A short primer on wavelets, second edition

Torben Klint

June 2001

Abstract

Wavelet theory is a new discipline, which has evolved at a tremendous rate since the early eighties. It has been developed in close collaboration between mathematicians, physicists and engineers, and despite its young age, it has already had a significant impact on pure mathematics as well as applications in physics and engineering science. This short survey aims at presenting the basics of wavelet theory in a way that on one hand is brief and to the point, without entering discussions on the mathematical subtleties of the subject, but on the other hand explains wavelet theory in sufficient depth so that the reader becomes able to make his/her way through the wavelet literature.

The text is organized as follows. After a short introduction in section 1, sections 2 through 5 treat the basics of wavelet theory. Section 6 gives some examples of applications of wavelets, and section 7 gives sources of wavelet material.

1 Introduction

The bottom line of wavelet theory is to represent or approximate a given function by a linear combination of simpler functions. This theme is well known from Fourier analysis. Let f be a periodic function, let its period be T. Then f can be represented by its *Fourier series*, which is a linear combination of sines and cosines oscillating at integer multiples of the angular frequency $\omega = \frac{2\pi}{T}$ of f:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\omega x) + b_n \sin(n\omega x).$$

In practical applications, one cannot handle infinite sums, and hence one approximates f by finite sections of the Fourier series:

$$f(x) \simeq f_N(x) \equiv \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(n\omega x) + b_n \sin(n\omega x).$$

A terminology which is often used is to say that the function f is split up into 'building blocks', which in the above case are the sines and cosines, oscillating with the frequencies $n\omega$.

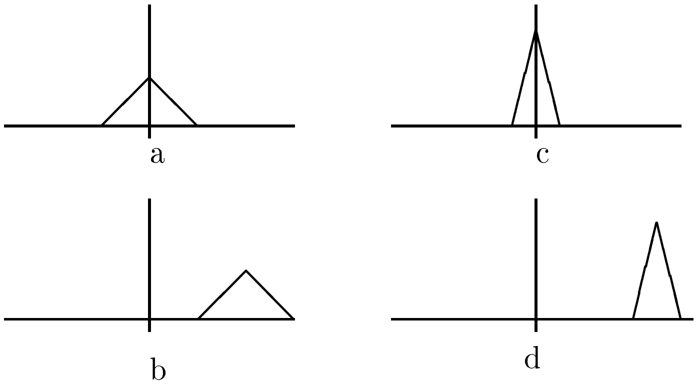


Figure 1: a): A function f. b): Translation. c): Dilation. d): Dilation and translation.

Now we leave the framework of periodic functions, and let ψ be a function defined on $\mathbb{R}$. We define the operators of *translation* T_k and *dilation* D_j as

$$T_k\psi(x) \equiv \psi(x-k), \quad \text{and } D_j\psi(x) \equiv 2^{j/2}\psi(2^j x)$$

where k and j are integers, called the *translation* and *dilation parameters*. The action of translation and dilation is depicted on figure 1. For a fixed scale parameter $j \in \mathbb{Z}$, we define the family of translates at scale j as

$$\psi_{j,k}(x) \equiv 2^{j/2}\psi(2^j x - k), \; k \in \mathbb{Z}, \tag{1}$$

Let f be a function that is to be approximated. This function could be, for instance, a speech signal, an image, or the solution to a differential or integral equation. In wavelet theory, f is approximated using the building blocks (1), and the representation of f has the form

$$f = \sum_{j,k \in \mathbb{Z}} d_{j,k}\psi_{j,k}.$$

In applications, f is approximated by a finite section of the above sum.

Hence, f is approximated by the building blocks consisting of translates and dilates of *one single function* ψ. If this function satisfies a number of requirements, it is called a *wavelet*. It turns out that this way of approximating functions yield a number of properties that are desirable for applications in various branches of engineering science.

2 Scaling functions, Multiresolution Analysis

As mentioned in the introduction, the bottom line of wavelet theory is to approximate functions in a special way. In order to make this precise, we have to specify the class of functions we want to approximate. The natural class of functions to consider in wavelet theory is $L_2(\mathbb{R})$, the class of square integrable functions $f : \mathbb{R} \mapsto \mathbb{C}$:

$$f \in L_2(\mathbb{R}) \;\Leftrightarrow \int_{\mathbb{R}} |f(x)|^2 dx < \infty.$$

One can also consider wavelet theory in the framework of functions being square integrable over some interval subset of $\mathbb{R}$, but since this subject is considerably more complicated than wavelet theory in $L_2(\mathbb{R})$, we postpone this discussion.

The function space $L_2(\mathbb{R})$ can be equipped with a *norm* and an *inner product*:

$$\|f\|_{L_2(\mathbb{R})} \equiv \sqrt{\int_{\mathbb{R}} |f(x)|^2 dx}$$

$$\langle f, g\rangle_{L_2(\mathbb{R})} \equiv \int_{\mathbb{R}} f(x)\overline{g(x)}dx$$

where the bar means complex conjugation. Using the inner product thus defined, we say that the $L_2(\mathbb{R})$ functions f and g are perpendicular if $\langle f, g\rangle_{L_2(\mathbb{R})} = 0$. A set of $L_2(\mathbb{R})$ functions $\{f_n\}_{n\in\mathbb{N}}$ is an *orthonormal system* if

$$\langle f_m, f_n\rangle = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}.$$

If furthermore $\{f_n\}$ has the property that any function f from a subspace $X \subseteq L_2(\mathbb{R})$ can be uniquely expressed as a (in general infinite) linear combination of the f_n, then we say that $\{f_n\}$ is an *orthonormal basis* for X.

This section is devoted to describing the concept of a Multiresolution Analysis, which we now define:

Definition 2.1 A Multiresolution Analysis (MRA) is a sequence $\{V_j\}$ of closed subspaces of $L_2(\mathbb{R})$, having the following properties:

(i) $\ldots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots$

(ii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L_2(\mathbb{R})$.

(iii) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$.

(iv) $f(x) \in V_j \Leftrightarrow f(2^{-j}) \in V_0 \; \forall j \in \mathbb{Z}$.

(v) $f(x) \in V_0 \Leftrightarrow f(x-k) \in V_0$ for all $k \in \mathbb{Z}$.

(vi) There exists a function $\phi \in L_2(\mathbb{R})$ with the property that the family of translates $\{\phi(x-k) : k \in \mathbb{Z}\}$ is an orthonormal basis for V_0. This function is called the *scaling function* for the MRA.

This definition deserves some remarks. The spaces $\{V_j\}$ should be thought of as a sequence of approximation spaces; in the sense that as the scale parameter j increases, the V_j are able to recapture an increasing amount of detail in any given $L_2(\mathbb{R})$ function f; the condition (ii) expresses that any $L_2(\mathbb{R})$-function f can be approximated arbitrarily well by choosing an element from a sufficiently large space V_j. The condition (iv) expresses that all the V_j spaces are nothing but dilated versions of the 'fundamental' space V_0. The condition (v) is called translation invariance. Together, (iv), (v) and (vi) imply that for any scale parameter j the family

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k),\ \ k \in \mathbb{Z}$$

is an orthonormal basis for V_j.

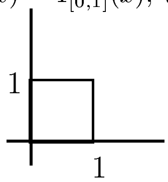
One can show that (i) is fulfilled if and only if the scaling function ϕ satisfies a *dilation equation*

$$\phi(x) = \sum_{n\in\mathbb{Z}} h_n \phi_{1,n}(x) = \sqrt{2}\sum_{n\in\mathbb{Z}} h_n \phi(2x-n). \tag{2}$$

Other names for the equation (2) are *refinement -, scaling -, and two-scale equation.* The coefficients h_n in (2) are called *filter coefficients* or the *mask* for the scaling function, and their role in wavelet theory and applications are paramount. For many wavelet families that are used in practical applications, one can say that the scaling function (and consequently the wavelet, see the next section) is constructed by determining a sequence of filter coefficients that satisfy some rather complicated conditions, which we shall not discuss here. The fundamental issue is that the scaling function (and the wavelet) is completely determined by the filter coefficients, and many applications of wavelet techniques involve only simple manipulations with the filter coefficients. Usually in wavelet books and articles, the filter coefficients of the scaling functions and wavelets considered there are given in the text, or in appendices.

Example 2.2 Consider the function $\phi(x) = 1_{[0,1]}(x)$, where

$$1_{[0,1]}(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}.$$



It is clear that the family of translates $\{\phi(x-k) : k \in \mathbb{Z}\}$ is an orthonormal system in $L_2(\mathbb{R})$. Approximating a function $f \in L_2(\mathbb{R})$ in the space V_0 spanned by translates of this square function amounts to approximating f by a piecewise constant function f_0, each piece has length 1. Approximating f in V_1 amounts to approximating f with a piecewise constant function, here each piece has length $\frac{1}{2}$. In general, approximating f in V_j amounts to approximating f with piecewise constant functions of which each piece has length 2^{-j}. As the scale parameter increases, we can recapture more detail in the function, and hence (ii) in the definition of a Multiresolution Analysis is fulfilled. The dilation equation for $\phi(x)$ reads

$$\phi(x) = \phi(2x) + \phi(2x-1) = \frac{1}{\sqrt{2}}\phi_{1,0}(x) + \frac{1}{\sqrt{2}}\phi_{1,1}(x).$$

3 Wavelets

Suppose that we have constructed a Multiresolution Analysis $\{V_j\}$ in $L_2(\mathbb{R})$, and equip $L_2(\mathbb{R})$ with the norm and inner product defined in the previous section. Henceforth, $\langle\cdot,\cdot\rangle$ always denotes the inner product in $L_2(\mathbb{R})$. Then we have that $V_j \subset V_{j+1}$ for any scale parameter $j \in \mathbb{Z}$. We define for every scale parameter $j \in \mathbb{Z}$ the *complement spaces* W_j through the relation

$$V_{j+1} = V_j \oplus W_j, \tag{3}$$

the interpretation of which we now describe. Let $f_{j+1} \in V_{j+1}$. Then f_{j+1} can be approximated by a function $f_j \in V_j$, but since V_j does not in general recapture all details of the space V_{j+1}, there are some details in f_{j+1} that is not contained in f_j. These missing details are contained in W_j, in the sense that we can find $g_j \in W_j$ so that

$$f_{j+1} = f_j + g_j, \tag{4}$$

and furthermore, f_j and g_j are perpendicular. Hence the complement space W_j should be thought of as a detail space that recaptures the amount of detail in V_{j+1} that V_j does not contain.

The question is now whether there exists a function ψ whose translates $\{\psi(x-k) : k \in \mathbb{Z}\}$ is an orthonormal basis for W_0, so that the complement spaces W_j has the same structure as the spaces V_j. The answer is affirmative, one choice of ψ is

$$\psi(x) = \sum_n (-1)^n h_{1-n}\phi_{1,n}(x) = \sqrt{2}\sum_n (-1)^n h_{1-n}\phi(2x-n) \tag{5}$$

where h_n are the filter coefficients for the scaling function ϕ. If the family of translates of ψ is an orthonormal basis for W_0, the function ψ is called a *wavelet*. The choice for constructing ψ such that the family of translates $\{\psi(x-k) : k \in \mathbb{Z}\}$ is an orthonormal basis for W_0 is not unique, but we shall stick to the above expression. The reader is referred to [1, ch. 5] for details.

The wavelet inherits the property (iv) in the definition of a Multiresolution Analysis, hence the family

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k),\ k \in \mathbb{Z}$$

is an orthonormal basis for W_j. The functions f_j and g_j in equation (4) can thus be expressed as

$$f_j = \sum_k c_{j,k}\phi_{j,k}, \text{ and } g_j = \sum_k d_{j,k}\psi_{j,k}$$

where $c_{j,k} = \langle f_j, \phi_{j,k} \rangle$ and $d_{j,k} = \langle f_j, \psi_{j,k} \rangle$. We shall refer to the $c_{j,k}$ as *scaling function coefficients*, and the $d_{j,k}$ as *wavelet coefficients.*

Choose now two integers $J > j_0$. Iteration of equation (3) becomes

$$V_J = V_{j_0} \oplus \bigoplus_{j=j_0}^{J-1} W_j, \tag{6}$$

in the sense that any $f_J \in V_J$ can be expressed as

$$f_J = f_{j_0} + \sum_{j=j_0}^{J-1} g_j \tag{7}$$

where

$$f_{j_0} \in V_{j_0},\ g_j \in W_j \text{ for } j = j_0, \ldots J-1,$$

and each of these functions are mutually perpendicular, and they can be represented in the orthonormal bases available in V_j and W_j. As j increases from j_0 to $J-1$, the g_j contain an increasing amount of detail of the function f_J. Using now equation (6) and (ii) in the definition of a Multiresolution Analysis we get

$$L_2(\mathbb{R}) = V_{j_0} \oplus \bigoplus_{j=j_0}^{\infty} W_j$$

and (iii) yields, when we let $j_0 \to \infty$,

$$L_2(\mathbb{R}) = \bigoplus_{j=-\infty}^{\infty} W_j.$$

This equation expresses that each $f \in L_2(\mathbb{R})$ can be expressed as

$$f = \sum_{j=-\infty}^{\infty} \sum_{k\in\mathbb{Z}} d_{j,k}\psi_{j,k}$$

where $d_{j,k} = \langle f, \psi_{j,k} \rangle$.

It is possible to design wavelets in order to meet requirements that might be desirable for a specific application. A property of the wavelet which is particularly desirable (in section 6 we shall see why) is the *property of vanishing moments.* We say that the wavelet ψ has n vanishing moments if

$$\int_{\mathbb{R}} x^l \psi(x) dx = 0,\ l = 0, \ldots n.$$

Another desirable property of the scaling function and wavelet is *compact support.* The support of a function f is defined as

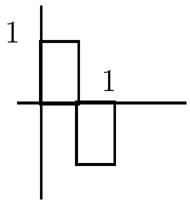
$$\text{supp} f \equiv \overline{\{x \in \mathbb{R} : f(x) \neq 0\}}.$$

A subset of $\mathbb{R}$ is compact if and only if it is closed and bounded, and since the support is defined as a closed set, a compactly supported function is a function that only takes on nonzero values on a bounded interval. In the design of wavelets, there is a tradeoff between various desirable properties of the scaling function and wavelet, which is described in the following diagram:

$$\begin{array}{ccc}
\text{Large number of} & & \text{Large support of wavelet} \\
\text{vanishing moments} & \Leftrightarrow & \text{and scaling function} \\
\Updownarrow & & \Updownarrow \\
\text{Many nonzero coefficients} & & \text{Smooth wavelet} \\
\text{in the dilation equation (2)} & \Leftrightarrow & \text{and scaling function}
\end{array}$$

Example 3.1 The wavelet corresponding to the scaling function $\phi(x) = 1_{[0,1]}(x)$ from example 2.2 is

$$\begin{aligned}
\psi(x) &\equiv \frac{1}{\sqrt{2}}\phi_{1,0}(x) - \frac{1}{\sqrt{2}}\phi_{1,1}(x) \\
&= \begin{cases} 1 & x \in [0, \frac{1}{2}[\\ -1 & x \in [\frac{1}{2}, 1[\end{cases}
\end{aligned}$$



This wavelet is called the *Haar wavelet*. The fact that translates and dilates of this function is an orthonormal basis for $L_2(\mathbb{R})$ was discovered in 1910, long before the concept of wavelets emerged.

As a final remark, we mention that there is a theorem by Daubechies [1, chapter 8] which states that symmetry, compact support, orthonormality of translates and real-valuedness of the scaling function and wavelet ϕ and ψ are restrictive requirements, in the sense that the only wavelet meeting these requirements is the Haar wavelet from example 3.1, and this family is not well suited for applications. If one wants compactly supported, real valued scaling function and wavelet whose translates are orthonormal, one must accept that these are nonsymmetric. We shall see examples of this below. On the other hand, if one wants symmetry of the wavelet, one has to relax at least one of the other requirements. One possibility is to relax the requirement that translates of ϕ resp. ψ should be orthonormal. Here, one usually requires the families $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ and $\{\psi(\cdot - k) : k \in \mathbb{Z}\}$ to constitute so-called *Riesz bases* for their spans. We shall not go into the mathematical details here, we only mention that a Riesz basis is a weaker basis concept than that of an orthonormal basis. From this point of view, a special branch of wavelet theory emerges, namely the theory of *biorthogonal wavelets*, see [1, ch. 8], or the paper [9].

4 Wavelet transforms

Assume that $\{V_j\}$ is a Multiresolution Analysis, and that the scaling function ϕ satisfies a dilation equation with filter coefficients h_n. Let the wavelet ψ be defined by equation (5).

4.1 Single scale and multiscale bases

Consider the equations (6) and (7). It states that any f_J in V_J can either be represented in the *single scale basis*

$$\{\phi_{J,k} : k \in \mathbb{Z}\} \tag{8}$$

of V_J, with respect to which f_J has the representation

$$f_J = \sum_k c_{J,k}\phi_{J,k}, \tag{9}$$

or f_J can be represented with respect to the *multiscale basis*

$$\{\phi_{j_0,k} : k \in \mathbb{Z}\} \cup \{\psi_{j_0,k} : k \in \mathbb{Z}\} \cup \ldots \cup \{\psi_{J-1,k} : k \in \mathbb{Z}\} \tag{10}$$

of $V_{j_0} \oplus \bigoplus_{j=j_0}^{J-1} W_j$, with respect to which f_J has the representation

$$f_J = \sum_k c_{j_0,k}\phi_{j_0,k} + \sum_{j=j_0}^{J-1}\sum_k d_{j,k}\psi_{j,k}. \tag{11}$$

In order to switch back and forth between single scale and multiscale representations of a given function, we need to know how to perform these change-of-basis operations. This is the role of the wavelet transform and its inverse.

4.2 The Wavelet transform

Definition 4.1 Let $J > j_0$ be integers, and let $f_J \in V_J$. Assume that f_J has the representation (9) with respect to the single scale basis (8), and that f_J has the representation (11) with respect to the multiscale basis (10). The mapping

$$\mathcal{W}_{J,j_0} : \{c_{J,k}\} \mapsto \{c_{j_0,k}, d_{j_0,k}, d_{j_0+1,k}, \ldots d_{J-1,k}\} \tag{12}$$

is called the *wavelet transform*, and the mapping

$$\mathcal{Y}_{j_0,J} : \{c_{j_0,k}, d_{j_0,k}, d_{j+1,k}, \ldots d_{J-1,k}\} \mapsto \{c_{J,k}\} \tag{13}$$

is called the *inverse wavelet transform.* The transforms $\mathcal{W}_{j,j-1}$ resp. $\mathcal{Y}_{j-1,j}$ are called the *one-step wavelet transform* resp. the *inverse one-step wavelet transform.*

We state without proof the following theorem.

Theorem 4.2 Assume that $\{V_j\}$ is a Multiresolution Analysis, and that the scaling function ϕ satisfies a dilation equation with filter coefficients $\{h_n\}$. On the one-step wavelet transform and its inverse, the following assertions hold:

(i) Assume that $f_{j+1} \in V_{j+1}$ has coordinates $\{c_{j+1,k}\}$ with respect to the single-scale basis $\{\phi_{j+1,k} : k \in \mathbb{Z}\}$ in V_{j+1}. Then $\mathcal{W}_{j+1,j}\{c_{j+1,k}\} = \{c_{j,k}, d_{j,k}\}$, where

$$c_{j,k} = \sum_n h_{n-2k} c_{j+1,n} \quad \text{and} \quad d_{j,k} = \sum_n (-1)^n h_{2k-n+1} c_{j+1,n}. \tag{14}$$

(ii) Assume that $f_{j+1} \in V_{j+1}$ has coordinates $\{c_{j,k}, d_{j,k}\}$ with respect to the multiscale basis $\{\phi_{j,k}, \psi_{j,k} : k \in \mathbb{Z}\}$ in $V_{j+1} = V_j \oplus W_j$. Then $\mathcal{Y}_{j,j+1}\{c_{j,k}, d_{j,k}\} = \{c_{j+1,k}\}$, where

$$c_{j,k} = \sum_n c_{j+1,n} h_{k-2n} + \sum_n (-1)^n d_{j,n} h_{2n-k+1}. \tag{15}$$

Hence, the one-step wavelet transform and its inverse is executed by performing a convolution-like operation of coordinates with the filter coefficients of the scaling function.

The general wavelet transform and its inverse is now performed by iterated use of the one-step wavelet transforms. As an example, let $J = 3$ and $j_0 = 0$, and let $f_3 \in V_3$ be represented in the single scale basis in V_3:

$$f_3 = \sum_k c_k^3 \phi_{3,k}.$$

The representation of f_3 with respect to the multiscale basis $\{\phi_{0,k}, \psi_{0,k}, \psi_{1,k}, \psi_{2,k} : k \in \mathbb{Z}\}$ is now obtained as indicated in the following diagram:

$$\begin{array}{ccccccc}
 & \mathcal{W}_{3,2} & & \mathcal{W}_{2,1} & & \mathcal{W}_{1,0} & \\
\{c_{3,k}\} & \rightarrow & \{c_{2,k}\} & \rightarrow & \{c_{1,k}\} & \rightarrow & \{c_{0,k}\} \\
 & \searrow & & \searrow & & \searrow & \\
 & & \{d_{2,k}\} & & \{d_{1,k}\} & & \{d_{0,k}\} \\
 & & & & \{d_{2,k}\} & & \{d_{1,k}\} \\
 & & & & & & \{d_{2,k}\}
\end{array} \tag{16}$$

Analogously, if the representation of f_3 in the multiscale basis $\{\phi_{0,k}, \psi_{0,k}, \psi_{1,k}, \psi_{2,k}\}$ is known, the representation of f_3 in the single scale basis in V_3 is obtained by applying the inverse one-step wavelet transform as indicated in the following diagram:

$$\begin{array}{ccccccc}
 & \mathcal{Y}_{0,1} & & \mathcal{Y}_{1,2} & & \mathcal{Y}_{2,3} & \\
\{c_{0,k}\} & \rightarrow & \{c_{1,k}\} & \rightarrow & \{c_{2,k}\} & \rightarrow & \{c_{3,k}\} \\
 & \nearrow & & \nearrow & & \nearrow & \\
\{d_{0,k}\} & & \{d_{1,k}\} & & \{d_{2,k}\} & & \\
\{d_{1,k}\} & & \{d_{2,k}\} & & & & \\
\{d_{2,k}\} & & & & & &
\end{array} \tag{17}$$

4.3 Complexity of the wavelet transform

One can show that as the scale parameter j increases, the functions $2^{j/2}\phi_{j,k}$ behave 'approximately as delta functions', in the sense that

$$\int_{\mathbb{R}} 2^{j/2}\phi_{j,k}(x)f(x)dx = 2^{j/2}\langle \phi_{j,k}, f\rangle \to f(2^{-j}k) \text{ as } j \to \infty.$$

Hence we can take the samples of a function $2^{j/2}f$ on a sufficiently fine dyadic grid of points $\{2^{-J}k : k \in \mathbb{Z}\}$ to be the scaling function coefficients in the single-scale representation in V_J: $c_{J,k} = 2^{j/2}f(2^{-J}k)$. Assume that f is a compactly supported function we want to represent in a multiscale basis. We choose a sufficiently large V_J-space, and if we let $|\text{supp}f|$ denote the length of the support of f, then f can be represented in the single scale basis in V_J using $N = 2^J|\text{supp}f|$ coefficients $c_{J,k}$, which we by the remark above can approximate by sampling the function f on the corresponding dyadic grid. One can show that for each step of the one-step wavelet transform, the $c_{j,k}$ split into $c_{j-1,k}$ and $d_{j-1,k}$ in such a way that there are approximately half as many $c_{j-1,k}$ as $c_{j,k}$, and approximately half as many $d_{j-1,k}$'s as $c_{j,k}$'s. An application of

$$\sum_{j=1}^{\infty} 2^{-j} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 1$$

then show that an execution of a wavelet transform requires $\mathcal{O}(N)$ floating point operations. An application of the Fast Fourier transform requires $\mathcal{O}(N\log_2 N)$ floating point operations.

4.4 Behaviour of wavelet coefficients

The wavelet cofficients $d_{j,k}$ in multiscale basis representations of $L_2(\mathbb{R})$ functions exhibit a remarkable behaviour that can be exploited successfully in a wide range of applications. Suppose that f is an $L_2(\mathbb{R})$ function that we want to represent in a multiscale basis, and that f in a neighborhood of some point, which we for simplicity take to be $x = 0$, is 'well behaved' in the sense that it can be approximated by a polynomial of a low degree s:

$$f(x) \simeq \sum_{p=0}^{s} a_p x^p, \text{ for } x \text{ close to } 0.$$

If we assume that ψ has n vanishing moments with $n \geq s$, and we study the wavelet coefficient $d_{0,0} = \langle f, \psi_{0,0} \rangle$, we get

$$\begin{aligned} d_{0,0} &\simeq \int_{\mathbb{R}} \psi(x) \sum_{p=0}^{s} a_p x^p dx \\ &= \sum_{p=0}^{s} a_p \int_{\mathbb{R}} x^p \psi(x) dx \\ &= 0. \end{aligned}$$

In general, the wavelet coefficients $d_{j,k}$ exhibit the behaviour that if the function f in some neighborhood has slow variations compared to the wavelet $\psi_{j,k}$, then the corresponding wavelet coefficient $d_{j,k}$ is zero or close to zero, and vice versa. This property gives rise to a wide range of applications, some of which we consider in section 6.

5 Some examples of wavelet families

In this section we briefly review some of the most commonly used wavelet families. Nearly all wavelets that are used in applications have in common that they cannot be expressed in closed form by standard functions such as polynomials, root functions, trigonometric or logarithmic functions.

5.1 The Daubechies wavelet family

The Daubechies wavelet family was the first example of a family of compactly supported wavelets. They are designed by constructing a finite sequence of filter coefficients $\{h_n\}$ satisfying some conditions that we shall mention now. Define

$$m(\xi) = \sum_{n} h_n \mathrm{e}^{-in\xi}, \tag{18}$$

which is a so-called trigonometric polynomial if only finitely many h_n's are nonzero. The function in equation (18) is often referred to as the *two-scale symbol*. The scaling function is now defined through its Fourier transform as the infinite product

$$\hat{\phi}(\xi) = \prod_{j=0}^{\infty} m(2^{-j}\xi).$$

It can be verified that ϕ thus defined satisfies a dilation equation with coefficients h_n. Sufficient conditions for $\{\phi(x-k) : k \in \mathbb{Z}\}$ to be an orthonormal system are

$$|m(\xi)|^2 + |m(\xi+\pi)|^2 = 1 \tag{19}$$
$$m(0) = 1 \tag{20}$$
$$m(\xi) \neq 0 \text{ on } [-\frac{\pi}{2}, \frac{\pi}{2}] \tag{21}$$

The Daubechies wavelets are now designed in such a way that $m(\xi)$ approximates $1_{[-\frac{\pi}{2},\frac{\pi}{2}]}(\xi)$. Some of the Daubechies wavelets are shown in figure (2). Note that these wavelets are nonsymmetrical, as mentioned in section 3. The Daubechies wavelet family are optimal with respect to vanishing moments and support. This means that among all wavelet families with a fixed number of vanishing moments, the Daubechies family is the one in which the wavelets have minimal support. Also, among all wavelet families with a given support width, the Daubechies wavelet family is the one with the largest number of vanishing moments. More information on the Daubechies wavelet family, such as the filter coefficients, can be obtained from [1].

5.2 Spline wavelets

A spline of order m is a function $s : \mathbb{R} \mapsto \mathbb{R}$ which is $m-2$ times continuously differentiable, and whose restrictions $s|_{[k,k+1[}$ are polynomials of maximal degree $m-1$. Splines are widely used in numerical analysis, for instance for constructing a function that interpolates a given set of data (x_i, y_i), where i belongs to some index set. There are various ways of constructing splines, but often they possess other desirable properties, such as symmetry and fast decay. One can construct (biorthogonal) wavelets that are splines, and thus combine the benefits of splines and wavelets. For a comprehensive account on spline wavelets, we refer to [7, ch. 4+6], or [8, ch. 3.3].

5.3 Wavelets for an interval

The wavelet bases considered so far are designed for the space $L_2(\mathbb{R})$. In the construction of wavelet bases, the Fourier transform on $L_2(\mathbb{R})$ plays a significant role, and thus it is not a trivial task to design wavelet bases for, say, an interval $[0,1]$. However, one way of doing so is to consider *periodized* versions of scaling functions and wavelets. For $f \in L_2(\mathbb{R})$, one defines the *periodization* of f as

$$f^{\text{per}}(x) = \sum_{n\in\mathbb{Z}} f(x+k).$$

Note that f^{per} is indeed 1-periodic. Suppose that $\{V_j\}_{j\in\mathbb{Z}}$ is a MRA in $L_2(\mathbb{R})$. Now the idea is to periodize the scaling function and wavelet

$$\phi^{\text{per}}(x) = \sum_{n\in\mathbb{Z}} \phi(x+k), \;\; \psi^{\text{per}}(x) = \sum_{n\in\mathbb{Z}} \psi(x+k),$$

and thereby approximate functions in $L_2([0,1])$ by linear combinations of translates and dilates of the periodized scaling function and wavelet. In this way, one can define the *periodized MRA* $\{\tilde{V}_j\}_{j\in\mathbb{Z}}$ as

$$\tilde{V}_j = \mathrm{span}\{\phi^{\mathrm{per}}_{j,k} : k \in \mathbb{Z}\},$$

and it turns out that $\{\tilde{V}_j\}_{j\in\mathbb{N}}$ inherits all the properties described in sections 2-4. For more information on periodized MRA, the reader is referred to [8, ch.2].

There is also other ways of constructing wavelets on the interval, but since these are considerably more complicated than the above case, we shall not mention them here, the reader is referred to the paper [10].

6 Applications

In the cases of wavelet applications in denoising and compression, the appealing property of wavelet techniques lies in the ability of wavelets to separate different kinds of information in a very effective way. If f is for instance a speech signal or an image, f is approximated by picking an f_J from a sufficiently large V_J space. Then one employs the wavelet transforms of section 4, thereby splitting f_J into parts

$$f_J = f_{j_0} + \sum_{j=j_0}^{J-1} g_j, \quad \text{with } f_{j_0} \in V_{j_0},\ g_j \in W_j,\ j = j_0 \ldots J-1. \tag{22}$$

where f_{j_0} is a coarse approximation to f_J, and the g_j's are a sequence of detail terms, containing increasing amounts of detail of the function f_J. The role played by the parts $f_{j_0}, g_{j_0}, \ldots, g_{J-1}$ depend on the given application.

6.1 Denoising

Suppose that f is a segment of a noisy speech signal, and that we want to denoise it. We sample the signal on a sufficiently fine grid, thereby getting a representation/approximation f_J of the signal in a space V_J, see subsection 4.3. The space V_J should be so large that it includes both the signal and the noise. Then we perform a wavelet transform, thereby representing f_J as in equation (22). After the right number of wavelet transforms, the coarse approximation f_{j_0} is the denoised signal, whereas the noise is contained in the detail terms g_j. Now we have separated the signal from the noise, whereby the task of denoising the signal is resolved.

6.2 Compression

Suppose that f represents an image, that we want to store on a computer, and we wish to use a small amount of memory. We sample the image on a sufficiently fine grid, thereby getting a representation/approximation f_J in a V_J space. Now we perform a wavelet transform, as in the previous subsection. The detail terms are represented in wavelet bases as

$$g_j = \sum_k d_{j,k} \psi_{j,k}.$$

Depending on the number of vanishing moments of the wavelet that has been chosen, the wavelet coefficients $d_{j,k}$ corresponding to areas where the image does not contain sharp edges, will be zero, or very small. Those that are below some predefined tolerance can be set equal to zero without losing much information; hence only the wavelet coefficients significantly larger than zero need to be stored. This process is called *thresholding* in the wavelet literature. Loosely speaking, wavelets concentrate their effort in analyzing the most important parts of the image. For more information on image compression, the reader could consult [6], and the references therein.

6.3 Differential and integral equations

Wavelets exhibit promising potential in the area of adaptive solution methods of operator equations. An adaptive method is a method whose numerical procedure depends on the solution being approximated. The idea is to use more computational power to resolve the solution in areas where it is relatively irregular, and less computational power to resolve the solution in areas where it is relatively smooth. These methods are iterative, and in each iteration one seeks information on the current error, and what one can do to improve the approximation in the next iteration.

Suppose that we want to solve an operator equation $\mathcal{L}u = f$, where $\mathcal{L}$ is a differential or an integral operator, f is a given right hand side, and u is the solution. We also assume that we have in hand a wavelet basis $\{\psi_\lambda\}$. Here is a coarse diagram of an adaptive wavelet method: For $i = 0, 1, \ldots$

1 If $i = 0$, determine an initial guess

$$u_0 \in X_0 = \text{span}\{\psi_\lambda : \ \lambda \in I_0\}$$

otherwise solve $\mathcal{L}_i u_i = f_i$ for $u_i \in X_i$.

2 Find the residual in the wavelet basis

$$r_i = \mathcal{L}(u - u_i) = f - \mathcal{L}u_i = \sum_\lambda d^i_\lambda \psi_\lambda.$$

If r_i is small, stop.

3 Construct new approximation space

$$X_{i+1} = X_i \cup \text{span}\{\psi_\lambda : |d^i_\lambda| \quad \text{'large'}\}$$

4 $i := i + 1$, go to 1.

The large wavelet coefficients in the residual tell us where the error of the current approximation is large, and the corresponding wavelet basis function should be included in the approximation space in the next iteration.

Wavelet exhibit a number of other properties that are useful in this area, the description of which lies beyond the scope of this note. We refer to [3], [4].

7 Where to find wavelet material

Wavelet digest is an internet journal, to which currently 15000 people subscribe. It gives information on wavelet related textbooks, articles, conferences, etc. The URL is `http://www.wavelet.org`.

On the URL `http://www.mathsoft.com/wavelets` there is a large number of papers on wavelets and applications in various areas, such as ordinary and partial differential equations, statistics, image and sound processing, etc.

Here is a list of journals in which wavelet articles are often published:

- Acta Numerica
- Applied and Computational Harmonic Analysis
- Constructive Approximation
- Journal of Fourier Analysis and Appliciations
- Journal of Mathematical Analysis and Applications
- SIAM Journal of Mathematics and Applications.
- SIAM Journal of Mathematical Analysis

References

[1] I. Daubechies: *Ten Lectures on Wavelets.* SIAM, 1992.

[2] A.K.Chan, J.C.Goswami: *Fundamentals of Wavelets.* Wiley Interscience 1999.

[3] W. Dahmen: *Wavelet and Multiscale methods for Operator Equations.* Acta Numerica 1997.

[4] A. Cohen, W. Dahmen, R. DeVore: *Adaptive Wavelet Methods for Elliptic Operator Equations, Convergence Rates.* Math. Comp. **70** (2001), 27–75.

[5] R. Devore and B. Lucier: *Wavelets.* Acta Numerica, Vol. 1, Cambridge University Press, p. 1-56. See also the wavelet resources website.

[6] G.Strang, T.Nguyen: *Wavelets and Filter Banks.* Wellesley-Cambridge press, Wellesley, MA, 1996.

[7] C. Chui: *An Introduction to Wavelets.* Academic press, 1992.

[8] P. Wojtaszczyk: *A Mathematical Introduction to Wavelets.* Cambridge University Press, 1997.

[9] A. Cohen, I. Daubechies, J-C. Feauveau: *Biorthogonal Bases of Compactly Supported Wavelets.* Comm. Pure and Appl. Math, **45**, 1992, p. 485-560.

[10] A. Cohen, I. Daubechies, P. Vial: *Wavelets on the Interval and Fast Wavelet Transforms.* Appl. Comp. Harm. Anal., **1**, 1993, p. 54-81.

Torben Klint
Deptartment of Mathematics
Building 303, room 117
Technical University of Denmark
2800 Kongens Lyngby
Phone: ++45 45253052
E-mail:`T.Klint@mat.dtu.dk`
Homepage: `http://www.mat.dtu.dk/persons/Jensen_Torben_Klint`

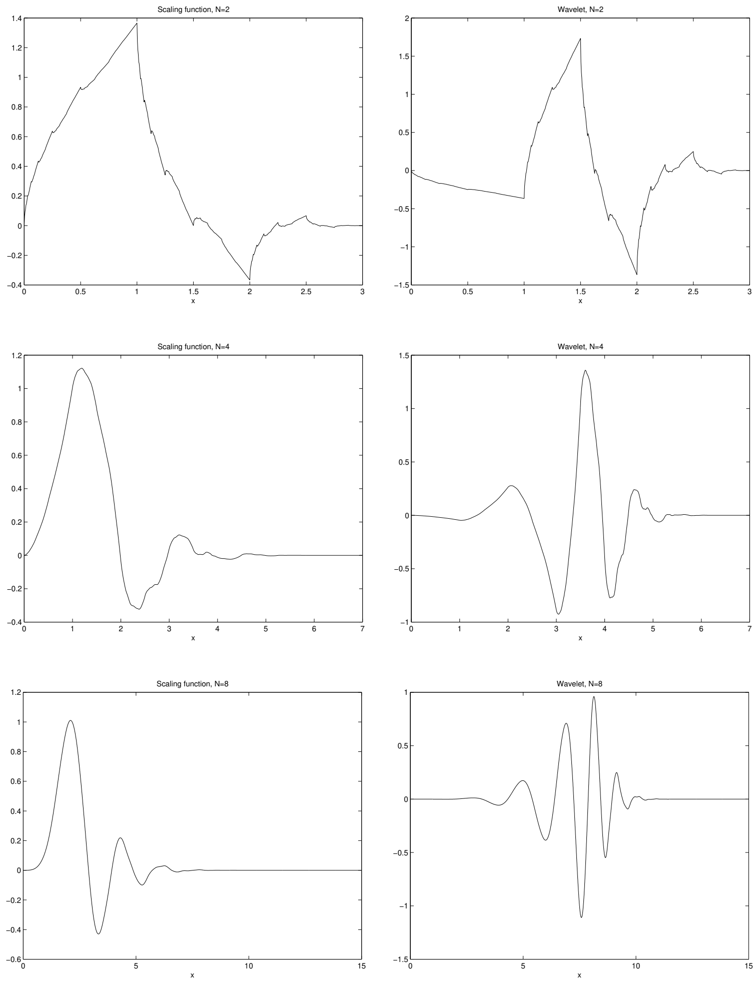


Figure 2: Some examples of Daubechies compactly supported scaling functions and wavelets. The number N is half of the number of coefficients in the dilation equation (2).