Rose-Hulman Undergraduate Mathematics Journal

Volume 21
Issue 1

Article 12

On the Construction and Mathematical Analysis of the Wavelet Transform and its Matricial Properties

Diego Sejas Viscarra
Universidad Mayor de San Simón, dsejas.math@pm.me

On the Construction and Mathematical Analysis of the Wavelet Transform and its Matricial Properties

Cover Page Footnote

I wish to thank my advisor, Santiago Relos Paco, for the mathematical guidance and support. Special thanks to my father, Gastón Sejas Cortés, who has provided valuable advise and moral support through the process of writing.

ℝose-Hulman Undergraduate Mathematics Journal
Volume 21, Issue 1, 2020

On the Construction and Mathematical Analysis of the Wavelet Transform and its Matricial Properties

By *Diego Sejas Viscarra*

Abstract. We study the properties of computational methods for the Wavelet Transform and its Inverse from the point of view of Linear Algebra. We present a characterization of such methods as matrix products, proving in particular that each iteration corresponds to the multiplication of an adequate unitary matrix. From that point we prove that some important properties of the Continuous Wavelet Transform, such as linearity, distributivity over matrix multiplication, isometry, etc., are inherited by these discrete methods.

This work is divided into four sections. The first section corresponds to the classical theoretical foundation of harmonic analysis with wavelets; it is used for clarity only. The second section presents the construction of the Discrete Wavelet Transform for vectors and its Inverse, emphasizing on storage efficiency. The third section presents the generalization of the Transform to matrices. It is equivalent to section 2, but some methods and tools used are slightly different, showing an alternative approach to the subject. The fourth section presents the main results of this work.

1 Introduction

The *Wavelet Transform* is a mathematical tool that has proved to be very useful on different contexts. For instance, a very important problem of analysis, namely, the representation of a signal[1] by a sequence of coefficients, is satisfied by the wavelet transform; moreover, the process is invertible, so a signal can be reconstructed from its representing sequence [Daubechies, 1992]. Since this process is based on the decomposition of functions into simpler objects (called *wavelets*), it also allows arbitrary approximations [Mallat, 1999] and the extraction of valuable information [Daubechies, 1992, Mallat, 1999].

Mathematics Subject Classification. Primary: 65T60, secondary: 45C40
Keywords. Matricial properties of the wavelet transform, wavelet transform, matrix product, linearity.

[1]The term *signal* is used as a synonym for *function*.

Although this is enough motivation to study the wavelet transform, there is another important reason: it is a known fact that the *Fourier transform* suffers a major drawback, i.e., the sinusoids it uses lack of time localization, so it let us know the frequency content—so to speak—of a signal, but not the lapse of time of the occurrence of a given frequency. An adaptation to solve this problem is the *Gabor transform*, also called *windowed Fourier transform*, which relies on the use of a "window", that is, a function with good time localization [Daubechies, 1992, Gomes and Velho, 1999, Sossa, 2009]. However, this method has its own drawback: If the details of function are much smaller than the width of the window, they will be detected but not localized; if they are much bigger, they won't be detected properly. The wavelet transform, however, solves these both problems, since it has good time localization and, there's no matter how big or small are the details of a function, they'll be properly detected.

Surprisingly, these are not the only reasons to study it, since the wavelet transform has a number of unexpected applications such as image compression [Beatty, 2004, Daubechies, 1992, Mallat, 1999, Bultheel, 2006, Sossa, 2009] (e.g., the FBI fingerprint database [Graps, 1995, Sossa, 2009]), signal denoising [Daubechies, 1992, Mallat, 1999, Bultheel, 2006, Graps, 1995, Sossa, 2009] and fast matrix multiplication [Beatty, 2004], among others.

1.1 Conventions and notations

Throughout this work, by *vector* we understand *column vector*. Vectors will be denoted by lowercase letters such as u, v, w, while matrices will be denoted by uppercase letters such as A, B, C. The elements of a vector are implicitly numbered starting with zero; likewise, the rows and columns of a matrix are numbered starting with zero. If v is a vector, its ith element will be denoted v_i; if A is a matrix, its (i, j)th element—the one corresponding to its ith row and jth column—will be denoted as a_i^j.

A thick dot like • indicates that a variable is to be placed in that position. For example, $f(\bullet)$ must be understood as $f(t)$ or $f(s)$, i.e., f is a function of one variable. Two thick dots indicate the positions of two different variables, unless they are located on different sides of the equal sign. For example $f(\bullet,\bullet)$ must be understood as $f(s,t)$, but $f(\bullet) = \sin(\bullet)$ must be understood as $f(t) = \sin(t)$.

Unless otherwise explicitly stated, we will assume the set of variation of a variable as the largest possible. For example, (a_i) stands for $(a_i)_{i\in\mathbb{Z}}$ in case i is an integer variable; $\int f(t)\,dt$ stands for the integral on the complete domain of f. If one variable is fixed in an expression while other is not, it will be indicated showing the non fixed variable; for example, $(a_{i,j})_j$ indicates i is fixed while j is variable; $\sum_{i=0}^{n-1} a_{i,j}$ indicates j is fixed and i is variable.

The notation $x \bmod y$ stands for the remainder of the integer division x/y. The complex conjugate of $z \in \mathbb{C}$ will be denoted as $\overline{z}$. Matrix transposition is indicated by an apostrophe.

Further notations will be introduced when needed.

1.2 Mathematical preliminaries

This subsection is mainly intended as a brief exposition of the basic concepts necessary for the following development. It also presents some results about the continuous wavelet transform that we will prove later for the discrete case. No proof will be presented here since this is only an introduction, but the interested reader can consult [Daubechies, 1992] or [Sejas, 2012] (in Spanish) for further detail.

1.2.1 On the the continuous wavelet transform. We start by considering the *space of square integrable signals on* $\mathbb{R}$, denoted $\mathscr{L}^2(\mathbb{R})$ and defined by

$$\mathscr{L}^2(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{C} : \int |f(t)|^2 \, dt < \infty \right\}.$$

With the usual operations of sum of functions and multiplication by a scalar, this is a complex vector space. Even more, with the scalar product $\langle \bullet \mid \bullet \rangle$ and the norm $\|\bullet\|$, defined for $f, g \in \mathscr{L}^2(\mathbb{R})$ by

$$\langle f \mid g \rangle = \int f(t) \overline{g(t)} \, dt$$

and

$$\|f\| = \int |f(t)|^2 \, dt,$$

respectively, this is a Hilbert space.

Among all signals in $\mathscr{L}^2(\mathbb{R})$, the ones of our particular interest right now are wavelets.

Definition 1.1. For a signal $\psi \in \mathscr{L}^2(\mathbb{R})$ we define

$$\boxed{C_\psi = 2\pi \int \frac{|\widehat{\psi}(\omega)|^2}{|\omega|} \, d\omega},$$

where $\widehat{\psi}$ denotes the Fourier Transform of ψ. Then, we say that ψ is a *wavelet* if it satisfies

$$\boxed{0 < C_\psi < \infty}. \tag{1}$$

The reason for this restriction will be clear later.

Example 1.2. Figure 1 shows two well-known wavelets. Observe their oscillatory behavior from which they obtain the name.

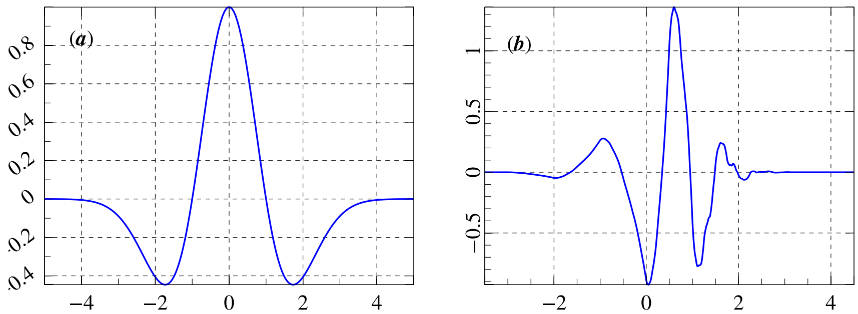

Figure 1: Two common wavelets: **(a)** the mexican hat and **(b)** Daubechies with four vanishing moments.

The intuitive idea behind the wavelet transform is very simple: A wavelet ψ is "concentrated" around a given frequency and a given time. By means of a dilation (alternatively, contraction) we decrease (alternatively, increase) this central frequency, and by means of a translation we modify the central time, thus obtaining a new wavelet—let's denote it $\psi_{a,b}$. Given a function $f \in \mathscr{L}^2(\mathbb{R})$, the scalar product $\langle f \mid \psi_{a,b}\rangle$ gives us a sort of measure of the behavior of f around the new time and frequency.

Definition 1.3. Let ψ be a wavelet. We call *family of wavelets* (*generated by* ψ) the family $(\psi_{a,b})$ defined by

$$\psi_{a,b}(\bullet) = |a|^{-1/2}\psi\left(\frac{\bullet - b}{a}\right),$$

for $a \in \mathbb{R}_0 = \mathbb{R} \setminus \{0\}$, called *scale parameter*, and $b \in \mathbb{R}$, called *translation parameter*. In this case, we say that ψ is the *mother wavelet*.

The factor $|a|^{-1/2}$ guaranties that $\|\psi_{a,b}\| = \|\psi\|$, which is a very important condition to obtain orthonormal wavelet bases.

Definition 1.4. Let ψ be a wavelet and $f \in \mathscr{L}^2(\mathbb{R})$. The (*continuous*) *wavelet transform of* f is the function $\mathrm{W}(f)$, defined for $a \in \mathbb{R}_0$ and $b \in \mathbb{R}$ by

$$\boxed{\mathrm{W}(f)(a,b) = \langle f \mid \psi_{a,b}\rangle = |a|^{-1/2}\int f(t)\,\overline{\psi\left(\frac{t-b}{a}\right)}\,dt}.$$

Example 1.5. Figure 2 shows two wavelet transforms of the signal

$$f(t) = \begin{cases} \cos\left(\frac{2\pi}{25}t\right) & \text{if } 200 \le t < 400, \\ \cos\left(\frac{2\pi}{70}t\right) & \text{if } 400 \le t < 800, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

You can observe the good time localization.

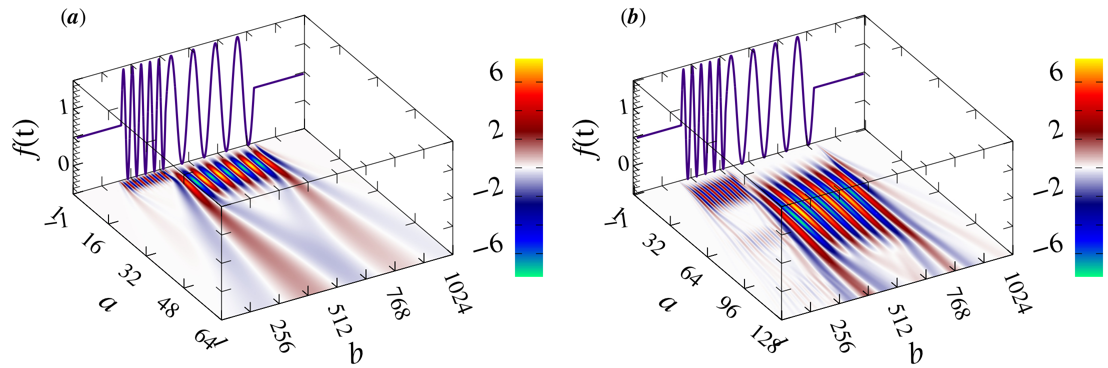

Figure 2: Two wavelet transforms of the signal (2), **(a)** using the mexican hat and **(b)** using Daubechies with four vanishing moments. (Data obtained with Scilab v6.0.1 and the Scilab Wavelet Toolbox v0.3.1-3.)

One of the most important properties of the wavelet transform is its invertibility. Indeed, for all $f \in \mathscr{L}^2(\mathbb{R})$,

$$\boxed{f = \frac{1}{C_\psi} \iint W(f)(a,b)\,\psi_{a,b}\,\frac{da\,db}{a^2}}. \tag{3}$$

Equation (3) is called *resolution of the identity*, and gives us a formula for the *inverse wavelet transform*, which we denote $W^{-1}(\bullet)$. Here we can see the reason for the restriction in (1) in **definition 1.1**.

Convergence of the double integral in (3) must be understood in the weak sense, i.e., taking the scalar product with a signal $g \in \mathscr{L}^2(\mathbb{R})$ on both sides, an equation is obtained that is proved to be true in the following result:

Theorem 1.6. *Let ψ be a wavelet. For $f, g \in \mathscr{L}^2(\mathbb{R})$ we have that*

$$C_\psi \langle f \mid g \rangle = \iint W(f)(a,b)\,\overline{W(g)(a,b)}\,\frac{da\,db}{a^2}. \tag{4}$$

There is also a slightly stronger sense of convergence of the right-hand side expression in (3):

Theorem 1.7. *Let ψ be a wavelet. For $f \in \mathscr{L}^2(\mathbb{R})$ we have that*

$$\lim_{\substack{A_1 \to 0 \\ A_2, B \to \infty}} \left\| f - \frac{1}{C_\psi} \iint_{\substack{A_1 \le |a| \le A_2 \\ |b| \le B}} W(f)(a,b)\,\psi_{a,b}\,\frac{da\,db}{a^2} \right\| = 0, \tag{5}$$

where the double integral stands for the unique element in $\mathscr{L}^2(\mathbb{R})$ which inner product with $g \in \mathscr{L}^2(\mathbb{R})$ is given by (4).

We can interpret (5) as a guarantee that any signal in $\mathscr{L}^2(\mathbb{R})$ can be arbitrarily approximated by a superposition of wavelets [Daubechies, 1992].

The following is a direct consequence of **theorem 1.6**.

Corollary 1.8. *Let* $\mathscr{L}^2(\mathbb{R}^2, d\mu)$ *denote the space of functions square integrable on* $(\mathbb{R}\setminus\{0\})\times\mathbb{R}$ *with respect to the measure* $d\mu = \frac{da\,db}{C_\psi a^2}$*, then:*

1. *If* $f, g \in \mathscr{L}^2(\mathbb{R})$*, then* $W(f), W(g) \in \mathscr{L}^2(\mathbb{R}^2, d\mu)$*, and*

$$\langle W(f) \mid W(g)\rangle = \langle f \mid g\rangle,$$

   i.e., $W(\bullet)$ *is an isometry.*

2. *If* $F, G \in \mathscr{L}^2(\mathbb{R}^2, d\mu)$*, then* $W^{-1}(F), W^{-1}(G) \in \mathscr{L}^2(\mathbb{R})$*, and*

$$\langle W^{-1}(F) \mid W^{-1}(G)\rangle = \langle F \mid G\rangle,$$

   i.e., $W^{-1}(\bullet)$ *is an isometry.*

1.2.2 On the discrete wavelet transform.

It is important to count with a discrete version of the wavelet transform for practical purposes. That is, we want to be able to compute the transform on a discrete subset of its domain, but in such a way that the important properties of representation and reconstruction of signals are not lost.

One possible discretization of the scale parameter is $a = a_0^j$, where $a_0 > 1$ is fixed and $j \in \mathbb{Z}$ is variable. To discretize the translation parameter, let's observe that for narrow wavelets (for small a), small translations are necessary in order to cover the complete domain, while for wide wavelets (for large a), only large translations are necessary. So a natural discretization is $b = kb_0a$, where $b_0 > 0$ is fixed and $k \in \mathbb{Z}$ is variable. The set of pairs $\left(a_0^j, kb_0a_0^j\right)$ will be called *discrete lattice* or *discrete grid*, and will be denoted Δ_{a_0,b_0}.

In order to compute the wavelet transform on a discrete lattice, only a subset of the whole family of wavelets is necessary

Definition 1.9. Let ψ be a wavelet and Δ_{a_0,b_0} a lattice. We call *family of wavelets generated by* ψ (*associated to* Δ_{a_0,b_0}) the doubly indexed sequence $(\psi_{j,k})$ defined by

$$\psi_{j,k}(\bullet) = a_0^{-j/2}\,\psi\!\left(a_0^{-j}\bullet - kb_0\right)$$

for $j, k \in \mathbb{Z}$. The function ψ is called *mother wavelet*.

The definition of the wavelet transform is not different at all in this case, but now we can regard it a sequence.

Definition 1.10. Let ψ be a wavelet and Δ_{a_0,b_0} a lattice. The *discrete wavelet transform of f (associated to Δ_{a_0,b_0})* is the doubly indexed sequence

$$\boxed{\mathrm{W}(f) = \left(\langle f \mid \psi_{j,k}\rangle\right)}.$$

In general, not every choice of ψ and Δ_{a_0,b_0} will guarantee the representation and reconstruction of signals by the wavelet transform; however, a very special—but not too restrictive case—where these properties are not lost is when the family of wavelets $(\psi_{j,k})$ is an orthonormal base [Mallat, 1999].

Definition 1.11. A family $(\mathrm{F}_{j,k})$ in $\mathscr{L}^2(\mathbb{R})$ is an orthonormal basis if the following conditions hold:

1. $\langle \mathrm{F}_{m,n} \mid \mathrm{F}_{p,q}\rangle = \delta_{m,p}\delta_{n,q}$, for every $m, n, p, q \in \mathbb{Z}$, where δ denotes Kronecker's delta function,[2] and
2. for $f \in \mathscr{L}^2(\mathbb{R})$,
$$\|f\|^2 = \sum_j \sum_k \left|\langle f \mid \mathrm{F}_{j,k}\rangle\right|^2.$$

The following result gives us an expression for the *discrete inverse wavelet transform*, which we will denote $\mathrm{W}^{-1}(\bullet)$.

Theorem 1.12. *If the family of wavelets $(\psi_{j,k})$ associated to the lattice Δ_{a_0,b_0} is an orthonormal basis of $\mathscr{L}^2(\mathbb{R})$, then, for $f \in \mathscr{L}^2(\mathbb{R})$,*

$$\boxed{f = \sum_j \sum_k \langle f \mid \psi_{j,k}\rangle \, \psi_{j,k}}, \tag{6}$$

where convergence holds in the $\mathscr{L}^2(\mathbb{R})$ sense.

Notice that (6) is very similar to (3); actually, this is a discrete version of the resolution of the identity.

Remark. In the case of the wavelet ψ having a compact support,[3] the wavelet transform can be applied to signals that are square integrable over finite intervals.

2 Construction of the vectorial method

This section deals with the construction of a fast and efficient method to compute the wavelet transform and its inverse, which can also be generalized to be applied on vectors. Our method differs slightly from the canonical one, in that we build and use periodic signals, instead of compact-supported ones, in order to obtain a discrete equivalent of the wavelet transform and its inverse.

[2] Defined by $\delta_{i,j} = 1$, if $i = j$, and $\delta_{i,j} = 0$, otherwise.

[3] The subset of the domain of a signal where it takes non-zero values.

2.1 The multiresolution analysis

The multiresolution analysis is a very important tool in wavelet analysis. Among its advantages we can count the construction of a fast and efficient algorithm to compute the wavelet transform and its inverse, a method to construct orthonormal wavelet bases, and the easy generalization of concepts to vectors and matrices.

The idea behind multiresolution analysis consists an decomposing the space $\mathscr{L}^2(\mathbb{R})$ into a decreasing sequence (V_j) of subspaces with some important properties. Given a signal $f \in \mathscr{L}^2(\mathbb{R})$, its orthogonal projection $\mathrm{P}_j f$ over V_j is an approximation with certain level of detail (level of resolution). The signal $\mathrm{P}_{j-1}f - \mathrm{P}_j f$ is in a space W_j—which is the orthogonal complement of V_j in V_{j-1}—, where an orthonormal wavelet basis exists. The coefficients of this signal with respect to this basis are the values of the wavelet transform with the corresponding level of resolution.

For what's left of this article, let's use the *dyadic lattice* $\Delta_{2,1}$. So we implicitly adopt the notation $\psi_{j,k}(\bullet) = 2^{-j/2}\psi(2^{-j}\bullet - k)$.

Definition 2.1. A *multiresolution analysis* (*of* $\mathscr{L}^2(\mathbb{R})$) is a sequence (V_j) of closed subspaces that satisfies:

1. $\mathrm{V}_j \subset \mathrm{V}_{j-1}$ for $j \in \mathbb{Z}$,
2. $\overline{\bigcup \mathrm{V}_j} = \mathscr{L}^2(\mathbb{R})$,
3. $\bigcap \mathrm{V}_j = \{0\}$,
4. $f(\bullet) \in \mathrm{V}_j$ if and only of $f(2^j\bullet) \in \mathrm{V}_0$ for $j \in \mathbb{Z}$, and
5. there exists $\phi \in \mathrm{V}_0$ such that $(\phi(\bullet - k))$ is an orthonormal basis of V_0.

Every V_j is called *scale space* and ϕ is called *scale function* or *father wavelet.*

Notice that the multiresolution aspect comes from condition item 4, that states that every V_j is a scaled version of V_0.

Example 2.2. The well-known Haar multiresolution analysis (also called *Daubechies with one vanishing moment*) is given by

$$\mathrm{V}_j = \left\{ f \in \mathscr{L}^2(\mathbb{R}) : f \text{ is constant over } \left[2^j k, 2^j(k+1)\right) \text{ for } k \in \mathbb{Z} \right\}$$

and

$$\phi(t) = \begin{cases} 1 & \text{if } t \in [0,1), \\ 0 & \text{otherwise.} \end{cases}$$

We can see that conditions items 4 and 5 of a multiresolution analysis imply that $(\phi_{j,k})_k$ is an orthonormal basis of V_j. Since $\phi \in V_0 \subset V_{-1}$, we can write

$$\boxed{\phi = \sum \langle \phi \mid \phi_{-1,k} \rangle \, \phi_{-1,k}}, \tag{7}$$

where convergence holds in the $\mathscr{L}^2(\mathbb{R})$ sense.

Definition 2.3. Let ϕ be the father wavelet of a multiresolution analysis. The sequence $(h_k) = \big(\langle \phi \mid \phi_{-1,k} \rangle\big)$ is called *scale filter*.

The following result is the link point between MRA and the wavelet transform. More on this theorem can be found in [Daubechies, 1992]

Theorem 2.4. *Let* (V_j) *be a multiresolution analysis and* (h_k) *the corresponding scale filter. Then there exists a wavelet* ψ *such that*

1. *the sequence* $(\psi_{j,k})_k$ *is an orthonormal basis of* W_j*;*
2. *the sequence* $(\psi_{j,k})$ *is an orthonormal basis of* $\mathscr{L}^2(\mathbb{R})$*;*
3. *for* $f \in \mathscr{L}^2(\mathbb{R})$ *and* $j \in \mathbb{Z}$*,*

$$P_{j-1} f - P_j f = \sum_k \langle f \mid \psi_{j,k} \rangle \, \psi_{j,k}. \tag{8}$$

One possible choice for ψ *is*

$$\psi = \sum_k (-1)^k \, \overline{h_{1-k}} \, \phi_{-1,k}.$$

Notice that (8) contains the values of the wavelet transform for $a = 2^j$ fixed and $b = k2^j$ variable with $k \in \mathbb{Z}$.

Definition 2.5. Let (V_j) be a multiresolution analysis with associated scale filter (h_k). The *wavelet induced by the MRA* (V_j) is the signal

$$\boxed{\psi = \sum_k (-1)^k \, \overline{h_{1-k}} \, \phi_{-1,k}}.$$

The sequence $(g_k) = \big((-1)^k \, \overline{h_{1-k}}\big)$ is called *wavelet filter*.

Example 2.6. Figure 3 show the scale functions and induced wavelets for the four first Daubechies MRA's.

The following result is a direct consequence of **theorem 1.12** and **theorem 2.4**.

Corollary 2.7. *If* ψ *is a wavelet induced by a multiresolution analysis, the corresponding wavelet transform is invertible.*

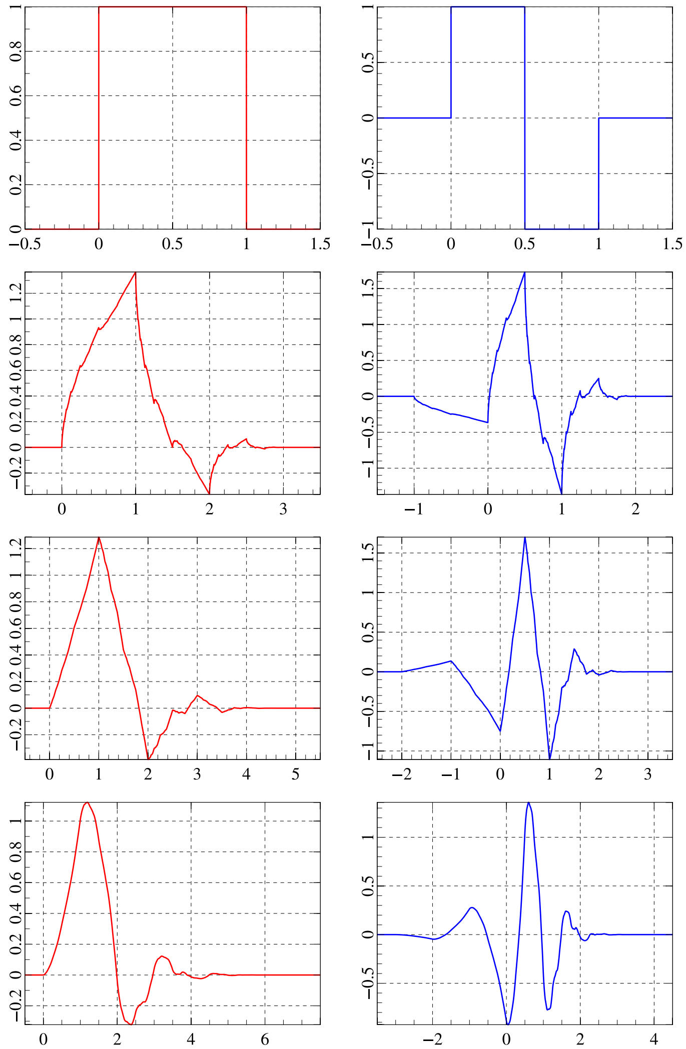

Figure 3: **Left to right:** Scale function and associated wavelet. **Top to bottom:** Daubechies with one, two, three and four vanishing moments.

2.2 The decomposition algorithm for vectors

2.2.1 The fast wavelet transform: The decomposition equations. *Fast wavelet transform* is the name given to a hierarchical algorithm based on the MRA. Right now we are only interested on the equations that characterize this method.

Let (V_j) be a MRA, ϕ the corresponding scale function, ψ the corresponding induced wavelet, and $f \in \mathscr{L}^2(\mathbb{R})$. Let's adopt the following notation:

- $c^j = \left(c^j_k\right) = \left(\langle f \mid \phi_{j,k}\rangle\right)$, and
- $d^j = \left(d^j_k\right) = \left(\langle f \mid \psi_{j,k}\rangle\right)$.

Let $j, k \in \mathbb{Z}$. It is evident that

$$c_k^{j+1} = \langle \mathrm{P}_j f \mid \phi_{j+1,k} \rangle. \tag{9}$$

We have that

$$\mathrm{P}_j f = \sum_p c_p^j \phi_{j,p}. \tag{10}$$

On the other hand, by (7),

$$\begin{aligned} \phi_{j+1,k} &= 2^{-(j+1)/2} \phi\left(2^{-(j+1)} \bullet - k\right) \\ &= 2^{-(j+1)/2} \sum_q h_q \phi_{-1,q}\left(2^{-(j+1)} \bullet - k\right) \\ &= \sum_q h_q 2^{-j/2} \phi\left(2^{-j} \bullet - (2k+q)\right) \\ &= \sum_q h_q \phi_{j,2k+q}. \end{aligned} \tag{11}$$

Replacing (10) and (11) in (9),

$$\begin{aligned} c_k^{j+1} &= \left\langle \sum_p c_p^j \phi_{j,p} \,\middle|\, \sum_q h_q \phi_{j,2k+q} \right\rangle \\ &= \sum_p \sum_q c_p^j \overline{h_q} \langle \phi_{j,p} \mid \phi_{j,2k+q} \rangle. \end{aligned}$$

Remembering that $(\phi_{j,p})_p$ is an orthonormal set,

$$\boxed{c_k^{j+1} = \sum_p \overline{h_{p-2k}} \, c_p^j}. \tag{12}$$

A similar procedure leads to

$$\boxed{d_k^{j+1} = \sum_p \overline{g_{p-2k}} \, c_p^j}. \tag{13}$$

Let us suppose that the sequence c^0 is known for some $f \in \mathscr{L}^2(\mathbb{R})$. By means of equations (12) and (13) we obtain a hierarchical decomposition scheme, called *pyramidal decomposition*, which is schematically represented in fig. 4. Every use of these equations (corresponding to one iteration) gives us the wavelet transform with a new level of resolution, i.e, the sequences d^j.

2.2.2 Construction of the vector decomposition algorithm. When studying a phenomenon with continuous behavior, it is generally impossible to have but a discrete finite sample of this behavior (a vector). It is thus very important to count with a version of the wavelet transform suitable to be applied to vectors.

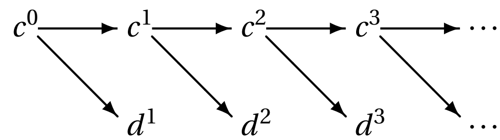

Figure 4: The classical representation of pyramidal decomposition.

For what's left of this work, by saying that a sequence s is *finite* we mean that it has a finite number of non-zero terms; in this case we may use the notation $s = (s_{n_1}, \dots, s_{n_2})'$, where s_{n_1} and s_{n_2} are the first and last non-zero terms, respectively, and define its length $|s| = n_2 - n_1 + 1$. By an abuse of notation, for every vector v, we shall also write $|v| = n_0$ to mean that it has n_0 elements.

Since we are trying to construct an algorithm, it is reasonable to think that the filters (h_k) and (g_k) are finite;[4] this implies that ϕ and ψ have compact support.[5]

Convention. Let's assume from now on that, by means of adequate translations, the scale filter and wavelet filter have the forms $(h_k) = (h_0, \dots, h_{n-1})'$ and $(g_k) = (g_0, \dots, g_{n-1})'$, respectively, where $n = |(h_k)|$.

Let $v \in \mathbb{C}^{n_0}$. Since the pyramidal decomposition can't be directly applied to vectors, but to signals, we define a function $f_v \in \mathscr{L}^2(\mathbb{R})$ that characterizes the elements of v. Intuition dictates that the natural way to do this is by defining

$$f_v = \sum_{k=0}^{n_0-1} v_k \, \phi_{0,k}.$$

By taking the scalar products $\langle f_v \mid \phi_{0,k} \rangle$ for $k \in \mathbb{Z}$, we have

$$c_k^0 = \begin{cases} v_k & \text{if } k \in \{0, \dots, n_0 - 1\}, \\ 0 & \text{otherwise,} \end{cases}$$

so this seems to be a good choice. However, this definition of f_v leads to an information increment problem. Indeed, let's consider the following example: Suppose $(h_k) = (h_0, \dots, h_5)'$, $(g_k) = (g_0, \dots, g_5)'$, and $v = (v_0, \dots, v_9)'$. If we proceed as described above, applying the decomposition equations (12) and (13) we get

$$c^1 = \left(c^1_{-2}, \dots, c^1_4\right)' \quad \text{and} \quad d^1 = \left(d^1_{-2}, \dots, d^1_4\right)'.$$

We see that starting with a vector with ten elements, we obtain two vectors[6] with seven elements. In general, the increment of information after n iterations with filters

[4] Notice that they both have the same length.

[5] This contributes to the convergence of series defined with this signals.

[6] Remember they are actually sequences.

with even length L will be $n(\mathrm{L}-2)$. This is an obvious problem when trying to implement the algorithm in a computer where the storage space is limited.

We propose a simple yet elegant and ingenious solution: Define f_v to be periodic.[7] In this case, the sequences c^j and d^j will be infinite but periodical, so only the non-redundant information must be stored.

Without any loss of generality we assume $n_0 = |v|$ is even; if this is not the case, a null element can always be added. We define

$$\boxed{f_v = \sum_k v_{k \bmod n_0} \phi_{0,k}}. \tag{14}$$

Then we have that

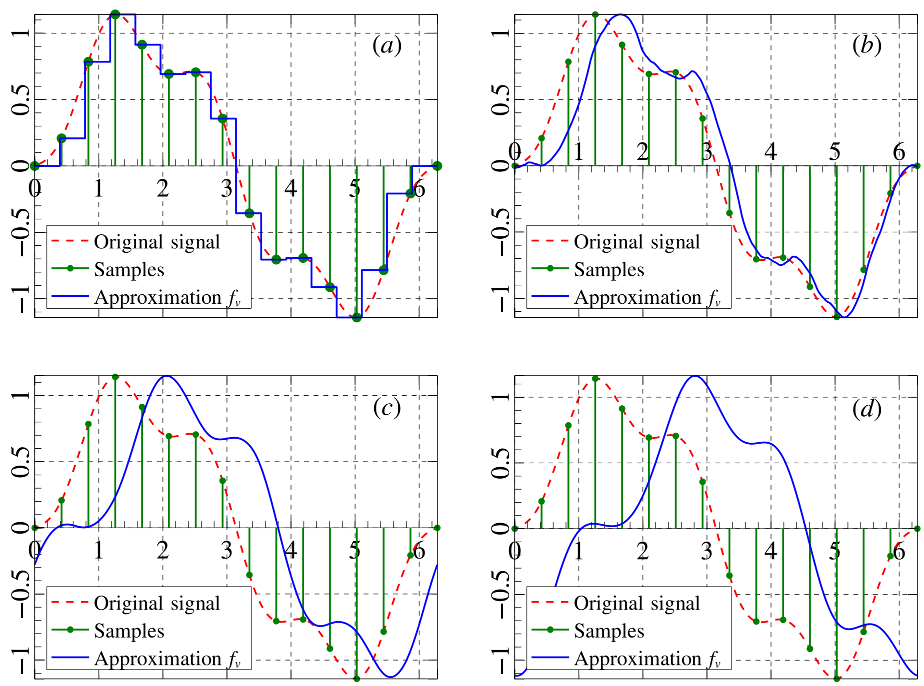

Figure 5: The function f_v as defined in (14) using different Daubechies filters: ***(a)*** one, ***(b)*** four, ***(c)*** ten, and ***(d)*** twenty vanishing moments. Notice that f_v approximates very well the original signal, although there is an interesting translation phenomenon directly proportional to the length of the filter.

$$c_k^0 = v_{k \bmod n_0}.$$

[7] This is also proposed independently of the author in [Gomes and Velho, 1999], where other alternatives for f_v are also studied.

Applying the decomposition equations (12) and (13) we get

$$c^1_k = c^1_{k \bmod (n_0/2)} \quad \text{and} \quad d^1_k = d^1_{k \bmod (n_0/2)}.$$

We can observe that it is necessary to store only n_0 numbers, so no extra storage space is needed.

In general, c^j and d^j have period $n_0/2^j$, so it is enough to store $c^1_0, \dots, c^1_{n_0/2-1}, d^1_0, \dots, d^1_{n_0/2-1}$ in the first iteration; $c^2_0, \dots, c^2_{n_0/4-1}, d^2_0, \dots, c^2_{n_0/4-1}, c^1_0, \dots, c^1_{n_0/2-1}$ in the second iteration, and so on[8]—exactly n_0 numbers every time.

2.2.3 The wavelet transform of a vector.

For the rest of this article, the notation $x_{n_1:n_2}$, where x is a vector or sequence, and $n_1, n_2 \in \mathbb{Z}$ with $n_1 \leq n_2$, is equivalent to $(x_{n_1}, \dots, x_{n_2})'$.

Definition 2.8. Let v be a vector with convenient length n_0. We call (*discrete*) *wavelet transform of v with n levels of resolution*, and denote $\mathrm{W}^n(v)$, the vector

$$\left(c^n_{0:n_0/2^n-1}, d^n_{0:n_0/2^n-1}, \dots, d^1_{0:n_0/2-1}\right)'$$

When $n = 1$ we can make omission and simply write $\mathrm{W}(v)$. We also agree that $\mathrm{W}^0(v) = v$.

Example 2.9. Let us consider

$$(h_k) = \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right)', \quad (g_k) = \left(\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right)', \quad v = (0,1,2,3)'.$$

We have the following:

- Wavelet transform of v with one level of resolution:

$$\mathrm{W}(v) = \left(\frac{1}{\sqrt{2}}, \frac{5}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)',$$

- wavelet transform of v with two levels of resolution:

$$\mathrm{W}^2(v) = \left(3, -2, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)'.$$

[8] We have implicitly assumed here that $|v| = k2^n$, for some $k \in \mathbb{N}$, with $k \geq 1$, and n is the number of iterations. In this case we say that v has *convenient length*.

Figure 6 gives a more graphical idea of the effect of the wavelet transform.

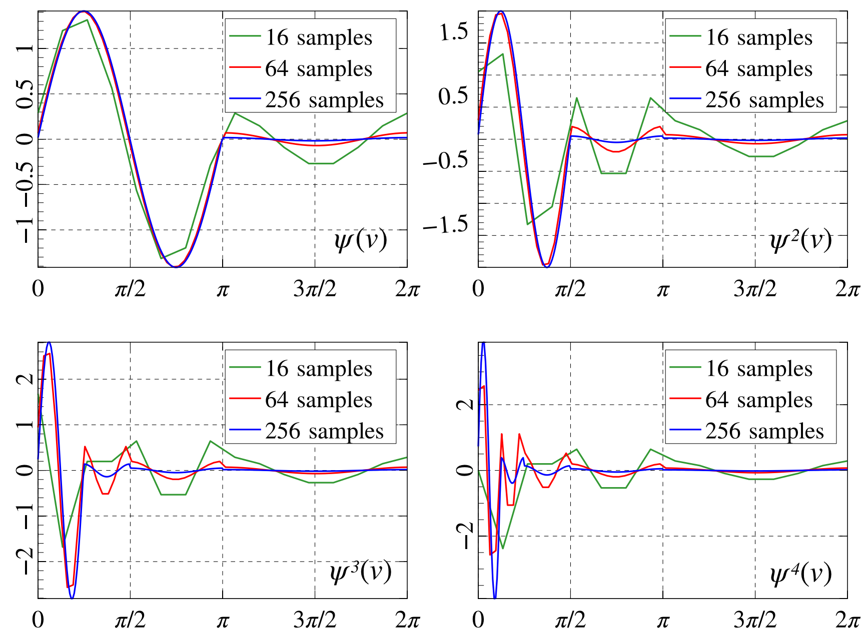

Figure 6: The wavelet transforms of discrete samplings of $\sin(t)$ on $[0,2\pi)$ with different levels of resolution.

2.3 The reconstruction algorithm for vectors

2.3.1 The fast inverse wavelet transform: The reconstruction equation. Let $j,k \in \mathbb{Z}$. It is evident that

$$c_k^j = \langle \mathrm{P}_j f \mid \phi_{j,k} \rangle. \tag{15}$$

By (8) we have that

$$\mathrm{P}_j f = \sum_p c_p^{j+1} \phi_{j+1,p} + \sum_p d_p^{j+1} \psi_{j+1,p}. \tag{16}$$

On the other hand, by (11),

$$\phi_{j+1,p} = \sum_q h_q \phi_{j,2p+q}, \tag{17}$$

and with a similar deduction,

$$\psi_{j+1,p} = \sum_q g_q \phi_{j,2p+q}. \tag{18}$$

Replacing (16), (17) and (18) in (15),

$$\begin{aligned} c_k^j &= \Big\langle \sum_p c_p^{j+1} \sum_q h_q \phi_{j,2p+q} + \sum_p d_p^{j+1} \sum_q g_q \phi_{j,2p+q} \,\Big|\, \phi_{j,k} \Big\rangle \\ &= \sum_p \sum_q c_p^{j+1} h_q \langle \phi_{j.2p+q} \mid \phi_{j,k} \rangle + \sum_p \sum_q d_p^{j+1} g_p \langle \phi_{j,2p+q} \mid \phi_{j,k} \rangle. \end{aligned}$$

Finally, remembering $(\phi_{j,k})_k$ is an orthonormal set,

$$\boxed{c_k^j = \sum_p h_{k-2p} c_p^{j+1} + \sum_p g_{k-2p} d_p^{j+1}}. \tag{19}$$

Suppose the sequences c_n, d^n, ..., d^1 are known for some $f \in \mathscr{L}^2(\mathbb{R})$ and $n \in \mathbb{N}$. By means of (19) we obtain a reconstruction scheme, called *pyramidal reconstruction*, which is schematically represented in fig. 7. After n iterations the sequence c^0 is obtained.

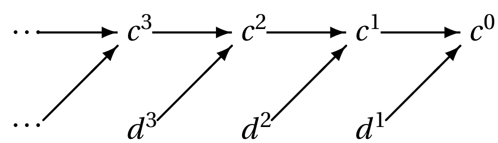

Figure 7: The classical representation of pyramidal reconstruction.

2.3.2 Construction of the vector reconstruction algorithm.

Let w be of the form

$$w = \left(u_0^m, \dots, u_{n_0/2^m-1}^m, w_0^m, \dots, w_{n_0/2^m-1}^m, \dots, u_0^1, \dots, u_{n_0/2-1}^1\right)'$$

for $m \in \mathbb{N}$ and n_0 a convenient number, i.e., $w = \mathrm{W}^m(v)$ for some $v \in \mathbb{C}^{n_0}$. Following the ideas previously presented regarding periodicity, we define for $k \in \mathbb{Z}$,

$$c_k^m = u^m_{k \bmod (n_0/2^m)}, \quad d_k^m = w^m_{k \bmod (n_0/2^m)}, \quad \dots \quad , \quad d_k^1 = w^1_{k \bmod (n_0/2)}.$$

By applying the reconstruction equation (19), after n iterations, with $n \le m$, we obtain

$$c_k^{m-n} = c^{m-n}_{k \bmod (n_0/2^{m-n})}, \quad d_k^{m-n} = d^{m-n}_{k \bmod (n_0/2^{m-n})}, \quad \dots \quad , \quad d_k^1 = d^1_{k \bmod (n_0/2)}.$$

As before, it is sufficient to store only the non-redundant information, exactly $n_0 = |w|$ elements in every iteration.

2.3.3 The inverse wavelet transform of a vector.

Definition 2.10. Let w be a vector of the form $w = \mathrm{W}^m(v)$ for some $m \in \mathbb{N}$ and $v \in \mathbb{C}^{n_0}$, and let $n \le m$. We call (*discrete*) *inverse wavelet transform of w with n levels of resolution*, and denote $\mathrm{W}^{-n}(w)$, the vector $\mathrm{W}^{m-n}(v)$.

3 Construction of the matricial method

This section presents a generalization of the constructed methods to matrices.

3.1 The bidimensional multiresolution analysis

The multiresolution used up to this point can be considered unidimensional (1D-MRA, from now on) due to the objects involved, e.g., functions over $\mathbb{R}$, vectors, etc. In order to be able to generalize the methods derived from it to matrices it is necessary to count with a bidimensional version (2D-MRA, from now on).

The *space of square integrable functions on* $\mathbb{R}^2$, denoted $\mathscr{L}^2(\mathbb{R}^2)$, is defined by

$$\mathscr{L}^2(\mathbb{R}^2) = \left\{ f : \mathbb{R}^2 \to \mathbb{C} : \int \left| f(s,t) \right|^2 ds\, dt < \infty \right\}.$$

Definition 3.1. A (*bidimensional*) *multiresolution analysis of* $\mathscr{L}^2(\mathbb{R}^2)$ is a sequence $(\mathbb{V}_j)$ of closed subspaces that satisfies:

1. $\mathbb{V}_j \subset \mathbb{V}_{j-1}$ for $j \in \mathbb{Z}$,
2. $\overline{\bigcup \mathbb{V}_j} = \mathscr{L}^2(\mathbb{R}^2)$,
3. $\bigcap \mathbb{V}_j = \{0\}$,
4. $f(\bullet) \in \mathbb{V}_j$ if and only if $f(2^j \bullet) \in \mathbb{V}_0$ for $j \in \mathbb{Z}$, and
5. there exists $\Phi \in \mathbb{V}_0$ such that $(\Phi(\bullet - k))_{k \in \mathbb{Z}^2}$ is an orthonormal basis of $\mathbb{V}_0$.

Every $\mathbb{V}_j$ is called *scale space* and Φ is called *scale function* or *father wavelet*.

There exists a tool that allows the generation of a 2D-MRA starting from a 1D-MRA: The tensor product $\otimes$, defined for F and G closed subspaces of $\mathscr{L}^2(\mathbb{R})$ by

$$\mathrm{F} \otimes \mathrm{G} = \left\{ h(s,t) = f(s)\, g(t) : f \in \mathrm{F} \text{ and } g \in \mathrm{G} \right\}.$$

($\mathrm{F} \otimes \mathrm{G}$ is a closed subspace of $\mathscr{L}^2(\mathbb{R}^2)$.)

Let's consider a 1D-MRA (V_j) with scale function ϕ and associated wavelet ψ. We define

- $\mathbb{V}_j = \mathrm{V}_j \otimes \mathrm{V}_j$,
- $\mathbb{W}_j = (\mathrm{W}_j \otimes \mathrm{V}_j) \oplus (\mathrm{V}_j \otimes \mathrm{W}_j) \oplus (\mathrm{W}_j \otimes \mathrm{W}_j)$,
- $\Phi(s,t) = \phi(s)\, \phi(t)$,
- $\Psi^h(s,t) = \phi(s)\, \psi(t)$,

- $\Psi^v(s,t) = \psi(s)\,\phi(t)$,

- $\Psi^d(s,t) = \psi(s)\,\psi(t)$.

(The superindexes h, v and d stand for *horizontal*, *vertical* and *diagonal*, respectively, which reason will be seen later.) Defined this way, $(\mathbb{V}_j)$ is a 2D-MRA with scale function Φ, and wavelets Ψ^h, Ψ^v and Ψ^d.[9] The wavelet transform is defined for $f \in \mathscr{L}^2(\mathbb{R}^2)$ by

$$\mathrm{W}(f) : (j,k) \mapsto \Big(\langle f \mid \Psi^h_{j,k} \rangle, \langle f \mid \Psi^v_{j,k} \rangle, \langle f \mid \Psi^d_{j,k} \rangle \Big).$$

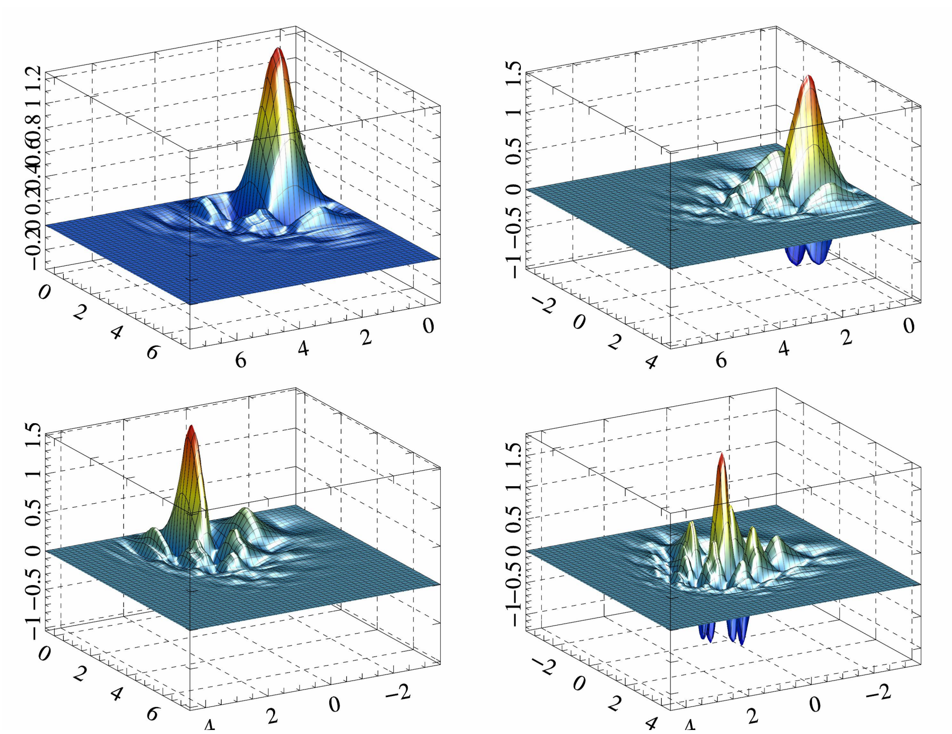

Figure 8: Bidimensional scale function and wavelets obtained from the Daubechies MRA with four vanishing moments. **Left to right and top to bottom:** Scale function Φ, and wavelets Ψ^h, Ψ^v and Ψ^d.

[9]According to a criterion which is not of our particular interest.

3.2 The matrix decomposition algorithm

3.2.1 Construction of the matrix decomposition algorithm. Let us adopt the following notation:

- $C^j = \big(\langle f \mid \Phi_{j,k}\rangle\big)$,
- $D^{h,j} = \big(\langle f \mid \Psi^h_{j,k}\rangle\big)$,
- $D^{v,j} = \big(\langle f \mid \Psi^v_{j,k}\rangle\big)$,
- $D^{d,j} = \big(\langle f \mid \Psi^d_{j,k}\rangle\big)$.

Let $A = (a_i^j)$ be a matrix with order $m_0 \times n_0$, where, without any loss of generality, we suppose that m_0 and n_0 are even. Following the same logics of the previous section, we define a signal f_A that characterizes the elements of A by letting

$$\boxed{f_A = \sum_{k_1}\sum_{k_2} a_{k_1 \bmod m_0}^{k_2 \bmod n_0}\, \Phi_{0,(k_1,k_2)}}. \tag{20}$$

Taking the scalar product with every $\Phi_{0,(k_1,k_2)}$ we get

$$C^0_{k_1,k_2} = a_{k_1 \bmod m_0}^{k_2 \bmod n_0}.$$

We can rewrite f_A in terms of Φ, Ψ^h, Ψ^v and Ψ^d. Indeed, due to the definition of Φ,

$$\begin{aligned} f_A &= \sum_{k_1}\sum_{k_2} C^0_{(k_1,k_2)}\, \Phi_{0,(k_1,k_2)} \\ &= \sum_{k_1}\sum_{k_2} C^0_{(k_1,k_2)} \big(\phi_{0,k_1}\phi_{0,k_2}\big) \\ &= \sum_{k_1}\Big(\sum_{k_2} C^0_{(k_1,k_2)}\, \phi_{0,k_2}\Big)\phi_{0,k_1}. \end{aligned} \tag{21}$$

For $k_1 \in \mathbb{Z}$ fixed, the expression in parenthesis defines a signal in $V_0 \subset \mathscr{L}^2(\mathbb{R})$, which we denote $f^r_{k_1}$—where the superindex r stands for *rows*—, that characterizes the vector $\big(C^0_{k_1,0},\dots,C^0_{k_1,n-1}\big)$, i.e., the k_1-th row of A. We use the 1D-MRA to decompose this signal:

$$f^r_{k_1} = \sum_k \langle f^r \mid \phi_{1,k}\rangle\, \phi_{1,k} + \sum_l \langle f^r \mid \psi_{1,l}\rangle\, \psi_{1,l}. \tag{22}$$

Replacing (22) in (21),

$$\begin{aligned} f_A &= \sum_{k_1}\Big(\sum_k \langle f^r \mid \phi_{1,k}\rangle\, \phi_{1,k} + \sum_l \langle f^r \mid \psi_{1,l}\rangle\, \psi_{1,l}\Big)\phi_{0,k_1} \\ &= \sum_{k_1}\Big(\sum_k \langle f^r_{k_1} \mid \phi_{1,k}\rangle \phi_{1,k}\Big)\phi_{0,k_1} + \sum_{k_1}\Big(\sum_l \langle f^r_{k_1} \mid \psi_{1,l}\rangle\Big)\phi_{0,k_1} \\ &= \sum_k\Big(\sum_{k_1} \langle f^r_{k_1} \mid \phi_{1,k}\rangle \phi_{0,k_1}\Big)\phi_{1,k} + \sum_l\Big(\sum_{k_1} \langle f^r_{k_1} \mid \psi_{1,l}\rangle \phi_{0,k_1}\Big)\psi_{1,l}. \end{aligned} \tag{23}$$

These last expressions in parenthesis define new signals, the first denoted f_k^c—where the superindex c stands for *columns*—, and the second denoted g_l^c, characterizing the elements of the vectors

$$\big(\langle f_0^r \mid \phi_{1,k}\rangle, \dots, \langle f_{m_0-1}^r \mid \phi_{1,k}\rangle\big),$$
$$\big(\langle f_0^r \mid \psi_{1,k}\rangle, \dots, \langle f_{m_0-1}^r \mid \psi_{1,k}\rangle\big),$$

respectively, i.e., f_k^c characterize the k-th column of the matrix resulting from the decomposed rows of A, where $0 \le k < n_0/2$, while g_l^c characterizes the l-th column of the same matrix, where $n_0/2 \le l < n_0$. Decomposing these signals, we get

$$f_k^c = \sum_p \langle f_k^r \mid \phi_{1,p}\rangle\, \phi_{1,p} + \sum_q \langle f_k^r \mid \psi_{1,q}\rangle\, \psi_{1,q}, \tag{24}$$
$$g_l^c = \sum_r \langle g_l^c \mid \phi_{1,r}\rangle\, \phi_{1,r} + \sum_s \langle g_l^c \mid \psi_{1,s}\rangle\, \psi_{1,s}. \tag{25}$$

Replacing (24) and (25) in (23),

$$f_{\mathrm{A}} = \sum_k \sum_p \langle f_k^c \mid \phi_{1,p}\rangle\, \phi_{1,p}\, \phi_{1,k} + \sum_k \sum_q \langle f_k^c \mid \psi_{1,q}\rangle\, \psi_{1,q}\, \phi_{1,k} + \sum_l \sum_r \langle g_l^c \mid \phi_{1,r}\rangle\, \phi_{1,r}\, \psi_{1,l} + \sum_l \sum_s \langle g_l^c \mid \psi_{1,s}\rangle\, \psi_{1,s}\, \psi_{1,l}.$$

Finally, remembering the definitions of Φ, Ψ^h, Ψ^v and Ψ^d, we get

$$f_{\mathrm{A}} = \sum_k \sum_p \langle f_k^c \mid \phi_{1,p}\rangle\, \Phi_{1,(p,k)} + \sum_k \sum_q \langle f_k^c \mid \psi_{1,q}\rangle\, \Psi^h_{1,(k,q)} + \sum_l \sum_r \langle g_l^c \mid \phi_{1,r}\rangle\, \Psi^v_{1,(l,r)} + \sum_l \sum_s \langle g_l^c \mid \psi_{1,s}\rangle\, \Psi^d_{1,(l,s)}.$$

Although this procedure appears to be very complicated, is actually very simple: Since the signals $f_{k_1}^r$ defined by the parenthesis in (21) characterize the rows of the matrix A, the decomposition made in (22) is equivalent to applying the vectorial wavelet transform to these rows; on the other hand, since the signals f_k^c and g_l^c defined by the parenthesis in (23) characterize the columns of the matrix resulting from the decomposed rows, the decomposition made in (24) and (25) is equivalent to apply the vectorial wavelet transform to these columns. As before, the periodicity guaranties that the quantity of information that needs to be stored is constant and equal to $m_0 n_0$. So the relevant (non-redundant) information can be stored in a matrix of the same dimensions as A, as indicated in fig. 9.

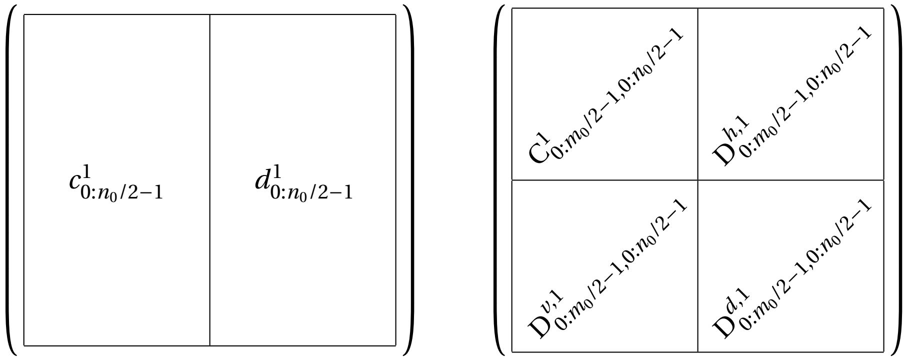

Figure 9: The way to store the relevant information from the wavelet transform of a matrix. **Left to right:** The information of the vectorial transform applied to the rows, and the information of the first iteration (after applying the vectorial transform to the columns of the matrix on the left).

The second iteration of this procedure consists simply on decomposing the left superior fourth part of the matrix obtained during the first iteration; the result can be stored in the corresponding submatrix. We have thus obtained the pyramidal decomposition for matrices, which is schematized in fig. 10.

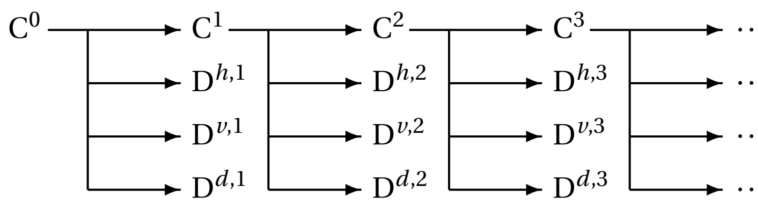

Figure 10: The classical representation of the pyramidal decomposition for matrices.

3.2.2 The wavelet transform of a matrix.

Definition 3.2. Let A be a matrix with convenient dimensions $m_0 \times n_0$. We call (*discrete*) *wavelet transform of* A *with n levels of resolution*, and denote $W^n(A)$, the matrix

$$\begin{pmatrix} C^n_{0:m_0/2^n-1,0:n_0/2^n-1} & D^{h,n}_{0:m_0/2^n-1,0:n_0/2^n-1} & \cdots & D^{h,1}_{0:m_0/2-1,0:n/2-1} \\ D^{v,n}_{0:m_0/2^n-1,0:n_0/2^n-1} & D^{d,n}_{0:m_0/2^n-1,0:n_0/2^n-1} & & \\ & \vdots & \ddots & \\ & D^{v,1}_{0:m_0/2-1,0:n_0/2-1} & & D^{d,1}_{0:m_0/2-1,0:n_0/2-1} \end{pmatrix}.$$

When $n = 1$ we can make omission and write simply $W(A)$. We also agree that $W^0(A) = A$.

Example 3.3. Let us consider

$$(h_k) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)', \quad (g_k) = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)', \quad \mathrm{A} = \begin{pmatrix} 64 & 2 & 3 & 61 \\ 9 & 55 & 54 & 12 \\ 17 & 47 & 46 & 20 \\ 40 & 26 & 27 & 37 \end{pmatrix}.$$

We have the following:

- Wavelet transform of A with one level of resolution: $\mathrm{W}(\mathrm{A}) = \begin{pmatrix} 65 & 65 & 8 & -8 \\ 65 & 65 & -8 & 8 \\ 1 & -1 & 54 & -50 \\ -1 & 1 & -22 & 18 \end{pmatrix}$,
- wavelet transform of A with two levels of resolution: $\mathrm{W}^2(\mathrm{A}) = \begin{pmatrix} 130 & 0 & 8 & -8 \\ 0 & 0 & -8 & 8 \\ 1 & -1 & 54 & -50 \\ -1 & 1 & -22 & 18 \end{pmatrix}$.

Example 3.4. A more graphical idea of the effect of the wavelet transform is presented in fig. 11.

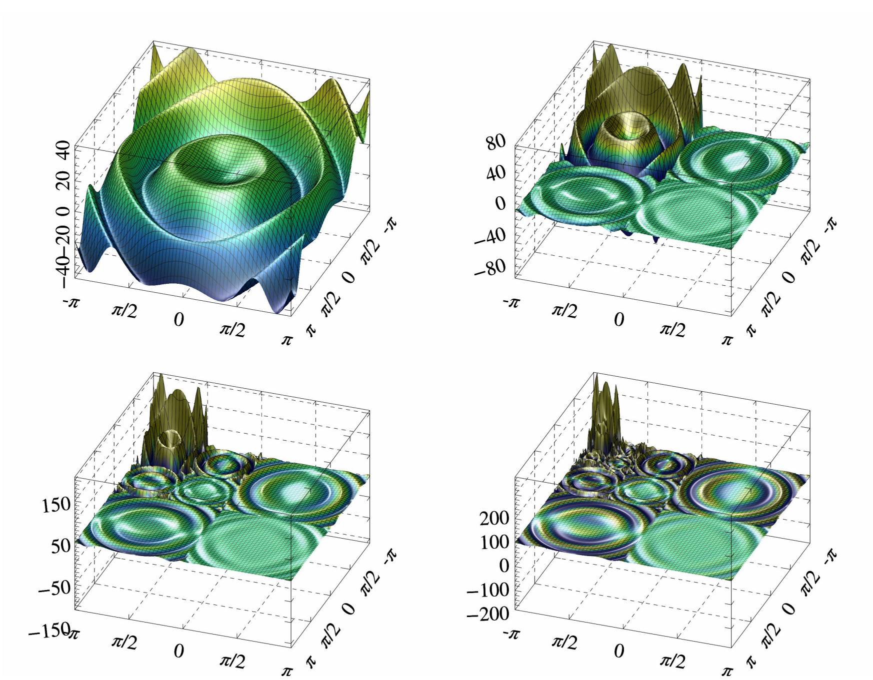

Figure 11: Wavelet transform of a discrete sampling of the signal $f(s,t) = \frac{3}{2}\sin(s^2+t^2) - \frac{2}{15}(s^3+t^3)$. **Left to right and top to bottom:** Zero, one, two and three levels of resolution.

Figure 12 shows the effect of the wavelet transform on an image (a matrix of colors). If we pay attention to the result of the first iteration, we can observe that the original image has been reduced by a factor of four and has been confined to the left superior section of the matrix; this section corresponds to the coefficients of the sequence C^1. The rest of the image has become dark; this section of the matrix corresponds to the coefficients of the sequences $D^{h,1}$, $D^{v,1}$ and $D^{d,1}$. The reason for them to be dark is that they contain coefficients close to zero, which is the number assigned to color black on the chosen scale.

Figure 12: Wavelet transform of an image using the MRA of Daubechies with four vanishing moments. **Left to right and top to bottom:** Original image, wavelet transform by rows, and wavelet transforms with one and two levels of resolution.

However, they are not completely black. Because of the way the wavelets Ψ^h, Ψ^v and Ψ^d have been defined, the bidimensional wavelet transform analyses an image in three different directions: horizontal, vertical and diagonal (that is where they get their superindexes). As a result, a sort of "directional border detection" is performed, as seen on fig. 13: The wavelet transform detects an abrupt change of color in the direction of analysis.

Figure 13: Detection of borders with the wavelet transform. The dark zones have been enhanced. **Left to right:** Original image and wavelet transform with one level of resolution.

3.3 The matrix reconstruction algorithm

Although it is completely possible to formally deduce the inversion algorithm of the matrix wavelet transform, it is completely unnecessary since the procedure is very clear: Apply the vector inverse wavelet transform, first to columns and the to rows.

3.3.1 The inverse wavelet transform of a matrix.

Definition 3.5. Let B be a matrix of the form $\mathrm{B} = \mathrm{W}^m(\mathrm{A})$ for some $m \in \mathbb{N}$ and some matrix A with convenient dimensions, and let $n \in \mathbb{N}$ such that $n \leq m$. We call (*discrete*) *inverse wavelet transform of* B *with* n *levels of resolution*, and denote $\mathrm{W}^{-n}(\mathrm{B})$, the matrix $\mathrm{W}^{m-n}(\mathrm{A})$.

4 The matricial properties of the wavelet transform

This section documents the properties of the methods we have built from the point of view of linear algebra. The main result is that the wavelet transform and its inverse can be written as matrix products. From this property we will derive others such as linearity, distributivity with respect to multiplication, isometry, etc. This section's material largely generalizes the results and proofs presented in [Beatty, 2004].

4.1 The wavelet transform as a matrix product

Definition 4.1. Let $n_0, n \in \mathbb{N}$ and $j \in \{1, \ldots, n\}$. We define the matrix $\Lambda^j_{n_0,n}$ by

$$\Lambda^j_{n_0,n} = \begin{pmatrix} \Lambda_{p_j \times q_j} & \mathrm{O}_{p_j \times r_j} \\ \mathrm{O}_{r_j \times q_j} & \mathrm{I}_{r_j} \end{pmatrix},$$

where

- $p_j = \begin{cases} q_j/2 + n_0/2^j & \text{if } j < n, \\ n_0/2^{n-1} & \text{if } j = n; \end{cases}$

- $q_j = \min\left\{r \in \mathbb{N} : r n_0/2^{n-1} \geq |(h_k)|\right\} n_0/2^{j-1}$;

- $r_j = \sum_{i=1}^{j-1} n_0/2^i = n_0\left(1 - \frac{1}{2^{j-1}}\right)$, where we adopt the convention that $\sum_{i=n_1}^{n_2} a_i = 0$ if $n_2 < n_1$;

- $\Lambda_{p_j \times q_j} = (\lambda_{i,k})$ is the matrix defined by

$$\lambda_{i,k} = \begin{cases} \overline{h_{(k-2i) \bmod q_j}} & \text{if } i < p_j - n_0/2^j, \\ \overline{g_{(k-2(i-p_j+n_0/2^j)) \bmod q_j}} & \text{if } i \geq p_j - n_0/2^j, \end{cases}$$

  for $i \in \{0,\ldots,p_j-1\}$, $k \in \{0,\ldots,q_j-1\}$, and some scale filter h and associated wavelet filter g;

- $\mathrm{O}_{n_1 \times n_2}$ is the matrix of order $n_1 \times n_2$, and

- I_{r_j} is the identity matrix of order $r_j \times r_j$.

It is impossible to clearly justify this definition right now, but the following example shows that these matrices are easily constructed.

Example 4.2. Let $(h_k) = (h_0,\ldots,h_3)'$. For $\Lambda^2_{8,3}$ we have $p_2 = 6$, $q_2 = 8$, $r_2 = 4$ and

$$\Lambda_{6\times 8} = \begin{pmatrix} \overline{h_0} & \overline{h_1} & \overline{h_2} & \overline{h_3} & \overline{h_4} & \overline{h_5} & \overline{h_6} & \overline{h_7} \\ \overline{h_6} & \overline{h_7} & \overline{h_0} & \overline{h_1} & \overline{h_2} & \overline{h_3} & \overline{h_4} & \overline{h_5} \\ \overline{h_4} & \overline{h_5} & \overline{h_6} & \overline{h_7} & \overline{h_0} & \overline{h_1} & \overline{h_2} & \overline{h_3} \\ \overline{h_2} & \overline{h_3} & \overline{h_4} & \overline{h_5} & \overline{h_6} & \overline{h_7} & \overline{h_0} & \overline{h_1} \\ \overline{g_0} & \overline{g_1} & \overline{g_2} & \overline{g_3} & \overline{g_4} & \overline{g_5} & \overline{g_6} & \overline{g_7} \\ \overline{g_6} & \overline{g_7} & \overline{g_0} & \overline{g_1} & \overline{g_2} & \overline{g_3} & \overline{g_4} & \overline{g_5} \end{pmatrix}$$

Definition 4.3. Let $n \in \mathbb{N}$ and $v = (v_0,\ldots,v_{n_0-1})'$ a vector with convenient length. We define the vector $\xi_n(v)$ by

$$[\xi_n(v)]_i = v_{i \bmod n_0}$$

for $i \in \{0,\ldots,r n_0 - 1\}$, where $r = \min\left\{m \in \mathbb{N} : m n_0/2^{n-1} \geq |(h_k)|\right\}$.

Example 4.4. Let $v = (v_0,\ldots,v_3)'$, $(h_k) = (h_0,\ldots,h_5)'$ and $n = 2$. Then $r = 3$ and

$$\xi_2(v) = (v_0,\ldots,v_3,v_0,\ldots,v_3,v_0,\ldots,v_3)'.$$

Remark. If we define the matrix $\mathbb{1}_{p,q}$, for all $p, q \in \mathbb{N}$, by

$$\mathbb{1}_{p,q} = \begin{pmatrix} \mathrm{I}_q \\ \vdots \\ \mathrm{I}_q \end{pmatrix}, \tag{26}$$

where there are exactly p identity matrices, then we can write

$$\xi_n(v) = \mathbb{1}_{r,n_0}\, v,$$

where r and n_0 are the same as in **definition 4.3**

The idea behind **definition 4.1** and **definition 4.3** is that the matrix $\Lambda^j_{n_0,n}$ corresponds to the jth iteration of the wavelet transform, while the vector $\xi_n(v)$ corresponds to the sequence c^0, such that

$$\begin{aligned} \mathrm{W}(v) &= \Lambda^1_{n_0,1}\,\xi_1(v), \\ \mathrm{W}^2(v) &= \Lambda^2_{n_0,2}\,\Lambda^1_{n_0,2}\,\xi_2(v), \\ \mathrm{W}^3(v) &= \Lambda^3_{n_0,3}\,\Lambda^2_{n_0,3}\,\Lambda^1_{n_0,3}\,\xi_3(v), \end{aligned}$$

and so on. (This result is formally established in **theorem 4.5**.) Here, the role of matrix $\Lambda^j_{n_0,n}$ is to compute exactly $p_j - n_0/2^j$ elements of the sequence c^j and exactly $n_0/2^j$ elements of the sequence d^j, while keeping constant the elements of the sequences $d^{j-1}, \dots, d^1$ computed by the previous matrices. For $j \in \{1, \dots, n-1\}$, $p_j - n_0/2^j \geq |(h_k)|$, so there are enough elements of c^j to compute c^{j+1}; for $j = n$, $p_j - n_0/2^j = n_0/2^n$, so the last matrix computes only the needed (non redundant) elements of c^n. On the other hand, the role of vector $\xi_n(v)$ is to contain enough elements of c^0—by repeating v a given number of times—to start the process.

Theorem 4.5. *Let $n \in \mathbb{N}$ and v be a vector with convenient length n_0. The wavelet transform of v with n levels of resolution can be written as a matrix product:*

$$\boxed{\mathrm{W}^n(v) = \prod_{j=n}^{1} \Lambda^j_{n_0,n}\,\xi_n(v)}.$$

Proof. Let $j \in \{1, \dots, n\}$, and p_j and q_j be as in **definition 4.1**. Since $q_j \geq |(h_k)|$, for $0 \leq i < p_j - n_0/2^j$,

$$c^j_i = \sum_{k=2i}^{2i+q_j-1} \overline{h_{k-2a}}\, c^{j-1}_{k \bmod q_j}.$$

By an adequate reordering,

$$\begin{aligned} c_i^j &= \sum_{b=0}^{q_j-1} \overline{h_{(k-2i) \bmod q_j}}\, c_k^{j-1} \\ &= \sum_{b=0}^{q_j-1} \lambda_{i,k}\, c_k^{j-1}. \end{aligned}$$

A similar reasoning leads to

$$d_i^j = \sum_{k=0}^{q_j-1} \overline{g_{(k-2i) \bmod q_j}}\, c_k^{j-1}$$

for $0 \le i < n_0/2^j$, or, by an adequate change of variable,

$$\begin{aligned} d_{i-p_j+n_0/2^j}^j &= \sum_{k=0}^{q_j-1} \overline{g_{(k-2(i-p_j+n_0/2^j)) \bmod q_j}}\, c_k^{j-1} \\ &= \sum_{k=0}^{q_j-1} \lambda_{i,k}\, c_k^{j-1} \end{aligned}$$

for $p_j - n_0/2^j \le i < p_j$. We conclude that

$$\left(c_{0:p_j-n_0/2^j-1}^j, d_{0:n_0/2^j-1}^j\right)' = \Lambda_{p_j \times q_j}\, c_{0:q_j-1}^{j-1}.$$

Therefore, since $q_1 = |\xi_n(v)|$, and, for $j \in \{2,\dots,n\}$, $q_j = q_{j-1}/2 = p_{j-1} - n_0/2^{j-1}$, then

$$\mathrm{W}^n(v) = \prod_{j=n}^{1} \xi_n(v).$$

□

Definition 4.6. Let $n \in \mathbb{N}$ and $\mathrm{A} = (a_{i,j})$ be a matrix with convenient order $m_0 \times n_0$. We define the matrix $\xi_n(\mathrm{A})$ by

$$[\xi_n(\mathrm{A})]_i^j = a_{i \bmod m_0}^{j \bmod n_0}$$

for $i \in \{0,\dots,rm_0-1\}$ and $j \in \{0,\dots,sn_0-1\}$, where

$$r = \min\left\{m \in \mathbb{N} : mm_0/2^{n-1} \ge |(h_k)|\right\}$$

and

$$s = \min\left\{m \in \mathbb{N} : mn_0/2^{n-1} \ge |(h_k)|\right\}.$$

Remark. Seen as a function, ξ_n operates on a matrix doing to its rows and columns the same it does to a vector, so, for every matrix A with order $m_0 \times n_0$, we can write

$$\xi_n(\mathrm{A}) = \mathbb{1}_{r,m_0}\,\mathrm{A}\,\mathbb{1}_{s,n_0}{}',$$

where r and s are as in **definition 4.6**

Theorem 4.7. *Let $n \in \mathbb{N}$ and A be a matrix with convenient order $m_0 \times n_0$. The wavelet transform of A with n levels of resolution can be written as a matrix product:*

$$\boxed{\mathrm{W}^n(\mathrm{A}) = \prod_{j=n}^{1} \Lambda^j_{m_0,n}\,\xi_n(\mathrm{A}) \prod_{j=1}^{n} \Lambda^j_{n_0,n}{}'}.$$

Proof. Since the wavelet transform of a matrix corresponds to just computing the vectorial transform of its rows and columns, this result follows from **theorem 4.5**. □

Corollary 4.8. *The wavelet transform of a matrix with n levels of resolution can be computed as a sequence of n iterations of the vectorial transform on its rows and n iterations on its columns, independently on how the first ones are combined with the second ones.*

Proof. This is a direct consequence of **theorem 4.7**, associativity of matrix multiplication and the fact that every multiplication by a matrix $\Lambda^j_{m,n}$ is equivalent to the jth iteration of the wavelet transform. □

Proposition 4.9. *The wavelet transform is a linear function of vectors (alternatively, matrices).*

Proof. This is evident from **theorem 4.5** and **theorem 4.7**, and from the linearity of ξ_n, seen as a function of vectors (alternatively, matrices). □

4.2 The inverse wavelet transform as a matrix product

Definition 4.10. Let $n_0, n \in \mathbb{N}$ and $j \in \{1, \ldots, n\}$. We define the matrix $\Omega^j_{n_0,n}$ by

$$\Omega^j_{n_0,n} = \begin{pmatrix} \Omega_{p_j \times q_j} & \mathrm{O}_{p_j \times r_j} \\ \mathrm{O}_{r_j \times q_j} & \mathrm{I}_{r_j} \end{pmatrix},$$

where

- $p_j = \begin{cases} n_0 & \text{if } j = 1, \\ q_j & \text{if } j > 1; \end{cases}$
- $q_j = \min\left\{r \in \mathbb{N} : r\,n_0/2^{n-1} \geq |(h_k)|\right\} n_0/2^{j-1};$

- $r_j = \sum_{i=1}^{j-1} 2^i q_j = q_j\left(2^{j-1}-1\right)$, where we adopt the convention that $\sum_{i=n_1}^{n_2} a_i = 0$ if $n_2 < n_1$;

- $\Omega_{p_j \times q_j} = (\omega_{i,k})$ is the matrix defined by

$$\omega_{i,k} = \begin{cases} h_{(i-2k) \bmod q_j} & \text{if } k < q_j/2, \\ g_{(i-2(k-q_j/2)) \bmod q_j} & \text{if } k \ge q_j/2, \end{cases}$$

  for $i \in \{0,\dots,p_j-1\}$, $k \in \{0,\dots,q_j-1\}$, and some scale filter h and associated wavelet filter g;

- $O_{r\times s}$ is the matrix with order $r \times s$, and

- I_{r_j} is the identity matrix with order $r_j \times r_j$.

Once more, an example will show how easy is to construct these matrices.

Example 4.11. Let $(h_k) = (h_0,\dots,h_3)'$. For $\Omega^2_{8,3}$ we have $p_2 = 8$, $q_2 = 8$, $r_2 = 8$ and

$$\Omega_{8\times 8} = \begin{pmatrix} h_0 & h_6 & h_4 & h_2 & g_0 & g_6 & g_4 & g_2 \\ h_1 & h_7 & h_5 & h_3 & g_1 & g_7 & g_5 & g_3 \\ h_2 & h_0 & h_6 & h_4 & g_2 & g_0 & g_6 & g_4 \\ h_3 & h_1 & h_7 & h_5 & g_3 & g_1 & g_7 & g_5 \\ h_4 & h_2 & h_0 & h_6 & g_4 & g_2 & g_0 & g_6 \\ h_5 & h_3 & h_1 & h_7 & g_5 & g_3 & g_1 & g_7 \\ h_6 & h_4 & h_2 & h_0 & g_6 & g_4 & g_2 & g_0 \\ h_7 & h_5 & h_3 & h_1 & g_7 & g_5 & g_3 & g_1 \end{pmatrix}.$$

Definition 4.12. Let $n \in \mathbb{N}$ and $w = (w_0,\dots,w_{n_0-1})'$ a vector with convenient length. We define the vector $\zeta_n(v)$ by

$$[\zeta_n(w)]_i = \begin{cases} w_{i \bmod n_0/2^n} & \text{if } 0 \le i < rn_0/2^n, \\ w_{i \bmod n_0/2^{n-p} + n_0/2^{n-p}} & \text{if } rn_0/2^{n-p} \le i < rn_0/2^{n-p-1},\ p \in \{0,\dots,n-1\}, \end{cases}$$

for $i \in \{0,\dots,rn_0-1\}$, where $r = \min\left\{m \in \mathbb{N} : mn_0/2^{n-1} \ge |(h_k)|\right\}$.

Remark. Notice that functions ξ_n and ζ_n are related by

$$\zeta_n(w) = \begin{pmatrix} \xi_0(w_{0:n_0/2^n-1}) \\ \xi_0(w_{n_0/2^n:n_0/2^{n-1}-1}) \\ \xi_1(w_{n_0/2^{n-1}:n_0/2^{n-2}-1}) \\ \xi_2(w_{n_0/2^{n-2}:n_0/2^{n-3}-1}) \\ \vdots \\ \xi_{n-1}(w_{n_0/2:n_0-1}) \end{pmatrix},$$

so we have

$$\zeta_n(v) = \operatorname{diag}\left(\mathbb{1}_{r,n_0/2^n}, \mathbb{1}_{r,n_0/2^n}, \mathbb{1}_{r,n_0/2^{n-1}}, \ldots, \mathbb{1}_{r,n_0/2}\right) v,$$

where diag indicates a block-wise diagonal matrix.

The idea behind **definition 4.10** and **definition 4.12** is that $\Omega^j_{n_0,n}$ inverts the jth iteration of the wavelet transform, while vector $\zeta_n(w)$ corresponds to $w = \mathrm{W}^n(v)$ for some $v \in \mathbb{C}^{n_0}$, such that

$$\begin{aligned}\mathrm{W}^{-1}(w) &= \Omega^1_{n_0,1}\,\zeta_1(w),\\ \mathrm{W}^{-2}(w) &= \Omega^1_{n_0,2}\,\Omega^2_{n_0,2}\,\zeta_2(w),\\ \mathrm{W}^{-3}(w) &= \Omega^1_{n_0,3}\,\Omega^2_{n_0,3}\,\Omega^3_{n_0,3}\,\zeta_3(w),\end{aligned}$$

and so on. (This result is formally established in **theorem 4.13**.) The role of vector $\zeta_n(v)$ is to contain enough elements of the sequences c^n, d^n, …, d^1, be repeating the corresponding parts of vector $w = \mathrm{W}^n(v)$ a given number of times.

The role of matrix $\Omega^j_{n_0,n}$ is to compute exactly p_j elements of the sequence c^{n-j} while keeping constant the elements corresponding to the sequences d^{n-j}, …, d^1.

Theorem 4.13. *Let $n \in \mathbb{N}$ and w be a vector with convenient length n_0. The inverse wavelet transform of w with n levels of resolution can be written as a matrix product:*

$$\boxed{\mathrm{W}^{-n}(w) = \prod_{j=1}^{n} \Omega^j_{n_0,n}\,\zeta_n(w)}.$$

Proof. Let $j \in \{1,\ldots,n\}$, and p_j and q_j be as in **definition 4.10**. Since $q_j \geq |(h_k)|$, for $0 \leq a < p_j$,

$$c^j_i = \sum_{k=\lfloor i/2\rfloor - q_j/2+1}^{\lfloor i/2\rfloor} h_{i-2k}\, c^{j+1}_{k \bmod q_j/2} + \sum_{k=\lfloor i/2\rfloor - q_j/2+1}^{\lfloor i/2\rfloor} g_{i-2k}\, d^{j+1}_{k \bmod q_j/2}.$$

By means of an adequate reordering,

$$c^j_a = \sum_{k=0}^{q_j/2-1} h_{(i-2k) \bmod q_j}\, c^{j+1}_k + \sum_{b=0}^{q_j/2-1} g_{(i-2k) \bmod q_j}\, d^{j+1}_k,$$

or, by a change of variable,

$$\begin{aligned}c^j_i &= \sum_{k=0}^{q_j/2-1} h_{(i-2k) \bmod q_j}\, c^{j+1}_k + \sum_{k=q_j/2}^{q_j-1} g_{(i-2(k-q_j/2)) \bmod q_j}\, d^{j+1}_k\\ &= \sum_{k=0}^{q_j/2-1} \omega_{i,k}\, c^{j+1}_k + \sum_{k=q_j/2}^{q_j-1} \omega_{i,k}\, d^{j+1}_k.\end{aligned}$$

We conclude that

$$c^{j}_{0:p_j-1} = \Omega_{p_j\times q_j}\left(c^{j+1}_{0:q_j/2-1}, d^{j+1}_{0:q_j/2-1}\right)'.$$

Therefore, since $q_n = \left|w_{0:rn_0/2^{n-1}-1} = \left(c^{n}_{0:rn_0/2^n-1}, d^{n}_{0:rn_0/2^n-1}\right)'\right|$, for $j \in \{2,\ldots,n\}$, and $q_{j-1} = 2q_j$, then

$$\mathrm{W}^{-n}(w) = \prod_{j=1}^{n} \Omega^{j}_{n_0,n}\zeta_n(w).$$

□

Definition 4.14. Let $n \in \mathbb{N}$ and $\mathrm{A} = (a_{i,j})$ a matrix with convenient order $m_0 \times n_0$. We define the matrix $\zeta_n(\mathrm{A})$ by

$$[\zeta_n(\mathrm{A})]^j_i = \begin{cases} a^{j \bmod n_0/2^n}_{i \bmod m_0/2^n} & \text{if } 0 \le i < rm_0/2^n \\ & \text{and } 0 \le j < sn_0/2^n; \\ a^{j \bmod n_0/2^{n-p}+n_0/2^{n-p}}_{i \bmod m_0/2^n} & \text{if } 0 \le i < rm_0/2^n \\ & \text{and } sn_0/2^{n-p} \le j < sn_0/2^{n-p-1}, \\ & \text{for } p \in \{0,\ldots,n-1\}; \\ a^{j \bmod n_0/2^n}_{i \bmod m_0/2^{n-p}+\mathrm{M}/2^{n-p}} & \text{if } rm_0/2^{n-p} \le i < rm_0/2^{n-p-1} \\ & \text{and } 0 \le j < sn_0/2^n, \text{ for } p \in \{0,\ldots,n-1\}; \\ a^{j \bmod n_0/2^{n-q}+n_0/2^{n-q}}_{i \bmod m_0/2^{n-k}+m_0/2^{n-p}} & \text{if } rm_0/2^{n-p} \le i < rm_0/2^{n-p-1} \\ & \text{and } sn_0/2^{n-q} \le j < sn_0/2^{n-q-1}, \\ & \text{for } p,q \in \{0,\ldots,n-1\}, \end{cases}$$

where

$$r = \min\left\{m \in \mathbb{N} : mm_0/2^{n-1} \ge |(h_k)|\right\}$$

and

$$s = \min\left\{m \in \mathbb{N} : mn_0/2^{n-1} \ge |(h_k)|\right\}.$$

Remark. Seen as a function, ζ_n operates on A doing to its rows and columns exactly what it does to a vector, so, for every matrix A with convenient order $m_0 \times n_0$,

$$\zeta_n(\mathrm{A}) = \operatorname{diag}\left(\mathbb{1}_{r,m_0/2^n}, \mathbb{1}_{r,m_0/2^n}, \ldots, \mathbb{1}_{r,m_0/2}\right) \mathrm{A} \operatorname{diag}\left(\mathbb{1}_{s,n_0/2^n}, \mathbb{1}_{s,n_0/2^n}, \ldots, \mathbb{1}_{s,n_0/2}\right)'.$$

The following results have similar proofs to the ones of their counterparts in the precedent subsection.

Theorem 4.15. *Let $n \in \mathbb{N}$ and A be a matrix with convenient order $m_0 \times n_0$. The inverse wavelet transform of A with n levels of resolution can be written as a matrix product:*

$$\boxed{\mathrm{W}^{-n}(\mathrm{A}) = \prod_{j=1}^{n} \Omega_{m_0,n}^{j} \zeta_n(\mathrm{A}) \prod_{j=n}^{1} \left(\Omega_{n_0,n}^{j}\right)'}.$$

Corollary 4.16. *The inverse wavelet transform of a matrix with n levels of resolution can be computed as a sequence of n iterations of the vectorial inverse transform on its rows and n iterations on its columns, independently on how the first ones are combined with the second ones.*

Proposition 4.17. *The inverse wavelet transform is a linear function of vectors (alternatively, matrices).*

4.3 Other matricial properties

Much of the complexity of the definitions of matrices $\Lambda_{n_0,n}^{j}$ and $\Omega_{n_0,n}^{j}$, and functions ξ_n and ζ_n comes from the need to obtain vectors and matrices "big enough" for the respective calculations. For $n \in \mathbb{N}$ and a vector v with convenient length n_0, if

$$n_0/2^{n-1} \geq |(h_k)|, \tag{27}$$

then $\xi_n(v) = \zeta_n(v) = v$, and matrices $\Lambda_{n_0,n}^{j}$ and $\Omega_{n_0,n}^{j}$ become much simpler. Same happens for a matrix with convenient order $m_0 \times n_0$ such that

$$m_0/2^{n-1} \geq |(h_k)| \quad \text{and} \quad n_0/2^{n-1} \geq |(h_k)|. \tag{28}$$

This subsection studies the consequences of these conditions.

Proposition 4.18. *Let $n \in \mathbb{N}$ and $n_0 \in \mathbb{N}$ be a convenient number. If (27) holds, then*

$$\boxed{(\Lambda_{n_0,n}^{j})^{-1} = \Omega_{n_0,n}^{j}}$$

for $j \in \{1, \ldots, n\}$.

Proof. We prove by induction on n. First, we notice that condition (27) implies that $\Lambda_{n_0,n-1}^{j} = \Lambda_{n_0,n}^{j}$ and $\Omega_{n_0,n-1}^{j} = \Omega_{n_0,n}^{j}$ for $j \in \{1, \ldots, n-1\}$.

By **theorem 4.5** and **theorem 4.13**, for every vector $v \in \mathbb{C}^{n_0}$,

$$\begin{aligned} v &= \mathrm{W}\big(\mathrm{W}^{-1}(v)\big) \\ &= \mathrm{W}\big(\Omega_{n_0,1}^{1}\, v\big) \\ &= \Lambda_{n_0,1}^{1}\, \Omega_{n_0,1}^{1}\, v, \end{aligned}$$

so $\Lambda_{n_0,1}^{1}\, \Omega_{n_0,1}^{1} = \mathrm{I}_{n_0}$. A similar procedure shows $\Omega_{n_0,1}^{1}\, \Lambda_{n_0,1}^{1} = \mathrm{I}_{n_0}$.

Let $n > 1$. Once again, by **theorem 4.5** and **theorem 4.13**, for every $v \in \mathbb{C}^{n_0}$,

$$\begin{aligned} v &= \mathrm{W}^n\big(\mathrm{W}^{-n}(v)\big) \\ &= \mathrm{W}^n\left(\prod_{j=1}^{n} \Omega_{n_0,n}^{j}\, v\right) \\ &= \prod_{j=n}^{1} \Lambda_{n_0,n}^{j} \prod_{j=1}^{n} \Omega_{n_0,n}^{j}\, v. \end{aligned}$$

By induction hypothesis, $\prod_{j=n-1}^{1} \Lambda_{n_0,n}^{j} \prod_{j=1}^{n-1} \Omega_{n_0,n}^{j} = \mathrm{I}_{n_0}$, so

$$v = \Lambda_{n_0,n}^{n}\, \Omega_{n_0,n}^{n}\, v.$$

We conclude that $\Lambda_{n_0,n}^{n}\, \Omega_{n_0,n}^{n} = \mathrm{I}_m$. A similar procedure shows $\Omega_{n_0,n}^{n}\, \Lambda_{n_0,n}^{n} = \mathrm{I}_m$. □

Corollary 4.19. *If n_0 is a even number such that $n_0 \geq |(h_k)|$, then*

$$\boxed{\sum_{i=0}^{n_0-1} \overline{[x]_i}\, [y]_{(i-2j) \bmod n_0} = \delta_{j,0}\, \delta_{x,y}},$$

for $j \in \{0, \dots, n_0/2-1\}$, where $x, y \in \{(h_k), (g_k)\}$ and δ is the Kronecker's delta function.

Proof. The result is obtained by multiplying

$$\Lambda_{n_0,1}^{1}\, \Omega_{n_0,n}^{1} = \mathrm{I}_{n_0}.$$

□

Proposition 4.20. *Let $n \in \mathbb{N}$ and $n_0 \in \mathbb{N}$ be a convenient number. If (27) holds, then*

$$\boxed{\overline{\left(\Lambda_{n_0,n}^{j}\right)}' = \Omega_{n_0,n}}$$

for $j \in \{1, \dots, n\}$.

Proof. This result is evident from **definition 4.1** and **definition 4.10**. □

Corollary 4.21. *Let $n \in \mathbb{N}$ and $n_0 \in \mathbb{N}$ be a convenient number. If (27) holds, then $\Lambda_{n_0.n}^{j}$ (alternatively, $\Omega_{n_0,n}^{j}$) is a unitary matrix.*

Proof. Direct consequence of **proposition 4.18** and **proposition 4.20**. □

The following result is presented in [Beatty, 2004], but it is proved there only for Haar wavelets, and for very particular versions of the wavelet transform, based on a couple of ideas called *multiresolution expansion* and ψ^n*-expansion*, which are analogous to multiresolution analysis. The proof presented here is not only more general, simpler and more elegant, but it also implies that result.

Proposition 4.22. *Let $n \in \mathbb{N}$, and A and B matrices with convenient orders $m_0 \times p_0$ and $p_0 \times n_0$, respectively. If the scale filter (h_k) is real and $q/2^{n-1} \geq |(h_k)|$ for $1 < q \in \{m_0, n_0, p_0\}$, then*

1. *the wavelet transform is distributive with respect to matrix multiplication:*

$$\boxed{\mathrm{W}^n(\mathrm{AB}) = \mathrm{W}^n(\mathrm{A})\,\mathrm{W}^n(\mathrm{B})};$$

2. *the inverse wavelet transform is distributive with respect to matrix multiplication:*

$$\boxed{\mathrm{W}^{-n}(\mathrm{AB}) = \mathrm{W}^{-n}(\mathrm{A})\,\mathrm{W}^{-n}(\mathrm{B})}.$$

Proof.

1. By **theorem 4.7** and associativity of matrix multiplication,

$$\mathrm{W}^n(\mathrm{A})\,\mathrm{W}^n(\mathrm{B}) = \prod_{j=n}^{1} \Lambda^j_{m_0,n}\,\mathrm{A}\left(\prod_{j=1}^{n} \left(\Lambda^j_{p_0,n}\right)' \prod_{j=n}^{1} \Lambda^j_{p_0,n}\right)\mathrm{B}\prod_{j=1}^{n} \left(\Lambda^j_{n_0,n}\right)'.$$

   Since the scale filter is real and by **corollary 4.21**, $\prod_{j=1}^{n} \left(\Lambda^j_{p_0,n}\right)' \prod_{j=n}^{1} \Lambda^j_{p_0,n} = \mathrm{I}_{n_0}$, so

$$\mathrm{W}^n(\mathrm{A})\,\mathrm{W}^n(\mathrm{B}) = \prod_{j=n}^{1} \Lambda^j_{m_0,n}\,\mathrm{AB}\prod_{j=1}^{n} \left(\Lambda^j_{n_0,n}\right)'.$$

   Therefore, once again by **theorem 4.7**,

$$\mathrm{W}^n(\mathrm{A})\,\mathrm{W}^n(\mathrm{B}) = \mathrm{W}^n(\mathrm{AB}).$$

2. The proof of the second part is similar.

□

The following result is a discrete equivalent of **corollary 1.8**

Corollary 4.23. *Let $n \in \mathbb{N}$ and $n_0 \in \mathbb{N}$ be a convenient number. If the scale filter (h_k) is real and (27) holds, then the wavelet transform with n levels of resolution and its inverse are isometries on $\mathbb{R}^{n_0}$, i.e.,*

$$\boxed{\langle \mathrm{W}^n(u) \mid \mathrm{W}^n(v)\rangle = \langle u \mid v\rangle},$$

and

$$\boxed{\langle \mathrm{W}^{-n}(u) \mid \mathrm{W}^{-n}(v)\rangle = \langle u \mid v\rangle},$$

for $u, v \in \mathbb{R}^{n_0}$, where $\langle \bullet \mid \bullet \rangle$ stands for the euclidean product.

Proof. This follows directly from the fact that matrices $\Lambda^j_{n_0,n}$ and $\Omega^j_{n_0,n}$ are unitary (**corollary 4.21**). Alternatively, is a direct consequence of the fact that the euclidean product can be written as a matrix product and **proposition 4.22**. □

5 Concluding remarks

We have presented a comprehensive study of the wavelet transform and its inverse from the point of view of linear algebra, putting emphasis on practical and computational aspects.

First, we have studied the canonical theoretical foundations. From that starting point, we have constructed versions of the transform and its inverse that preserve storage space during computation, just by constructing periodical functions f_v and f_A, as defined in (14) and (20), that represent vectors v and matrices A, respectively. Finally, we have studied the linear properties of these tools.

The main results presented here are that the wavelet transform and its inverse can be written as matrix products. From that property we have derived others very important, like distributivity over multiplication, isometry, which are inherited from their continuous version.

6 Acknowledgements

This work is derived from the author's undergraduate thesis [Sejas, 2012] as students at Universidad Mayor de San Simón. I am particularly thankful to my advisor, Santiago Relos Paco, for the mathematical guidance and support. Special thanks to my father, Gastón Sejas Cortés, who has provided valuable advise and moral support through the process of writing.

The graphics of this article were created with the MathGL library [Balakin, 2016], using the LaTeX package mglTeX [Sejas and Balakin, 2016], which was written by the author of this work. Some computations were worked out using the software Scilab v6.0.1, through the *Scilab Wavelet Toolbox* v0.3.1-3.

References

[Balakin, 2016] Balakin, A. A. (2016). *The* MathGL *library*. http://mathgl.sourceforge.net/doc_en/Main.html.

[Beatty, 2004] Beatty, D. (2004). *Applications of Wavelets to Image Processing and Matrix Multiplication.* Master thesis, Texas Tech University.

[Bultheel, 2006] Bultheel, A. (2006). Wavelets with Applications in Signal and Image Processing. Online. https://people.cs.kuleuven.be/~adhemar.bultheel/WWW/WAVE/alltxt.pdf.

[Daubechies, 1992] Daubechies, I. (1992). *Ten Lectures on Wavelets.* Number 61 in CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia.

[Gomes and Velho, 1999] Gomes, J. and Velho, L. (1999). From Fourier Analysis to Wavelets. Course Notes - SIGGRAPH 99, Instituto de Matemática Pura e Aplicada, IMPA, Rio de Janeiro, Brazil.

[Graps, 1995] Graps, A. (1995). An introduction to wavelets. *IEEE Computational Sciences and Engineering*, 2:50–61.

[Mallat, 1999] Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Elsevier (USA), second edition.

[Sejas, 2012] Sejas, D. (2012). *Sobre la Construcción y Análisis Matemático de la Transformada Wavelet y sus Propiedades Matriciales*. proyecto de grado para optar al diploma académico de licenciado en matemática, Universidad Mayor de San Simón.

[Sejas and Balakin, 2016] Sejas, D. and Balakin, A. (2016). *The mglTEX package*. www.ctan.org/pkg/mgltex. Current version: v4.2.

[Sossa, 2009] Sossa, D. A. (2009). *Una Introducción a la Teoría de las Wavelets y su Aplicación a la Compresión de Imágenes*. proyecto de grado para optar al diploma académico de licenciado en matemática, Universidad Mayor de San Simón.

Diego Sejas Viscarra
Universidad Mayor de San Simón
dsejas.math@pm.me