therefore we can find $\widehat{h}$ from

$$\begin{aligned}\widehat{h}(\omega) &= \frac{\sqrt{2}(\mathcal{F}\varphi)(2\omega)}{(\mathcal{F}\varphi)(\omega)} = \frac{\sqrt{2}e^{-2i\pi\omega}\frac{\sin(2\pi\omega)}{2\pi\omega}}{e^{-i\pi\omega}\frac{\sin(\pi\omega)}{\pi\omega}} \\ &= \frac{\sqrt{2}}{2}e^{-i\pi\omega}\frac{2\sin(\pi\omega)\cos(\pi\omega)}{\sin(\pi\omega)} = \sqrt{2}e^{-i\pi\omega}\cos(\pi\omega),\end{aligned}$$

and then use (11) to find

$$\widehat{g}(\omega) = \nu e^{2i\pi\omega} \overline{\sqrt{2}e^{-i\pi\left(\frac{1}{2}+\omega\right)}\cos\left(\pi\left(\frac{1}{2}+\omega\right)\right)} = -i\nu\sqrt{2}e^{3i\pi\omega}\sin(\pi\omega).$$

The corresponding wavelet is given via its L^2-Fourier transform by

$$\begin{aligned}(\mathcal{F}\psi)(\omega) &= -i\nu\sqrt{2}e^{\frac{3i\pi\omega}{2}}\sin\left(\frac{\pi\omega}{2}\right)\cdot\frac{1}{\sqrt{2}}e^{-\frac{i\pi\omega}{2}}\frac{\sin(\pi\omega/2)}{\pi\omega/2} \\ &= -i\nu e^{i\pi\omega}\frac{\sin^2(\pi\omega/2)}{\pi\omega/2} \\ &= -i\nu e^{i\pi\omega}\frac{e^{\frac{i\pi\omega}{2}}-e^{-\frac{i\pi\omega}{2}}}{2i}\frac{\sin(\pi\omega/2)}{\pi\omega/2} \\ &= -\frac{\nu}{2}\left(e^{\frac{3i\pi\omega}{2}}\operatorname{sinc}(\omega/2) - e^{\frac{i\pi\omega}{2}}\operatorname{sinc}(\omega/2)\right)\end{aligned}$$

Taking the inverse Fourier transform yields

$$\psi(x) = -\frac{\nu}{2}\left(2\mathbb{1}_{[-\frac{1}{2},\frac{1}{2})}\left(2x+\frac{3}{2}\right) - 2\mathbb{1}_{[-\frac{1}{2},\frac{1}{2})}\left(2x+\frac{1}{2}\right)\right) = -\nu\left(\mathbb{1}_{[-1,-\frac{1}{2})} - \mathbb{1}_{[-\frac{1}{2},0)}\right)(x)$$

We choose $\nu = -1$ (note $|\nu| = |-1| = 1$). Then $\psi = \mathbb{1}_{[-1,-\frac{1}{2})} - \mathbb{1}_{[-\frac{1}{2},0)}$ we derived from the MRA $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ is the Haar wavelet considered at the beginning of this chapter.

2. (Shannon wavelet) For $j \in \mathbb{Z}$, let $\mathcal{V}_j$ be the set of functions in $L^2(\mathbb{R})$ whose L^2-Fourier transform has support contained in $[-2^{j-1}, 2^{j-1}]$ (i.e., the FT is zero everywhere outside this interval). Let $\varphi(x) = \frac{\sin(\pi x)}{\pi x}$. Then, by Proposition 2.1, $\{\varphi(\,\cdot - k) : k \in \mathbb{Z}\}$ is an ONB for $\mathcal{V}_0$. It follows that $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ is an MRA with scaling function $\varphi(x) = \frac{\sin(\pi x)}{\pi x}$. Then $(\mathcal{F}\varphi)(\omega) = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}$, and so the filter h satisfies
$$\sqrt{2}\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(2\omega) = \sqrt{2}\mathbb{1}_{[-\frac{1}{4},\frac{1}{4}]}(\omega) = \widehat{h}(\omega)\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(\omega).$$
As $\widehat{h}$ is a 1-periodic function, we find $\widehat{h}(\omega) = \sqrt{2}\,\mathbb{1}_{\bigcup_{k\in\mathbb{Z}}\left[-\frac{1}{4}+k,\frac{1}{4}+k\right]}(\omega)$. We now use (11) to find
$$\begin{aligned}\widehat{g}(\omega) &= \nu e^{2i\pi\omega}\sqrt{2}\,\mathbb{1}_{\bigcup_{k\in\mathbb{Z}}\left[-\frac{1}{4}+k,\frac{1}{4}+k\right]}\left(\frac{1}{2}+\omega\right)\\ &= \nu\sqrt{2}e^{2i\pi\omega}\,\mathbb{1}_{\bigcup_{k\in\mathbb{Z}}\left[\frac{1}{4}+k,\frac{3}{4}+k\right]}(\omega).\end{aligned}$$
The Fourier transform of the associated wavelet ψ is given by
$$\begin{aligned}(\mathcal{F}\psi)(\omega) &= \nu\sqrt{2}e^{i\pi\omega}\,\mathbb{1}_{\bigcup_{k\in\mathbb{Z}}\left[\frac{1}{4}+k,\frac{3}{4}+k\right]}\left(\frac{\omega}{2}\right)\cdot\frac{1}{\sqrt{2}}\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}\left(\frac{\omega}{2}\right)\\ &= \nu e^{i\pi\omega}\,\mathbb{1}_{[-\frac{1}{2},-\frac{1}{4}]\cup[\frac{1}{4},\frac{1}{2}]}\left(\frac{\omega}{2}\right)\\ &= \nu e^{i\pi\omega}\,\mathbb{1}_{[-1,-\frac{1}{2}]\cup[\frac{1}{2},1]}(\omega)\end{aligned}$$
Taking the inverse Fourier transform yields
$$\psi(x) = \nu\left(\mathcal{F}^{-1}(\mathbb{1}_{[-1,1]} - \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]})\right)\left(x+\frac{1}{2}\right) = \nu\frac{-\sin(2\pi x) - \cos(\pi x)}{\pi(x+\frac{1}{2})}.$$
We choose $\nu = 1$ to ensure ψ is real-valued. The resulting wavelet ψ is called the *Shannon wavelet.*

3. (Spline wavelets) Fix an integer $m \geqslant 0$. For $j \in \mathbb{Z}$ let $\mathcal{V}_j$ be the space of functions that are $m-1$ times continuously differentiable (where we take "-1 continuously differentiable" to mean "piecewise constant") and equal to a polynomial of degree m on the intervals $\left[2^{-j}k, 2^{-j}(k+1)\right)$, for $k \in \mathbb{Z}$. Let $\Delta^0 = \mathbb{1}_{[0,1]}$, $\Delta^1 = \mathbb{1}_{[0,1]} * \mathbb{1}_{[0,1]}$, and for $m \geqslant 2$:
$$\Delta^m = \Delta^{m-1} * \mathbb{1}_{[0,1]}.$$
One can show that the closed linear span of $\mathcal{B}_0^\Delta := \{\Delta^m(\,\cdot - k) : k \in \mathbb{Z}\} \subset \mathcal{V}_0$ is $\mathcal{V}_0$ — we assume this fact without proof. Unfortunately, the set $\mathcal{B}_0^\Delta$ does not form an ONB for $\mathcal{V}_0$, because it is not an orthonormal system. However, not all is lost, because we can "rebalance" Δ^m in the Fourier domain to yield a function φ such that $\{\varphi(\,\cdot - k) : k \in \mathbb{Z}\}$ is indeed an ONB for $\mathcal{V}_0$. We first derive
$$\widehat{\Delta^m}(\omega) = \left(\widehat{\mathbb{1}_{[0,1]}}(\omega)\right)^{m+1} = e^{-i\pi(m+1)}\left(\operatorname{sinc}(\pi\omega)\right)^{m+1}$$
Specifically, we define φ^m via its L^2-Fourier transform according to $(\mathcal{F}\varphi^m)(\omega) = \widehat{\Delta^m}(\omega)/\sigma^m(\omega)$, where σ^m is given by
$$\sigma^m(\omega) = \sqrt{\sum_{k\in\mathbb{Z}}\left|\widehat{\Delta^m}(\omega+k)\right|^2}.$$

It is possible to compute σ^m explicitly, and thus show that it is a 1-periodic function bounded from both below and above by positive real numbers. It follows that φ^m is, indeed, an L^2 function. We then have

$$\sum_{k\in\mathbb{Z}} |(\mathcal{F}\varphi^m)(\omega+k)|^2 = \frac{1}{|\sigma^m(\omega)|^2} \sum_{k\in\mathbb{Z}} |\widehat{\Delta^m}(\omega+k)|^2 = 1,$$

for all $\omega \in \mathbb{R}$, and so by Proposition 2.1, $\mathcal{B}_0^\varphi := \{\varphi^m(\,\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal system. To show that it is an ONB of $\mathcal{V}_0$, it remains to show that $[\mathcal{B}_0^\varphi] = \mathcal{V}_0 = [\mathcal{B}_0^\Delta]$. To this end, write $\sigma^m(\omega) = \sum_{k\in\mathbb{Z}} \alpha_k\, e^{2\pi i k\omega}$ as a Fourier series, where $\{\alpha_k\}_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$. Then we can write

$$\widehat{\Delta^m}(\omega) = \sigma^m(\omega)(\mathcal{F}\varphi^m)(\omega) = \sum_{k\in\mathbb{Z}} \alpha_k\, e^{2\pi i k\omega}(\mathcal{F}\varphi^m)(\omega).$$

Now, multiply both sides by $e^{-2\pi i r\omega}$, for an arbitrary $r \in \mathbb{Z}$ to get

$$e^{-2\pi i r\omega}\widehat{\Delta^m}(\omega) = \sum_{k\in\mathbb{Z}} \alpha_k\, e^{2\pi i(k-r)\omega}(\mathcal{F}\varphi^m)(\omega). \tag{12}$$

We recognize the right-hand side of (12) as the L^2-Fourier transform of $\sum_{k\in\mathbb{Z}} \alpha_k\, \varphi^m(\,\cdot + k - r)$ (which is an unconditionally convergent sum, due to the fact that $\{\alpha_k\}_{k\in\mathbb{Z}}$ is square-summable, and $\mathcal{B}_0$ is an orthonormal system), and the left-hand side of (12) is simply the L^2-Fourier transform of $\Delta^m(\,\cdot - r)$. Therefore, by the injectivity of the L^2-Fourier transform, we have

$$\Delta^m(\,\cdot - r) = \sum_{k\in\mathbb{Z}} \alpha_k\, \varphi^m(\,\cdot + k - r),$$

where the sum on the right-hand side converges unconditionally in $L^2(\mathbb{R})$. As $[\mathcal{B}_0^\varphi]$ is a linear space, and $r \in \mathbb{Z}$ was arbitrary, we obtain $\operatorname{span}\{\Delta^m(\,\cdot - r) : r \in \mathbb{Z}\} \subset [\mathcal{B}_0^\varphi]$. As $[\mathcal{B}_0^\varphi]$ is closed, we have $\overline{\operatorname{span}\{\Delta^m(\,\cdot - r) : r \in \mathbb{Z}\}} \subset [\mathcal{B}_0^\varphi]$, and therefore $[\mathcal{B}_0^\Delta] \subset [\mathcal{B}_0^\varphi]$.

By considering the Fourier series $1/\sigma^m(\omega) = \sum_{k\in\mathbb{Z}} \beta_k\, e^{2\pi i k\omega}$ of the function $1/\sigma^m$, we can carry out a similar derivation to show $[\mathcal{B}_0^\varphi] \subset [\mathcal{B}_0^\Delta]$. Therefore $\mathcal{B}_0^\varphi$ is, indeed, an ONB for $\mathcal{V}_0$, and so $\{\mathcal{V}_j\}_{j\in\mathbb{Z}}$ is an MRA with scaling function φ^m. The function Δ^m is called the *cardinal B-spline wavelet of order* m, and the corresponding orthonormal wavelet ψ^m is called the *Battle-Lemarie wavelet of order* m.

3 Fast Wavelet Transform

Suppose ψ is an orthonormal wavelet derived from an MRA $\{V_j\}_{j\in\mathbb{Z}}$ with scaling function φ. By performing a change of variable $t \leftarrow 2(2^j t - n)$ in (4) and (5), we get

$$\frac{1}{\sqrt{2}}\varphi(2^j \cdot -n) = \sum_{k\in\mathbb{Z}} h[k]\varphi(2^{j+1} \cdot -(k+2n))$$
$$\frac{1}{\sqrt{2}}\psi(2^j \cdot -n) = \sum_{k\in\mathbb{Z}} g[k]\psi(2^{j+1}t - (k+2n))$$

A further change of variables $n \leftarrow n - 2k$ and multiplication by $2^{(j+1)/2}$ yields

$$\varphi_{j,n} = \sum_{k\in\mathbb{Z}} h[k-2n]\varphi_{j+1,k} \tag{13}$$
$$\psi_{j,n} = \sum_{k\in\mathbb{Z}} g[k-2n]\varphi_{j+1,k}, \tag{14}$$

for all $n, j \in \mathbb{Z}$, where $\varphi_{j,k} = 2^{j/2}\varphi(2^j \cdot -k)$ and $\psi_{j,k} = 2^{j/2}\psi(2^j \cdot -k)$. Fix a $j \in \mathbb{Z}$ and a function $f \in V_{j+1}$. We then have the following expansion of f in the ONB $\{\varphi_{j+1,k} : k \in \mathbb{Z}\}$ for V_{j+1}:

$$f = \sum_{n\in\mathbb{Z}} a_{j+1}[n]\varphi_{j+1,n}. \tag{15}$$

But, as $V_{j+1} = V_j \oplus W_j$ is an orthogonal direct sum, the set $\{\varphi_{j,k} : k \in \mathbb{Z}\} \cup \{\psi_{j,k} : k \in \mathbb{Z}\}$ is also an ONB for V_{j+1}, and so we have

$$f = P_{V_j} f + P_{W_j} f = \sum_{k\in\mathbb{Z}} a_j[k]\varphi_{j,k} + \sum_{k\in\mathbb{Z}} w_j[k]\psi_{j,k}. \tag{16}$$

Using (13) and (14), we can express the expansion coefficients $\{a_j[k]\}_{k\in\mathbb{Z}}$ and $\{w_j[k]\}_{k\in\mathbb{Z}}$ in terms of $\{a_{j+1}[n]\}_{n\in\mathbb{Z}}$, as follows:

$$\begin{aligned}
a_j[k] = \langle f, \varphi_{j,k}\rangle &= \left\langle \sum_{n\in\mathbb{Z}} a_{j+1}[n]\varphi_{j+1,n}, \sum_{\ell\in\mathbb{Z}} h[\ell-2k]\varphi_{j+1,\ell} \right\rangle \\
&= \sum_{n\in\mathbb{Z}}\sum_{\ell\in\mathbb{Z}} a_{j+1}[n]\overline{h[\ell-2k]}\, \underbrace{\langle \varphi_{j+1,n}, \varphi_{j+1,\ell}\rangle}_{=\delta_{n,\ell}} = \sum_{n\in\mathbb{Z}} a_{j+1}[n]\overline{h[n-2k]}, \\
w_j[k] = \langle f, \psi_{j,k}\rangle &= \left\langle \sum_{n\in\mathbb{Z}} a_{j+1}[n]\varphi_{j+1,n}, \sum_{\ell\in\mathbb{Z}} g[\ell-2k]\varphi_{j+1,\ell} \right\rangle \\
&= \sum_{n\in\mathbb{Z}}\sum_{\ell\in\mathbb{Z}} a_{j+1}[n]\overline{g[\ell-2k]}\, \underbrace{\langle \varphi_{j+1,n}, \varphi_{j+1,\ell}\rangle}_{=\delta_{n,\ell}} = \sum_{n\in\mathbb{Z}} a_{j+1}[n]\overline{g[n-2k]},
\end{aligned}$$

Therefore, we have obtained that the "coarse" approximation coefficients $\{a_j[k]\}_{k\in\mathbb{Z}}$ and the wavelet coefficients $\{w_j[k]\}_{k\in\mathbb{Z}}$ can be computed from the "fine" approximation coefficients $\{a_{j+1}[n]\}_{n\in\mathbb{Z}}$ recursively using

$$a_j[k] = \sum_{n\in\mathbb{Z}} \overline{h[n-2k]}\, a_{j+1}[n] \tag{17}$$
$$w_j[k] = \sum_{n\in\mathbb{Z}} \overline{g[n-2k]}\, a_{j+1}[n]. \tag{18}$$

These relations correspond to low and high pass filtering operations followed by a subsampling step:

$$\begin{aligned} \mathbf{a}_j &= (\mathbf{a}_{j+1} * \widetilde{\mathbf{h}}) \downarrow 2 \\ \mathbf{w}_j &= (\mathbf{a}_{j+1} * \widetilde{\mathbf{g}}) \downarrow 2, \end{aligned} \tag{19}$$

where $\widetilde{\mathbf{h}}$ and $\widetilde{\mathbf{g}}$ are the filters defined by $\widetilde{h}[k] = \overline{h[-k]}$, $\widetilde{g}[k] = \overline{g[-k]}$, and $\downarrow 2$ designates the downsampling operation, i.e., $(x \downarrow 2)[n] = x[2n]$. Intuitively, this corresponds to breaking down the approximation coefficients $\mathbf{a}_{j+1}$ at a higher resolution $2^{-(j+1)}$ into approximation coefficients $\mathbf{a}_j$ at a lower resolution 2^{-j}, plus the wavelet coefficients $\mathbf{w}_j$ making up for the loss of detail when going from the higher resolution $2^{-(j+1)}$ to the lower resolution 2^{-j}. Therefore the wavelet coefficients $\mathbf{w}_j$ can also be interpreted as *detail* coefficients.

Conversely, the coarse approximation coefficients $\{a_j[k]\}_{k\in\mathbb{Z}}$ and the detail coefficients $\{w_j[k]\}_{k\in\mathbb{Z}}$ can be used to reconstruct the fine approximation coefficients $\{a_{j+1}[n]\}_{n\in\mathbb{Z}}$. To see this, we begin by substituting the identities (13) and (14) into (16):

$$\begin{aligned} f &= \sum_{k\in\mathbb{Z}} a_j[k] \sum_{n\in\mathbb{Z}} h[n-2k]\varphi_{j+1,n} + \sum_{k\in\mathbb{Z}} w_j[k] \sum_{n\in\mathbb{Z}} g[n-2k]\varphi_{j+1,n} \\ &= \sum_{n\in\mathbb{Z}} \left[\sum_{k\in\mathbb{Z}} a_j[k]h[n-2k] + \sum_{k\in\mathbb{Z}} w_j[k]g[n-2k] \right] \varphi_{j+1,n}. \end{aligned}$$

Comparing this expression with (15) to obtain and using the fact that the expansion coefficients of a function in an ONB are unique, we deduce

$$a_{j+1}[n] = \sum_{k\in\mathbb{Z}} a_j[k]h[n-2k] + \sum_{k\in\mathbb{Z}} w_j[k]g[n-2k].$$

We thus have the following reconstruction formula

$$\mathbf{a}_{j+1} = (\mathbf{a}_j \uparrow 2) * \mathbf{h} + (\mathbf{w}_j \uparrow 2) * \mathbf{g}, \tag{20}$$

where the upsampling $\uparrow 2$ operator is given by:

$$(x \uparrow 2)[n] = \begin{cases} x[p], & n = 2p, \\ 0, & n = 2p+1. \end{cases}$$

The relations (19) and (20) are called the *forward wavelet transform* and the *inverse wavelet transform*, respectively.

3.1 Discrete Wavelet Transform

We wish to apply the wavelet construction to the case of finite-dimensional discrete signals $\mathbf{x} \in \mathbb{R}^N$. For the sake of simplicity, assume that $N = 2^J$. Given a signal

$$\mathbf{x} = \begin{bmatrix} x[0] & x[1] & \dots & x[2^J-1] \end{bmatrix}^T,$$

we interpret the values $x[n]$ as dyadic averages over a uniform grid on $[0,1]$ of mesh size 2^{-J} of some function $f \in L^2(\mathbb{R})$. That is,

$$f = \sum_{k=0}^{2^J-1} x[k]\, \varphi_{J,k} \quad \in \mathcal{V}_J.$$

Furthermore, assume φ and ψ have compact supports with

$$\operatorname{supp}(\varphi) \subset [-S, T], \quad \text{and} \quad \operatorname{supp}(\psi) \subset \left[-\frac{S+T+1}{2}, \frac{S+T-1}{2}\right],$$

for some $S, T \in \mathbb{N}$ (this can be achieved by taking φ to be a Daubechies scaling function). Then $\operatorname{supp}(f) \subset [-S2^{-J}, 1-2^{-J}+T2^{-J}]$, and f thus constructed can now be decomposed into a coarse part in $\mathcal{V}_{J-1}$ and a detail part in $\mathcal{W}_{J-1}$, which contains the error made when going from fine to coarse:

$$f = \sum_{k=-OH(J-1)}^{2^{J-1}-1+OH(J-1)} a_{J-1}[k]\,\varphi_{J-1,k} \quad + \sum_{k=-OH(J-1)}^{2^{J-1}-1+OH(J-1)} w_{J-1}[k]\,\psi_{J-1,k},$$

where the sequences $\mathbf{a}_{J-1} = \{a_{J-1}[k]\}_k$ and $\mathbf{w}_{J-1} = \{w_{J-1}[k]\}_k$ are given by

$$\begin{aligned} \mathbf{a}_{J-1} &= (\mathbf{a}_J * \widetilde{\mathbf{h}}) \downarrow 2 \\ \mathbf{w}_{J-1} &= (\mathbf{a}_J * \widetilde{\mathbf{g}}) \downarrow 2, \end{aligned}$$

with $\mathbf{a}_J = \mathbf{x} = \{x[k]\}_{k\in\{0,\dots,2^J-1\}}$, $\widetilde{\mathbf{h}} = \{h[-k]\}_{k\in\mathbb{Z}}$, $\widetilde{\mathbf{g}} = \{g[-k]\}_{k\in\mathbb{Z}}$, and $OH(j) = \lceil \operatorname{const}(S,T) \cdot 2^j \rceil$ is the number of "overhead coefficients" at level j. In the discrete world, we have decomposed the sequence $\mathbf{a}_J = \mathbf{x}$ into two sequences $\mathbf{a}_{J-1}$ and $\mathbf{w}_{J-1}$, which contains respectively the "coarse" information and the "fine" information at resolution 2^{-J}. We can repeat this procedure and decompose $\mathbf{a}_{J-1}$ into $\mathbf{a}_{J-2}$ and $\mathbf{w}_{J-2}$, and so on. In this end, we obtain the decomposition $\begin{bmatrix}\mathbf{a}_0^T & \mathbf{w}_0^T & \mathbf{w}_1^T & \dots & \mathbf{w}_{J-1}^T\end{bmatrix}^T$. The map

$$\mathbf{a}_J \to \begin{bmatrix}\mathbf{a}_0^T & \mathbf{w}_0^T & \mathbf{w}_1^T & \dots & \mathbf{w}_{J-1}^T\end{bmatrix}^T$$

is called the Discrete Wavelet Transform (DWT). When φ and ψ have compact supports (as we assumed for this discussion), the DWT and its inverse can be implemented in $\mathcal{O}\big(N \cdot (S+T)\big)$. Thus, the complexity of the FWT is linear in the size of the signal to be transformed. Compare this with the complexity $\mathcal{O}(N \log N)$ of the Fast Fourier Transform. The decomposition and reconstruction algorithms are summarized as Algorithms 1 and 2, respectively.

4 Vanishing moments

Most applications of wavelet bases exploit their ability to approximate particular classes of functions with few nonzero wavelet coefficients. Therefore, we would like a wavelet ψ such that most of the coefficients $\langle f, \psi_{j,k} \rangle$ of a function f are close to zero. Functions of certain degree of smoothness, such as $f \in C^n(\mathbb{R})$, for some $n \in \mathbb{N}$, can be locally approximated by their Taylor polynomial around a point $x_0 \in \mathbb{R}$ of order n:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o\big((x-x_0)^n\big) \tag{21}$$

Therefore, if $j, k \in \mathbb{Z}$ are such that $\psi_{j,k}$ is concentrated around x_0, then

$$\langle f, \psi_{j,k} \rangle = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} \langle (\,\cdot - x_0)^k, \psi \rangle + o\big(|\operatorname{supp}(\psi_{j,k})|^n\big). \tag{22}$$

We see that a desireable property for ψ would be that its inner product with polynomials of degree $k \leqslant n$ is zero.

Algorithm 1: Wavelet decomposition

Input : $N = 2^J$, the signal length,
$\mathbf{x} \in \mathbb{R}^N$, the signal to be decomposed,
$\mathbf{g} = \{\mathbf{g}[n]\}_{n\in\mathbb{Z}}$, the mirror filter,
$\mathbf{h} = \{\mathbf{h}[n]\}_{n\in\mathbb{Z}}$, the conjugate mirror filter.

Output : $\mathbf{c} \in \mathbb{R}^N$, the wavelet decomposition.

Initialize $\mathbf{a}_J = \mathbf{x}$.
for $j = J, J-1, \ldots, 2, 1$ **do**
    Compute the approximation coefficients $\mathbf{a}_{j-1} = (\mathbf{a}_j * \widetilde{\mathbf{h}}) \downarrow 2$, where $\widetilde{\mathbf{h}}$ is the filter defined by $\widetilde{h}[n] = h[-n]$ for all $n \in \mathbb{Z}$.
    Compute the wavelet coefficients $\mathbf{w}_{j-1} = (\mathbf{a}_j * \widetilde{\mathbf{g}}) \downarrow 2$, where $\widetilde{\mathbf{g}}$ is the filter defined by $\widetilde{g}[n] = g[-n]$ for all $n \in \mathbb{Z}$.
end for
return $\mathbf{c} = \begin{bmatrix} \mathbf{a}_0^T & \mathbf{w}_0^T & \mathbf{w}_1^T & \ldots & \mathbf{w}_{J-1}^T \end{bmatrix}^T$

Algorithm 2: Wavelet reconstruction

Input : $N = 2^J$, the signal length,
$\mathbf{c} = \begin{bmatrix} \mathbf{a}_0^T & \mathbf{w}_0^T & \mathbf{w}_1^T & \ldots & \mathbf{w}_{J-1}^T \end{bmatrix}^T \in \mathbb{R}^N$, the wavelet decomposition,
$\mathbf{g} = \{\mathbf{g}[n]\}_{n\in\mathbb{Z}}$, the mirror filter,
$\mathbf{h} = \{\mathbf{h}[n]\}_{n\in\mathbb{Z}}$, the conjugate mirror filter.

Output : $\mathbf{x} \in \mathbb{R}^N$, the reconstructed signal.

Initialize $\mathbf{a}_J = \mathbf{x}$.
for $j = 0, 1, \ldots, J-1$ **do**
    Compute $\mathbf{a}_{j+1} = (\mathbf{a}_j \uparrow 2) * \mathbf{h} + (\mathbf{w}_j \uparrow 2) * \mathbf{g}$.
end for
return $\mathbf{x} = \mathbf{a}_J$

Definition 3. We say that a real wavelet ψ has p vanishing moments if $\int_{\mathbb{R}} x^k \psi(x) \mathrm{d}x = 0$, for all $k \in \{0, 1, \dots, p-1\}$.

We continue our discussion with the additional assumption that ψ is a real wavelet with $n+1$ vanishing moments. Now, (22) reads

$$|\langle f, \psi_{j,k} \rangle| = o\big(|\mathrm{supp}(\psi_{j,k})|^n\big) \ll \|f\|_{L^2(\mathbb{R})} |\mathrm{supp}(\psi_{j,k})|,$$

for $j \in \mathbb{Z}$ large enough s.t. $|\mathrm{supp}(\psi_{j,k})| \ll \|f\|_{L^2(\mathbb{R})}$. Therefore, if f is approximated well by a polynomial of degree n around a point x_0, then the coefficients $\langle f, \psi_{j,k} \rangle$ around x_0 at fine enough scales 2^{-j} will be much smaller then the "typical magnitude" $\|f\|_{L^2(\mathbb{R})} |\mathrm{supp}(\psi_{j,k})|$ of a coefficient at scale 2^{-j}. This means that, as the resolution 2^{-j} of the approximation increases (i.e., $j \to \infty$), the coefficients $\langle f, \psi_{j,k} \rangle$ will be mostly very close to zero, spiking up only around the points x_0 where (21) is violated. The coefficients that are very close to zero can then be discarded, and the remaining coefficients (provided there are not too many points x_0 where (21) is violated) will constitute a good compressed representation of f.

The following proposition tells us that vanishing moments are typically a consequence of smoothness and decay properties of the wavelet ψ.

Theorem 4 (Vanishing moments, [2, Thm.3.4])**.** Suppose ψ is an orthonormal wavelet. For $n \in \mathbb{N}$ assume that

(i) $\psi \in C^n(\mathbb{R})$ (n times continuously differentiable),

(ii) $\psi^{(s)}$ is a bounded funtion on $\mathbb{R}$, for $s \in \{0, 1, \dots, n\}$, and

(iii) There exist constants $c > 0$ and $\alpha > n+1$ such that $|\psi(x)| \leqslant \frac{c}{(1+|x|)^\alpha}$, for all $x \in \mathbb{R}$.

Then, ψ has $n+1$ vanishing moments.

The following theorem establishes the existence of a class of wavelets with all the desireable properties we have discussed, namely real-valued, compactly supported, and with vanishing moments.

Theorem 5 (Daubechies wavelets [2, Section 2.3])**.** Let $2A$ be a positive even integer. Then there exists an MRA with a scaling function φ and wavelet ψ, called the *Daubechies wavelet with A vanishing moments* with the following properties:

(i) φ and ψ are real valued,

(ii) φ is supported on $[-(2A-1), 0]$ and ψ is supported on $[-A, A-1]$, and

(iii) ψ has A vanishing moments.

The wavelet ψ in the previous theorem is often referred to as either $D2A$ or dbA, thus $D4$ is the same wavelet as $db2$, $D6$ is the same wavelet as $db3$, and so on.

We now formalize the above discussion and derive nonlinear approximation rates for Daubechies wavelets.

Definition 4. Let $f : [a, b] \to \mathbb{R}$ be a function and $\alpha > 0$. We say that f is uniformly Lipschitz-α if there exists a $K > 0$ and polynomials p_v of degree $\lfloor \alpha \rfloor$, for $v \in [a, b]$, such that

$$|f(t) - p_v(t)| \leqslant K|t - v|^\alpha, \quad \text{for } t, v \in [a, b].$$

The smallest such value of K is called the *homogeneous Hölder-α norm* $\|f\|_{\tilde{C}^\alpha}$. The *Hölder-α norm* is then given by

$$\|f\|_{C^\alpha} = \|f\|_{\tilde{C}^\alpha} + \|f\|_{L^\infty}.$$

Examples.

1. Let $\varphi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be the scaling function of an MRA, and suppose that $\int_{\mathbb{R}} |x||\varphi(x)|\mathrm{d}x < \infty$. Consider a function $f \in C^1(\mathbb{R})$ such that $\|f'\|_{L^\infty(\mathbb{R})} < \infty$ and a point $t_0 \in \mathbb{R}$. We show that, in the limit as $j \to \infty$ and $2^{-j}k \to t_0$, we have $\langle f, 2^{\frac{j}{2}}\varphi_{j,k}\rangle \to f(t_0)$, allowing us to argue that $2^{\frac{j}{2}}\varphi_{j,k} \approx \delta(\cdot - 2^{-j}k)$ for large positive j, and so we can interpret the right-hand side of(15) as approximating

$$2^{-\frac{j+1}{2}} \sum_{n\in\mathbb{Z}} f(2^{-(j+1)}n)\,\delta(\cdot - 2^{-(j+1)}n),$$

i.e., a discretization of f.

To this end, we first observe that, by Theorem 2, we have $\int_{\mathbb{R}} \varphi(x)\mathrm{d}x = \widehat{\varphi}(0) = 1$. We now use the mean value theorem to calculate as follows:

$$\begin{aligned}
|\langle f, 2^{\frac{j}{2}}\varphi_{j,k}\rangle - f(t_0)| &= \left|\int_{\mathbb{R}} f(t)2^j\varphi(2^j(t-2^{-j}k))\mathrm{d}t - f(t_0)\right| \\
&\overset{x=2^jt-n}{=} \left|\int_{\mathbb{R}} f(2^{-j}x+2^{-j}k)\varphi(x)\mathrm{d}t - f(t_0)\int_{\mathbb{R}}\varphi(x)\mathrm{d}x\right| \\
&\leqslant \int_{\mathbb{R}} \left|f(2^{-j}x+2^{-j}k) - f(t_0)\right| |\varphi(x)|\,\mathrm{d}x \\
&\overset{\text{MVT}}{\leqslant} \int_{\mathbb{R}} \|f'\|_{L^\infty(\mathbb{R})} \left|2^{-j}x + 2^{-j}k - t_0\right| |\varphi(x)|\,\mathrm{d}x \\
&= \|f'\|_{L^\infty(\mathbb{R})}\left(\underbrace{2^{-j}}_{\to 0}\underbrace{\int_{\mathbb{R}}|x||\varphi(x)|\mathrm{d}x}_{<\infty} + \underbrace{\left|2^{-j}k - t_0\right|}_{\to 0}\underbrace{\|\varphi\|_{L^1(\mathbb{R})}}_{<\infty}\right) \\
&\longrightarrow 0,
\end{aligned}$$

as desired.

2. Let ψ be an orthonormal wavelet. Then the coefficients of a bounded function $f : [a,b] \to \mathbb{R}$ in the wavelet basis $\{\psi_{j,k} : j,k \in \mathbb{Z}\}$ decay exponentially with wavelet scale j at rate $2^{-j/2}$. Indeed, for all $j,k \in \mathbb{Z}$, we have

$$\begin{aligned}
|\langle f, \psi_{j,k}\rangle| &= \left|\int_a^b f(t)\cdot 2^{j/2}\psi(2^jt-k)\mathrm{d}t\right| \\
&\overset{x=2^jt-k}{\leqslant} \int_{2^ja-k}^{2^jb-k} \left|f\left(\frac{x+k}{2^j}\right)\right| 2^{-j/2}|\psi(x)|\mathrm{d}x \\
&\leqslant \|\psi\|_{L^1}\|f\|_{L^\infty}2^{-j/2}
\end{aligned}$$

3. Let ψ be an orthonormal wavelet with q vanishing moments supported on $[-q, q-1]$, let $[a,b] \subset \mathbb{R}$ be an interval, and let $\alpha > 0$ be such that $\lfloor\alpha\rfloor < q$. We show that

the coefficients of a uniformly Lipschitz-α function $f : [a, b] \to \mathbb{R}$ in the wavelet basis $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ decay exponentially with wavelet scale at rate $2^{-j(\alpha+1/2)}$. Concretely, we show that there exists a $B > 0$ such that, for all $j \in \mathbb{N}$ and $k \in \mathbb{Z}$ with $\mathrm{supp}(\psi_{j,k}) \subset [a, b]$, we have

$$|\langle f, \psi_{j,k}\rangle| \leqslant B\|f\|_{\tilde{C}^\alpha} 2^{-j(\alpha+1/2)}.$$

To this end, note that $2^{-j}k \in \mathrm{supp}(\psi_{j,k}) \subset [a, b]$, and so there exists a polynomial $p_{2^{-j}k}$ of degree $\lfloor \alpha \rfloor$ such that

$$|f(t) - p_{2^{-j}k}(t)| \leqslant \|f\|_{\tilde{C}^\alpha} |t - 2^{-j}k|^\alpha, \quad \text{for } t \in [a, b].$$

Now, as ψ has $q > \lfloor \alpha \rfloor$ vanishing moments, we have $\langle \psi_{j,k}, p_{2^{-j}k}\rangle = 0$, and so

$$\begin{aligned}
|\langle f, \psi_{j,k}\rangle| &= |\langle f - p_{2^{-j}k}, \psi_{j,k}\rangle| \\
&= \left| \int_a^b (f(t) - p_{2^{-j}k}(t))\, \overline{\psi_{j,k}(t)}\, \mathrm{d}t \right| \\
&\leqslant \|f\|_{\tilde{C}^\alpha} \int_a^b \left|t - 2^{-j}k\right|^\alpha \cdot 2^{j/2} \left|\psi(2^j t - k)\right| \mathrm{d}t \\
&\overset{x = 2^j t - k}{=} \|f\|_{\tilde{C}^\alpha} \int_{2^j a - k}^{2^j b - k} \left|2^{-j}x\right|^\alpha \cdot 2^{j/2} \left|\psi(x)\right| \mathrm{d}x \\
&\leqslant 2^{-j(\alpha+1/2)} \|f\|_{\tilde{C}^\alpha} \underbrace{\int_{\mathrm{supp}(\psi)} |x|^\alpha \left|\psi(x)\right| \mathrm{d}x}_{B=}.
\end{aligned}$$

5 Nonlinear approximation

Definition 5. Let $\mathcal{B} = \{g_m\}_{m=1}^\infty$ be an *ONB* for a Hilbert space $\mathcal{H}$, and let $f \in \mathcal{H}$. We write $f_\mathcal{B}[k] = \langle f, g_{m_k}\rangle$ for the coefficient of rank k:

$$|f_\mathcal{B}[k]| \geqslant |f_\mathcal{B}[k+1]|, \quad k \geqslant 1.$$

For $M \in \mathbb{N}$, the *nonlinear approximation* and *nonlinear approximation error* are given respectively by

$$f_M = \sum_{k=1}^M f_\mathcal{B}[k] g_{m_k}, \quad \text{and} \quad \epsilon_n(M, f) = \sum_{k=M+1}^\infty |f_\mathcal{B}[k]|^2.$$

We state the following general theorem proved in lectures.

Theorem 6. Let $\mathcal{B}$ and f be as in Definition 5, and suppose that $r > 1/2$ and $C > 0$ are such that $|f_\mathcal{B}[k]| \leqslant Ck^{-r}$, for all $k \geqslant 1$. Then

$$\epsilon_n(M, f) \leqslant \frac{C^2}{2r-1} M^{-(2r-1)}, \quad \text{for all } M \in \mathbb{N}.$$

We are now ready to derive the nonlinear approximation rate of piecewise regular functions in orthonormal bases of Daubechies wavelets.

Theorem 7. Let ψ be the Daubechies wavelet with q vanishing moments with corresponding scaling function φ, and set $\mathcal{B} = \{\varphi_{0,k} : k \in \mathbb{Z}\} \cup \{\psi_{j,k} : j \geqslant 0, k \in \mathbb{Z}\}$. Furthermore, let $\alpha \in (1/2, q)$, $\Delta = \{0 = a_1 < a_2 < \cdots < a_K = 1\} \subset \mathbb{R}$, and let $f : [0,1] \to \mathbb{R}$ be uniformly Lipschitz-α on $[a_\ell, a_{\ell+1}]$, for $\ell \in [K-1]$. Then, there exists a constant B depending on φ, ψ, α, and K such that the nonlinear approximation error of f in the basis $\mathcal{B}$ satisfies

$$\epsilon_n(M, f) \leqslant B \sup_{\ell \in [K-1]} \|f\|_{C^\alpha([a_\ell, a_{\ell+1}])} \cdot M^{-2\alpha}.$$

Proof. For $j \in \mathbb{Z}$, partition the indices of wavelets at scale j into

$$\begin{aligned} \mathrm{I}_j &= \{k \in \mathbb{Z} : \mathrm{supp}(\psi_{j,k}) \cap \Delta \neq \varnothing\}, \quad \text{and} \\ \mathrm{II}_j &= \{k \in \mathbb{Z} : \mathrm{supp}(\psi_{j,k}) \cap \Delta = \varnothing\}. \end{aligned}$$

We call $\psi_{j,k} \in \mathrm{I}_j$ and $\psi_{j,k} \in \mathrm{II}_j$ wavelets of type I and type II, respectively. Furthermore, write $C = 2q - 1 = |\,\mathrm{supp}(\varphi)| = |\,\mathrm{supp}(\psi)|$ and $F = \sup_{\ell \in [K-1]} \|f\|_{C^\alpha([a_\ell, a_{\ell+1}])}$. By the examples above, we have

$$|\langle f, \psi_{j,k}\rangle| \leqslant B_1 F\, 2^{-j/2}, \qquad \text{for } k \in \mathrm{I}_j, \text{ and} \tag{23}$$

$$|\langle f, \psi_{j,k}\rangle| \leqslant B_2 F\, 2^{-j(\alpha+1/2)}, \qquad \text{for } k \in \mathrm{II}_j, \tag{24}$$

where $B_1 = \|\psi\|_{L^1}$ and $B_2 = \int_{\mathrm{supp}(\psi)} |x|^\alpha\, |\psi(x)|\, \mathrm{d}x$. Moreover, we have

$$\begin{aligned} \#(\mathrm{I}_j) &\leqslant K \cdot |\,\mathrm{supp}(\psi_{j,k})|/2^{-j} = KC, \text{ and} \\ \#(\mathrm{II}_j) &\leqslant 2^j + 2C, \end{aligned}$$

for all $j \in \mathbb{Z}$. Write $f_{\mathcal{B},\mathrm{I}}[r] = \langle f, \psi_{j_r^{\mathrm{I}}, k_r^{\mathrm{I}}}\rangle$ for the r^{th} largest coefficient in magnitude among $\{\langle f, \psi_{j,k}\rangle : j \geqslant 0, k \in \mathrm{I}_j\}$, and similarly define $f_{\mathcal{B},\mathrm{II}}[r] = \langle f, \psi_{j_r^{\mathrm{II}}, k_r^{\mathrm{II}}}\rangle$ as the r^{th} largest coefficient in magnitude among $\{\langle f, \psi_{j,k}\rangle : j \geqslant 0, k \in \mathrm{II}_j\}$.

As there are at most jKC wavelets of type I at scales $\{0, 1, \ldots, j-1\}$, and at scales $\ell \geqslant j$ we have the estimate (23), we find that

$$f_{\mathcal{B},\mathrm{I}}[jKC] \leqslant B_1 F\, 2^{-j/2},$$

and so

$$f_{\mathcal{B},\mathrm{I}}[m] \leqslant B_1 F\, 2^{-\lfloor \frac{m}{KC} \rfloor/2} \leqslant \underbrace{2^{1/2} B_1}_{B_3=} F\, 2^{-\frac{m}{2KC}}, \quad \text{for } m \in \mathbb{N}. \tag{25}$$

On the other hand, as there are at most $(2^{j-1} + 2C) + (2^{j-2} + 2C) + \cdots + (1 + 2C) = 2^j - 1 + 2jC < 2^j(1 + C)$ wavelets of type II at scales $\{0, 1, \ldots, j-1\}$, and at scales $\ell \geqslant j$ we have (24), we find that

$$f_{\mathcal{B},\mathrm{II}}[2^j(1 + C)] \leqslant B_2 F\, 2^{-j(\alpha+1/2)},$$

and so

$$\begin{aligned} f_{\mathcal{B},\mathrm{II}}[m] &\leqslant B_2 F\, 2^{-\left\lfloor \log_2\left(\frac{m}{1+C} \vee 1\right)\right\rfloor(\alpha+1/2)} \\ &\leqslant 2^{\alpha+1/2} B_2 F \left(\left(\frac{m}{1+C}\right) \vee 1\right)^{-(\alpha+1/2)} \\ &\leqslant \underbrace{2^{\alpha+1/2} B_2 (1 + C)^{\alpha+\frac{1}{2}}}_{B_4=} F m^{-(\alpha+1/2)}, \quad \text{for } m \in \mathbb{N}. \end{aligned} \tag{26}$$

Using (25) and (26) combined with the fact that at most $C+1$ of the scaling coefficients $\langle f, \varphi_{0,k}\rangle$, $k \in \mathbb{Z}$, are nonzero, we find that

$$\begin{aligned} f_{\mathcal{B}}[m] &\leqslant \max\left\{ f_{\mathcal{B},\mathrm{I}}\left[\left\lfloor \frac{m-(C+1)+1}{2} \right\rfloor\right], f_{\mathcal{B},\mathrm{II}}\left[\left\lfloor \frac{m-(C+1)+1}{2} \right\rfloor\right] \right\} \\ &\leqslant F \cdot \max\left\{ B_3\, 2^{-\frac{1}{2KC}\lfloor \frac{m-C}{2} \rfloor}, B_4 \left\lfloor \frac{m-C}{2} \right\rfloor^{-(\alpha+1/2)} \right\} \\ &\leqslant F \cdot B_5 (m-C)^{-(\alpha+1/2)}, \quad \text{for } m > C+1, \end{aligned}$$

where B_5 depends on B_3, B_4, K, C, and α. We can therefore find a constant B_6 depending on φ, ψ, α, and K such that $f_{\mathcal{B}}[m] \leqslant B_6 F m^{-(\alpha+\frac{1}{2})}$, for $m \in \mathbb{N}$. It now follows by Theorem 6 that

$$\epsilon_n(M, f) \leqslant \frac{B_6^2}{2\alpha} \cdot F \cdot M^{-2\alpha}, \quad \text{for } M \in \mathbb{N},$$

as desired.

□

6 Generalization of Wavelets

Although one-dimensional wavelets have good resolution properties in both the time and frequency domains, they lack the ability to resolve signals along arbitrary directions in $2D$ and $3D$. In addition, a large number of wavelet coefficients are required to account for edges, i.e., singularities along lines and curves. Several extensions of the wavelet transform have been proposed to incorporate the resolution of singularities along lines and curves: Among the most famous ones, one can cite curvelets, ridgelets, shearlets.

References

[1] S. Mallat, *A wavelet tour of signal processing: the sparse way*, ser. Third edition. Burlington, MA, USA: Academic Press, Elsevier, 2009.

[2] *A first course on wavelets*. CRC press, 1996.

[3] Y. Meyer, *Wavelets and operators*, ser. Cambridge Studies in Advanced Mathematics. Cambridge, UK: Cambridge University Press, 1992, vol. 37.