Learning to swim in a sea of wavelets

Adhemar Bultheel

December 14, 1994

Abstract

We give some introductory notes about wavelets, motivating and deriving the basic relations that are used in this context. These notes should be considered as in introduction to the literature. They are far from complete but we hope it can motivate some readers to get involved with a quite interesting piece of mathematics which is the result of a lucky mariage between the results of the signal processing community and results in multiresolution analysis. We try to give answers to the questions: What are wavelets? What is their relation to Fourier analysis? Where do the scaling function and the wavelet function come from? Why can they be useful? What is a wavelet transform? Where and how are they applied?

Contents

1 History

Some ideas related to wavelets already existed at the beginning of the century, but the real development came only in the mid-eighties.

Besides a paper by Frazier and Jawerth (1985) [12], wavelets were in the initial stage developed in France, the so called "French school" lead by J. Morlet, A. Grossmann and Y. Meyer.

Wavelets, or "Ondelettes" as they are called in French were used at the beginning of the eighties by J. Morlet, a geophysicist, as a tool for signal analysis for seismic data. The numerical success of this application prompted A. Grossmann (a theoretical physicist) and J. Morlet to make a theoretical study of the wavelet transforms (1984) [13, 15]. In 1985, Y. Meyer, an harmonic analyst, pointed out the strong connection with the existing analysis techniques of singular integral operators.

Ingrid Daubechies became involved in 1986 and this started an interaction between signal analysis and the mathematical aspects of dilations and translations [7, 10].

Also Stephane Mallat became involved when he noticed the connection with multiresolution analysis [22].

A major breakthrough was provided in 1988 when Daubechies managed to construct a family of orthonormal wavelets with compact support [8], a result inspired by the work of Meyer and Mallat in the field of multiresolution analysis. Since then mathematicians, physicists and applied scientists became more and more excited about the ideas. See for example [6, 17, 5].

2 Motivation

We assume that the reader is familiar with Fourier analysis. For reference, we added some formulas in appendix A. We use $\mathcal{F}$ to denote the Fourier transform and $L_2 = L_2(\mathbb{R})$ for the space of square integrable functions on the real axis $\mathbb{R}$.

We shall sometimes refer to a function as being a signal. The norm $\|f\|$ will always refer to the 2-norm. It is called the energy of the signal and f being square integrable thus means that it has finite energy.

Let δ denote the Dirac delta function, then one knows that

$$f(x) = \int_{-\infty}^{\infty} f(u)\delta(x-u)d\xi$$

One could say that $f(x)$ is decomposed with respect to an orthonormal basis $\{\delta(x-u)\}_{u\in\mathbb{R}}$ of L_2. The "coordinates" with respect to this basis are just the function values $f(u)$. The basis functions are extremely local in the x-domain, but in the frequency domain, its Fourier transform is supported on the whole real axis: $\mathcal{F}\delta(x-u) = e^{-i\xi u}/\sqrt{2\pi}$, $\xi \in \mathbb{R}$. Thus, even though it is as narrow as possible in the x-domain, each basis function encloses all possible frequencies. This is an extreme situation.

At the other extreme, we may consider another orthonormal set of basis functions $\{e^{ixu}\}_{u\in\mathbb{R}}$ for L_2. These give an "expansion" of $f(x)$ as

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(u)e^{iux}du$$

where $\hat{f}$ denotes the Fourier transform of f. These $\hat{f}(u)$ are the coordinates for this basis. Now the basis functions are each associated with just one frequency since $\mathcal{F}e^{iux} = \delta(\xi - u)/\sqrt{2\pi}$, however in the x-domain, they are supported on the whole of $\mathbb{R}$.

The most interesting wavelets are somewhere in between these two extremes. They are local in both the x- and the frequency domain. They become really interesting if on top of that they are still an orthonormal set. We say that a function is local when most of its energy (consider e.g. its norm as a measure for the energy) is located in a finite interval. Such functions are zero (compact support) or decay quickly outside this interval. One can imagine such a function as an oscillation, with significant lobes inside a finite interval and practically zero outside. This should explain the term wavelet. The property of being local in both x-domain and frequency domain will make wavelets very suitable to analyse a signal by decomposing it into components which are local in both domains. This means that we can look at the signal in a finite time slice (a window moving over the signal) and at the same time select only the frequencies within a certain frequency band (as if there is also a window sliding over all possible frequencies). The sliding window in the x-domain corresponds to translations of the wavelet; the dilations of the wavelet will have a windowing effect in the frequency domain (see below). In the sequel of this paper we shall usually have wavelets with a compact support in the x-domain and local, i.e. with almost compact support in the frequency domain. We can not have compact support in both domains by the Fourier-Heisenberg uncertainty principle.

The previous analysis for L_2 can be repeated for 2π-periodic functions defined on $\mathbb{R}$. The space of square integrable 2π-periodic functions is denoted by $L^2_{2\pi}$ and it has the orthonormal basis $\{e^{inx}\}_{n\in\mathbb{Z}}$. Its Fourier expansion with respect to this basis is (a summation with no boundaries will always be supposed to range over $\mathbb{Z}$)

$$f(x) = \sum_n \hat{f}_n e^{inx}$$

with Fourier coefficients (the Fourier transform is a function of the discrete variable n)

$$\hat{f}_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx =: \left\langle f(x), e^{inx}\right\rangle.$$

Again the basis functions $\{w_n(x) = e^{inx} : n \in \mathbb{Z}\}$ form an orthonormal set: $\langle w_m, w_n\rangle = \delta_{nm}$. There is something remarkable about these basis functions. All the functions w_n are obtained as (integer) *dilations* of the same function $w(x) = e^{ix}$. This means that $w_n(x) = w(nx)$. Thus Fourier analysis tells us that every $f \in L^2_{2\pi}$ can be written as a linear combination of integer dilations of one basic function $w(x) = e^{ix}$. For large (small) n the function w_n corresponds to a high (low) frequency.

This is now another feature we want to have in wavelets: they form a basis which can be generated by dilations of one “mother” function. For $L_2 = L_2(\mathbb{R})$, the basis of sinusoidal waves $\{e^{inx}\}$ is not appropriate since it does not even belong to L_2. Instead of waves with an infinite support, we want “wavelets” which are local, i.e. with an (almost) compact support. So let $\psi(x)$ be the compactly supported mother function for L_2. However, because of the compact support, it will never be possible to expand every $f \in L_2$ by dilations of ψ, since all these will have compact support too. Therefore dilations are not enough and we need translations as well.

On the other hand, translations alone are not enough. Since ψ has compact support, one could hope to generate L^2 by shifting it (e.g. over all integer values). Thus we consider the functions $\psi_{0k}(x) = \psi(x-k)$, for $k \in \mathbb{Z}$. This is not enough though because also the Fourier transform $\hat{\psi}$ is supposed to be local and thus practically all its frequencies are from a fixed frequency band (i.e., it is practically band limited). Therefore all these ψ_{0k} contain frequencies from practically the same frequency band and there are of course functions in L_2 which are not band limited. Thus we shall have to consider both dilations and translations of the mother function. Thus, while the translates of ψ cuts the x-axis into small (i.e. local) pieces, the dilations of ψ correspond to a division of the frequency range into octaves. For computational efficiency, the scaling factors are often chosen to be powers of 2. So we shall consider the functions $\psi(2^jx-k)$. This is now a 2-parameter set of functions. In L^2, it is easily seen that $\|\psi(2^jx-k)\|^2 = 2^{-j}\|\psi\|^2$. Hence, the functions

$$\psi_{nk}(x) = 2^{n/2}\psi(2^nx-k)$$

will have unit length if $\|\psi\| = 1$. An important problem will be to choose ψ such that they also form an orthonormal set, i.e., such that in L^2 with inner product $\langle f, g\rangle := \int_{\mathbb{R}} f(x)\overline{g}(x)dx$,

$$\langle\psi_{jk}, \psi_{lm}\rangle = \delta_{jl}\delta_{km}.$$

In that event $f \in L^2$ will have the expansion

$$f(x) = \sum_{j,k} c_{jk}\psi_{jk}(x) \quad \text{with} \quad c_{jk} = \langle f, \psi_{jk} \rangle \,.$$

We mention here that also B-splines or finite elements form a set of basis functions with a compact support. Thus they are also closely related to wavelet analysis and there exists a florishing literature where splines and wavelets are integrated. See for example [4]. These basis functions are not orthogonal to their translates though. However splines are known for their ability to produce smooth approximations. Wavelets on the other hand will not be so smooth. This can be a disadvantage, but if the objective is to detect sharp edges (e.g. in image processing) this turns into an advantage. About this aspect see later in this paper. In the present paper we shall not discuss spline wavelets, cutting off an important branch of the existing literature on classical wavelets.

3 Discrete versus continuous wavelet transforms

The previous analysis brought us to the basis $\{\psi_{nk}\}$ in two discrete parameters n and k. The set of corresponding coefficients $c_{nk} = \langle f, \psi_{nk} \rangle$ is called the wavelet transform of f. The inner product is $\langle f, \psi_{nk} \rangle = \int_{-\infty}^{\infty} f(x)\overline{\psi_{nk}(x)}dx$. Often the function (signal) will be sampled, i.e., it takes the value $f(j)$ for $x = j$, $j \in \mathbb{Z}$ and is zero everywhere else. Then of course $f \in \ell_2(\mathbb{Z})$ and the integral has to be replaced by an infinite sum. For practical computations, one shall not work with infinitely long vectors $(f(j))_{j=-\infty}^{\infty}$ but only with vectors of finite length: $(f(j))_{j=0}^{N}$ say. For convenience (like in discrete Fourier transform (DFT)) N is chosen to be a power of 2: $N = 2^K$ and the (discrete) signal is supposed to be periodically extended beyond this range to all $j \in \mathbb{Z}$. This is the practical application of the discrete wavelet transform (DWT).

In fact, the term DWT is somewhat ambiguous. One could refer by this to the transform where the signal is considered as a discrete function (a sequence) as we just did, or one could use this to refer to the case where the signal is continuous but where the two parameters generating the wavelet basis (n and k in our notation) are discrete, as we did in the previous section. The continuous variant of the latter signification is the one where the discrete parameters n and k are replaced by continuous ones (we shall call them a and b). The analysis of the continuous wavelet transforms can be made in parallel with the discrete version. Somewhat naively we could say that summations over n and/or k should be replaced by integrals over a and/or b. Which formalism is chosen is usually depending on the applications one has in mind or the personal taste and background of the author. In this paper we have chosen for the disrete formalism. One argument being that our interest goes to applications which should be numerically implemented on a computer in which case the integrals will be replaced by sums anyway. This need not be restricted though to the dyadic points we shall consider here. In certain applications we need redundancy and some oversampling is done (when you want for example an error correcting mechanism for the music signal of a CD recording).

However, to show that it is perfectly possible to give continuous analogs by replacing the discrete parameters n, k by continuous parameters a, b, and for the sake of comparison, we briefly mention some formulas here.

In the continuous case we consider a family of functions, depending on two continuous parameters a and b, which are all connected with one single function that is scaled and shifted.

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}}\psi(\frac{x-b}{a}), \quad a \neq 0, \quad b \in \mathbb{R} \tag{3.1}$$

with $\psi \in L_2 = L_2(\mathbb{R})$.

As an example, take

$$\psi(x) = (1-x^2)e^{-\frac{1}{2}x^2} \quad \text{(Mexican hat)} \tag{3.2}$$

Note that $\psi(x)$ is the second derivative of $\exp(\frac{1}{2}x^2)$. Hence we know from Fourier analysis that

$$\hat{\psi}(\xi) = \xi^2 e^{-\frac{1}{2}\xi^2}$$

The function and two of its dilations are plotted in figure 1. Observe that varying b just shifts

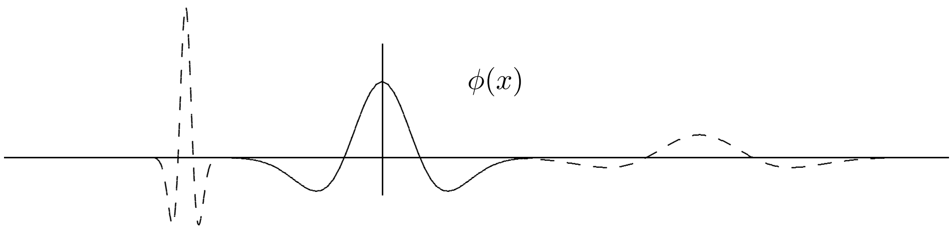


Figure 1: The Mexican hat function and two of its translated dilations

$\psi_{a,b}(x)$ along the x-axis while a small shrinks the picture to a narrow one (high frequency) while a large spreads out the picture (low frequency).

Note that ψ nor $\hat{\psi}$ really have a compact support but they decay very fast and the supports are nearly compact.

In analogy with the discrete version, one defines the *continuous wavelet transform* of f as

$$F(a,b) = \int_{-\infty}^{\infty} f(x)\overline{\psi_{a,b}(x)}dx = \langle f, \psi_{a,b}\rangle\,, \quad f \in L_2$$

When setting $a = 2^{-j}$ and $b = 2^{-j}k$, one is back in the discrete case.

For the inverse transform to exist, one should require that ψ satisfies

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty \tag{3.3}$$

where $\hat{\psi}$ is the Fourier transform of ψ. The condition (3.3) is necessary for the inverse transformation to exist (see (3.4) below).
Note that $C_\psi < \infty$ implies $\hat{\psi}(0) = 0 = \int_{-\infty}^{\infty} \psi(x)dx$. Thus ψ should change sign (to have a mean value 0). Note that for the Mexican hat function, $C_\psi = 2\pi$.

The inverse transform is given by

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} F(a,b)\psi_{a,b}(x)db \right] \frac{da}{a^2}. \tag{3.4}$$

In what follows, we shall be interested in the case where $b \in \mathbb{R}$ and $a > 0$ (only the values $2^j > 0$ in the discrete version). The restrictive condition (3.3) then has to be replaced by

$$\frac{1}{2}C_\psi = \int_0^{\infty} \frac{|\hat{\psi}(\xi)|^2}{\xi} d\xi < \infty$$

with reconstruction formula

$$f(x) = \frac{2}{C_\psi} \int_0^{\infty} \left[\int_{-\infty}^{\infty} F(a,b)\psi_{a,b}(x)db \right] \frac{da}{a^2}.$$

More about continuous wavelet transformations and their relation with short time Fourier transforms can be found in for example [14, 27].

4 Multiresolution analysis

The Fourier like framework we need to study wavelets is multiresolution analysis. We need *multi*resolution because the resolution, i.e., the details of the function that we can see will be governed by the frequencies, i.e. by our dilations though the parameter n. On the other hand, for each resolution we have a space of basis functions obtained by translation of a basic function obtained with the parameter k. Thus we have several spaces at a different resolution: a multiresolution. For an excellent treatment of multiresolution analysis, see [16]. The subsequent notes are based on this paper.

This analysis consists in breaking up L_2 in a sequence of nested closed subspaces V_j

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$$

so that $\bigcup_{n\in\mathbb{Z}} V_j$ is dense in L_2 and $\bigcap_{n\in\mathbb{Z}} V_n = \{0\}$. Moreover one should have

$$\begin{array}{lcl} f(x) \in V_n & \Leftrightarrow & f(2x) \in V_{n+1}, \quad n \in \mathbb{Z} \\ f(x) \in V_0 & \Leftrightarrow & f(x-k) \in V_0, \quad k \in \mathbb{Z} \end{array}$$

and there should exist a function ϕ, such that $\{\phi(x-n)\}_{n\in\mathbb{Z}}$ forms an orthonormal basis of V_0.

Based on this definition, one can prove the following properties. Suppose P_n is the projection of L_2 onto V_n then

$$\begin{array}{l} P_n P_m = P_m P_n = P_n, \quad m \geq n \\ \lim\limits_{n\to-\infty} P_n f = 0, \quad \lim\limits_{n\to\infty} P_n f = f, \quad f \in L_2 \\ P_{n+1} = D_{\frac{1}{2}} P_n D_2 \end{array}$$

with D_a a dilation operator: $(D_a f)(x) = |a|^{-1/2} f(x/a)$.

Example 4.1 (Box functions)

The V_n are the spaces of L_2 functions generated by the basis consisting of the functions which are piecewise constant on the intervals $[2^{-n}k, 2^{-n}(k+1)[$. Thus if χ_{nk} are the characteristic functions for these intervals, then

$$V_n = \operatorname{span}\{\chi_{nk} : k \in \mathbb{Z}\}.$$

It is an easy exercise to prove that these form indeed a multiresolution, where we may choose $\phi(x)$ to be the characteristic function of the unit interval $[0, 1[$. One may check that the projection P_n is defined by

$$(P_n f)(x) = 2^n \int_{2^{-n}k}^{2^{-n}(k+1)} f(y)dy \quad , \quad x \in [2^{-n}k, 2^{-n}(k+1)[$$

and the functions $\phi_{0k}(x) = \phi(x-k)$ form an orthogonal basis for the space V_0.

5 The father function or scaling function

The function ϕ is called the *father function* or *scaling function*. Denote its shifted versions as $\phi_{0n} = \phi(x-n)$, then any $f \in V_0$ can be written as

$$f = \sum_n a_n \phi_{0n}, \quad (a_n) \in \ell^2.$$

Now, since $\phi \in V_0$, also $\phi(\frac{x}{2}) \in V_{-1} \subset V_0$. Thus we may also expand this function in V_0:

$$\phi(\frac{x}{2}) = \sum_n c_n \phi(x-n), \quad x \in \mathbb{R}$$

In other words, the father function satisfies the *dilation equation*

$$\boxed{\phi(x) = \sum_n c_n \phi(2x - n), \quad x \in \mathbb{R}, \quad n \in \mathbb{Z}.}$$

This dilation equation is often called a *two-scale relation* (TSR). To avoid trivialities we look for a solution with $\int_{-\infty}^{\infty} \phi(x)dx \neq 0$. Suppose we normalize ϕ so that $\sqrt{2\pi}\hat{\phi}(0) = \int_{-\infty}^{\infty} \phi(x)dx = 1$, then

$$2\int_{-\infty}^{\infty} \phi(x)dx = \sum_n c_n \int_{-\infty}^{\infty} \phi(2x-n)d(2x-n),$$

thus

$$\boxed{\sum_n c_n = 2}$$

Suppose we can solve the dilation equation for some choice of the coefficients c_n, then we may consider the functions

$$\boxed{\phi_{nk}(x) = 2^{n/2}\phi(2^n x - k), \quad n, k \in \mathbb{Z}} \tag{5.1}$$

Note that $V_n = \text{span}\{\phi_{nk} : k \in \mathbb{Z}\}$.

6 Solution of the dilation equation

Let us first look at some example solutions of the equation

$$\phi(x) = \sum_k c_k \phi(2x - k), \quad \sum_k c_k = 2.$$

Example 6.1 Set $c_0 = 2$ and all other $c_k = 0$. Then a solution is $\phi = \delta$, the Dirac delta function, since indeed $\delta(x) = 2\delta(2x)$. This example shows that a solution is not always a smooth function. Also $\phi = 0$ is a solution but $\hat{\phi}(0) = 0$! The Dirac function resulting from this example is pathological, in the sense that it does not have all the usual properties a scaling function for a wavelet will have. So it is not considered to correspond to a wavelet.

Example 6.2 Set $c_0 = c_1 = 1$. The solution is the *box function*

$$\phi(x) = \chi_{[0,1[}(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{otherwise} \end{cases}$$

The proof can be checked easily on a picture. See figure 2. This example results in the Haar basis.

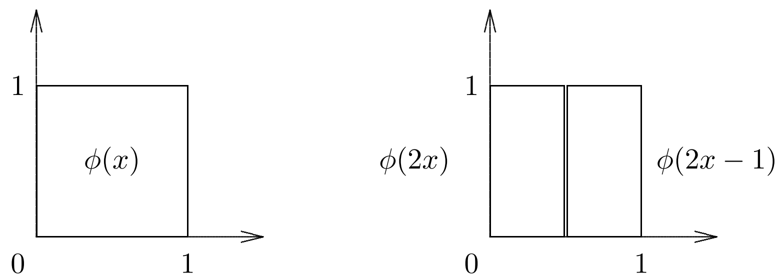


Figure 2: The box function and the dilation equation.

Example 6.3 For $c_1 = 1$, $c_0 = c_2 = \frac{1}{2}$, the solution is the *hat function*

$$\phi(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2 - x, & 1 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

One can check graphically that the hat function satisfies the dilation equation. See the figure 3 (for the wavelet function ψ see below).

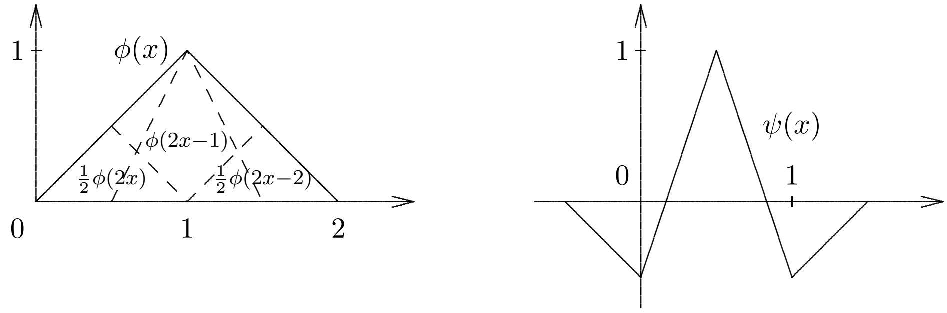


Figure 3: The hat function and the corresponding wavelet

Of course such a graphical construction can only be done for simple examples. We need a more systematic approach. There are several general construction methods. We give some examples.

Construction 1: (By iteration)
One way to find $\phi(x)$ is by iterating $\phi_j(x) = \sum_k c_k \phi_{j-1}(2x - k)$.

Example 6.4 Take for example with $\phi_0 =$ the box function $\chi_{[0,1[}$.
For $c_0 = 2$, the box function gets taller and thinner, so it goes to the Dirac function.
For $c_0 = c_1 = 1$, the box remains invariant $\phi_j = \phi_0$, $j \ge 0$.
For $c_1 = 1$, $c_0 = c_2 = \frac{1}{2}$, the hat function appears as $j \to \infty$.

Example 6.5 Using a computer program with graphical possibilities, one can try the same with $c_0 = c_3 = \frac{1}{4}$, $c_1 = c_2 = \frac{3}{4}$. The solution is a quadratic spline.

$$\phi(x) = \begin{cases} x^2, & 0 \le x \le 1 \\ 2 - 2x^2 + 6x - 3, & 1 \le x \le 2 \\ (3 - x)^2, & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases} \quad \text{(quadratic spline)}$$

Example 6.6 Another example corresponds to the choice $c_0 = c_4 = \frac{1}{8}$, $c_1 = c_3 = \frac{1}{2}$, $c_2 = \frac{3}{4}$. The solution is the cubic B-spline.

Example 6.7 An interesting example is obtained by choosing $c_0 = \frac{1}{4}(1 + \sqrt{3})$, $c_1 = \frac{1}{4}(3 + \sqrt{3})$, $c_2 = \frac{1}{4}(3 - \sqrt{3})$, $c_3 = \frac{1}{4}(1 - \sqrt{3})$. The solution is somewhat surprising. The corresponding wavelet is called D_4 (D for Daubechies and 4 because only 4 coefficients are nonzero). The result is plotted in the first part of figure 4. For the corresponding wavelet function ψ see below.

Construction 2: (By Fourier analysis).
Defining the Fourier transform

$$\hat{\phi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(x) e^{-ix\xi} dx,$$

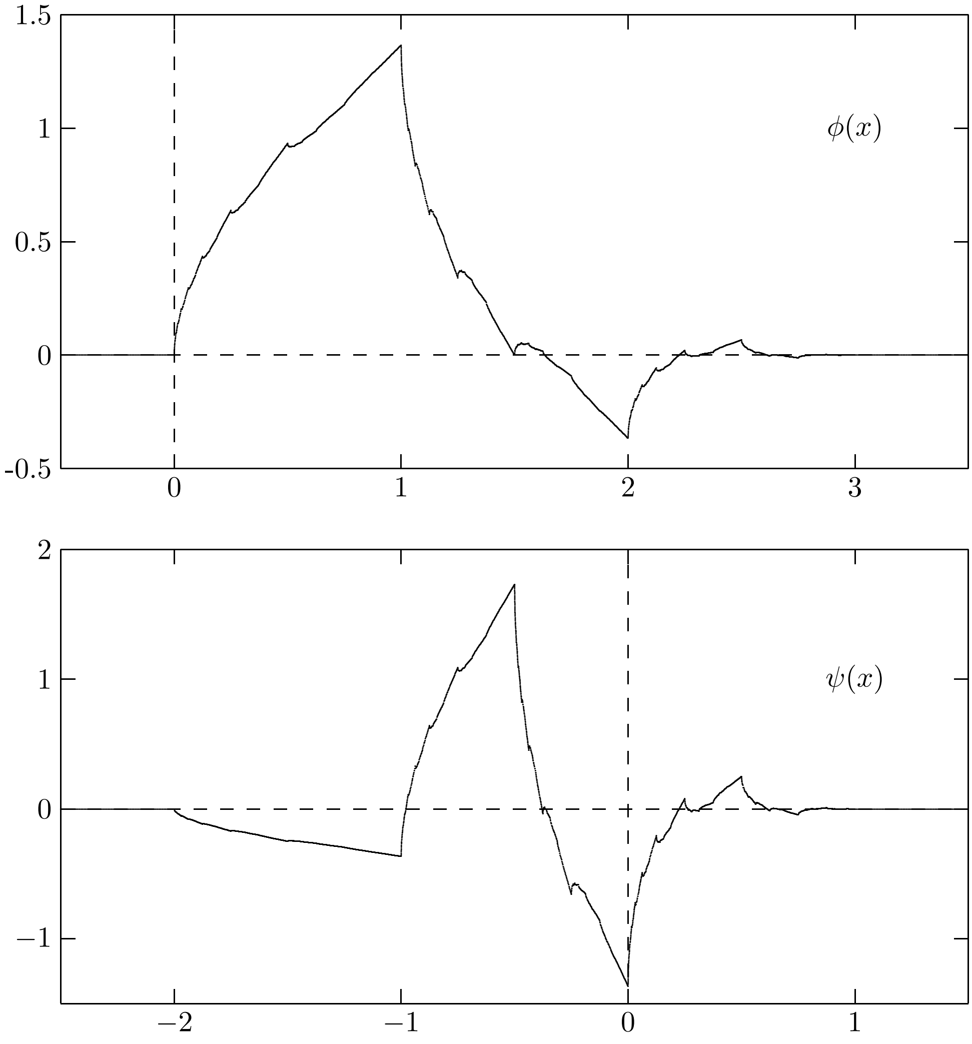


Figure 4: The Daubechies D_4 scaling function and wavelet.

the dilation equation gives

$$\begin{aligned}\hat{\phi}(\xi) &= \sum_n \frac{c_n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(2x-n)e^{-ix\xi}dx = \frac{H(\frac{1}{2}\xi)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(y)e^{-iy\xi/2}dy \\ &= H(\tfrac{1}{2}\xi)\hat{\phi}(\tfrac{1}{2}\xi)\end{aligned}$$

where $H(\xi) = \frac{1}{2}\sum_n c_n e^{-in\xi}$. Note that $H(0) = 1$. Iterating the above result and using $\hat{\phi}(0) = 1/\sqrt{2\pi}\int_{-\infty}^{\infty} \phi(x)dx = 1/\sqrt{2\pi}$, we find

$$\hat{\phi}(\xi) = \frac{1}{\sqrt{2\pi}} \prod_{j=1}^{\infty} H(2^{-j}\xi).$$

It can be shown rigorously that this infinite product makes indeed sense but we shall not do this here.

Example 6.8 $c_0 = 2$, then $H(\xi) = 1$, $\hat{\phi}(\xi) = 1/\sqrt{2\pi}$ and this is indeed the Fourier transform of the Dirac function.

Example 6.9 $c_0 = 1 = c_1$ (box function). The product of the H-functions ($H(\xi) = (1 + e^{-\xi})/2$) is a geometric series.

$$H(\tfrac{1}{2}\xi)H(\tfrac{1}{4}\xi) = \frac{1}{4}(1 + e^{-i\xi/2})(1 + e^{-i\xi/4}) = \frac{1 - e^{-i\xi}}{4(1 - e^{-i\xi/4})}.$$

The product of N such functions is

$$\prod_{k=1}^{N} H(2^{-k}\xi) = \frac{1 - e^{-i\xi}}{2^N(1 - e^{-i\xi/2^N})}$$

which for $N \to \infty$ approaches

$$\sqrt{2\pi}\hat{\phi}(\xi) = \frac{1 - e^{-i\xi}}{i\xi} = \int_0^1 e^{i\xi x}dx = \sqrt{2\pi}\mathcal{F}\chi_{[0,1[}$$

and this identifies $\hat{\phi}$ as the Fourier transform of the box function.

The Fourier analysis approach now gives easily the following examples which you may check.

Example 6.10 The hat function comes from squaring the previous $H(\xi)$, hence squaring $\prod_1^{\infty} H(2^{-j}\xi)$.

Example 6.11 The cubic spline comes from squaring again.

Construction 3: (Recursion)
Suppose $\phi(x)$ is known at integer values $x = k$. Then the dilation equation defines $\phi(x)$ at half integers $x = k/2$. Repeating this process yields $\phi(x)$ at all dyadic points $x = k/2^j$. This is a fast algorithm and it is often used in practice.

Example 6.12 For D_4, we can find starting values at $\phi(1)$ and $\phi(2)$ as follows. We shall show in the next Theorem that $\operatorname{supp}\phi(x) \subset [0, 3]$. It can also be shown that at the boundaries, $\phi(0) = \phi(3) = 0$, so that of all values $\phi(k)$, $k \in \mathbb{Z}$, only $\phi(1)$ and $\phi(2)$ are nonzero. Hence, the dilation equation gives

$$\begin{aligned}\phi(1) &= c_1\phi(1) + c_0\phi(2) \\ \phi(2) &= c_3\phi(1) + c_2\phi(2)\end{aligned} \equiv \begin{bmatrix} \phi(1) \\ \phi(2) \end{bmatrix} = C \begin{bmatrix} \phi(1) \\ \phi(2) \end{bmatrix}$$

Thus $[\phi(1) \quad \phi(2)]^T$ is an eigenvector for the matrix $C = \begin{bmatrix} c_1 & c_0 \\ c_3 & c_2 \end{bmatrix}$. Its eigenvalues are $\lambda = 1$ and $\lambda = \frac{1}{2}$. For $\lambda = 1$, the eigenvector is $\phi(1) = cc_0$, $\phi(2) = cc_3$. The constant c is chosen to normalize the vector. As we shall see later in Lemma 9.5, the values of $\phi(k)$ should sum up to 1. Hence $c = (c_0 + c_3)^{-1}$. This eventually gives the required function values to start with. The next values at 1/2 and 3/2 are given by

$$\begin{aligned} \phi(\frac{1}{2}) &= c_0\phi(1) \\ \phi(\frac{3}{2}) &= c_2\phi(1) + c_1\phi(2) \end{aligned}$$

etcetera.

We saw from the examples that they all had compact support. One can show in general

Theorem 6.1 *If $\phi(x) = \sum_n c_n\phi(2x-n)$ with $c_n = 0$ for $n < N^-$ and $n > N^+$, then* $\mathrm{supp}(\phi) \subset [N^-, N^+]$. *(The support is the smallest closed interval, outside of which the function is zero.)*

Proof. We use construction 1 : Let $\phi_0 = \chi_{[-\frac{1}{2},\frac{1}{2}[}$ and iterate

$$\phi_j(x) = \sum_n c_n\phi_{j-1}(2x-n)$$

Denote $\mathrm{supp}(\phi_j) = [N_j^-, N_j^+]$, then

$$N_j^- = \frac{1}{2}(N_{j-1}^- + N^-) \quad , \quad N_j^+ = \frac{1}{2}(N_{j-1}^+ + N^+)$$

with

$$N_0^- = -\frac{1}{2} \quad , \quad N_0^+ = \frac{1}{2}.$$

An easy induction shows that

$$N_j^- = 2^{-j}N_0^- + \left(\frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^j}\right)N^-,$$

which converges for $j \to \infty$ to N^-. A similar argument shows that $\lim_{j\to\infty} N_0^+ = N^+$. This proves the theorem. $\square$

7 Interpretation of multiresolution

We have seen that a basis for V_0 is given by translates of the father function ϕ. The reciprocal of the translation distance is called the *resolution* of this basis. One could say that the resolution gives the number of basis functions per unit length. Let us set by definition the resolution of V_0 to be 1. The projection P_0f gives an approximation of f at resolution 1. The projection P_jf of f onto V_j gives the approximation of f at resolution 2^j.

The higher j, the higher the resolution, i.e., the more basis functions per unit length. This means that we look at the signal in more detail. Compare the continuous expression (3.1) with the discrete analog (5.1). We could say that by choosing $a = 2^{-j}$, we choose a certain magnification for the microscope we use to look at the signal. The smaller a, the larger the magnification is, thus the more detail we want to catch. Thus if we slide our signal under our microscope it should be with small steps for a large magnification and with coarser steps for a low magnification. The parameter

b in (3.1) which governs the position should therefore be chosen proportional to a: small $a \Rightarrow$ large magnification $\Rightarrow$ small steps $\Rightarrow$ small b. Thus choosing $b = ak$ is a natural thing to do. If we thus discretize a and b as $a = 2^{-j}$ and $b = ak$, then $\phi_{jk}(x)$ looks like a discretized version of $\psi_{a,b}$. This is somewhat misleading however because the father function ϕ does *not* give the wavelets as we have introduced them before. Recall that we required that $\int \psi(x)dx = 0$. For the father function, it holds that $\int \phi(x)dx = 1$. What will really give the wavelets is the mother function ψ to be studied in the next section. Like the father ϕ generated orthonormal basis functions for the V_j, the mother function will generate an orthonormal basis for the orthocomplements W_j of the V_j.

When we consider f as a (time) signal, then the approximation at resolution 2^{j-1} is a *blurred* version of the signal approximation at resolution 2^j. Its difference $P_j f - P_{j-1} f$ is the detail signal and belongs to the orthogonal complement W_{j-1} of V_{j-1} in V_j. In the next section we shall study these orthocomplements and construct orthonormal bases for the W_j.

It is possible to construct approximants for f at different scales. Suppose we know $f_j = P_j f \in V_j$ at resolution 2^j, then it is possible to construct the approximant $f_{j-1} = P_{j-1} f \in V_{j-1}$ from this because $V_{j-1} \subset V_j$. The function f_{j-1} contains only half the number of coefficients per unit length as f_j (it is a blurred version). This forms the basis of wavelet expansions as given later. For more details see section 13.

8 The mother function

We know that in multiresolution analysis

$$V_n \subset V_{n+1}.$$

Suppose W_n is the orthogonal complement of V_n in V_{n+1}:

$$V_{n+1} = V_n \oplus W_n.$$

Thus

$$V_0 \oplus \sum_{k=0}^{n} W_k = \bigoplus_{-\infty}^{n} W_k = V_{n+1} \quad \text{and} \quad \bigoplus_{-\infty}^{\infty} W_k = L_2.$$

Now consider the function

$$\psi(x) = \sum_n (-1)^n \bar{c}_{1-n} \phi(2x - n) \in V_1. \tag{8.1}$$

We shall explain in the next section where this definition comes from. In this section we first try to get a bit familiar with the function ψ. It can be proved that (see section 10 below)

$$\psi_{0k}(x) = \psi(x - k)$$

forms an orthonormal basis for W_0 and that more generally, the *wavelets*

$$\boxed{\psi_{nk}(x) = 2^{n/2} \psi(2^n x - k) \quad , \quad n, k \in \mathbb{Z}}$$

are such that $\{\psi_{nk} : k \in \mathbb{Z}\}$ forms an orthonormal basis for W_n.

The function $\psi(x)$ is called the *mother function* or the *wavelet* (function). The mother functions for the previous examples can now be considered. One can easily check the following examples (do it !).

Example 8.1 For the box function ($c_0 = c_1 = 1$)

$$\psi(x) = \begin{cases} 1, & 0 \le x < 1/2 \\ -1, & 1/2 \le x < 1 \end{cases}$$

Figure 5 gives a plot of this wavelet. It is called the Haar wavelet.

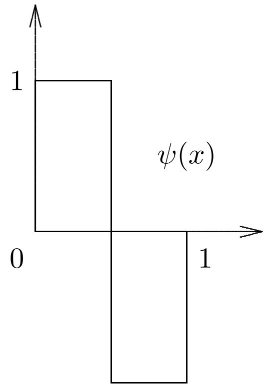


Figure 5: The Haar wavelet.

Example 8.2 The hat function ($c_0 = c_2 = \frac{1}{2}, c_1 = 1$) leads to

$$\psi(x) = \begin{cases} -1/2 - x, & -1/2 \le x \le 0 \\ 3x - 1/2, & 0 \le x \le 1/2 \\ 5/2 - 3x, & 1/2 \le x \le 1 \\ x - 3/2, & 1 \le x \le 3/2 \end{cases}$$

The wavelet is plotted in figure 3.

Note: Not everyone agrees to call the ψ of example 8.2 a wavelet (the scaling function ϕ is not orthogonal to its integer translates. The corresponding function ψ happens to be orthogonal to its integer translates, but most importantly, the $\psi(x-k)$ are not orthogonal to the $\phi(x-l)$.

Example 8.3 For Daubechies D_4, $\psi(x)$ is plotted in figure 4.

In general, when

$$c_n = 0 \quad \text{for} \quad n < N^- \quad \text{and} \quad n > N^+,$$

one can show that

$$\operatorname{supp}(\psi) \subset \left[\frac{1}{2}(1 - N^+ + N^-), \frac{1}{2}(1 + N^+ - N^-)\right].$$

This follows from

$$\psi(x) = \sum_n (-1)^n \bar{c}_{1-n} \phi(2x - n)$$

and the fact that $\operatorname{supp}(\phi) \subset [N^-, N^+]$. First note that $\operatorname{supp}\phi \subset [N^-, N^+]$ implies that $\operatorname{supp}\phi(2x - n)$ as a function of x is $[(N^- + n)/2, (N^+ + n)/2]$. On the other hand c_{1-n} is only nonzero for $n \in [1 - N^+, 1 - N^-]$. Therefore $\operatorname{supp}\psi$ is exactly as stated. We leave the details to the reader.

Let us develop now the analysis for ψ more exactly. It will give answers to questions such as: Where does the defining relation (8.1) of the ψ come from? Do the ψ_{nk} for $k \in \mathbb{Z}$ form indeed an orthonormal basis for the W_n? etc. We shall do this in section 10. First we need more properties of the father function, which we shall derive in the next section.

9 More properties of the scaling function

We know that

$$\hat{\phi}(2\xi) = H(\xi)\hat{\phi}(\xi), \quad H(\xi) = \frac{1}{2}\sum_k c_k e^{-ik\xi} \in L^2_{2\pi}.$$

We shall prove

Theorem 9.1 *The Fourier transform of the father function satisfies*

$$\boxed{\sum_{-\infty}^{\infty} |\hat{\phi}(\xi + 2k\pi)|^2 = \frac{1}{2\pi}.} \tag{9.1}$$

Proof. Use the fact that $\phi(x-k)$ forms an orthonormal basis in V_0, then

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} e^{-im\xi}d\xi = \delta_{0m} &= \int_{\mathbb{R}} \phi(x)\overline{\phi}(x-m)dx \\
&= \int_{\mathbb{R}} e^{-im\xi}|\hat{\phi}(\xi)|^2 d\xi \\
&= \sum_{k=-\infty}^{\infty} \int_{k2\pi}^{(k+1)2\pi} e^{-im\xi}|\hat{\phi}(\xi)|^2 d\xi \\
&= \int_0^{2\pi} e^{-im\xi}(\sum_{k=-\infty}^{\infty} |\hat{\phi}(\xi + 2k\pi)|^2)d\xi.
\end{aligned}$$

The second line is because the Fourier transform defines an isomorphism, the last line because the Fourier transform is continuous. This proves (9.1). □

This has the following consequence.

Corollary 9.2 *The function* $H(\xi) = \frac{1}{2}\sum_k c_k e^{-ik\xi}$ *satisfies*

$$\boxed{|H(\xi)|^2 + |H(\xi + \pi)|^2 = 1} \tag{9.2}$$

Proof. Recall $\hat{\phi}(2\xi) = H(\xi)\hat{\phi}(\xi)$ so that

$$\frac{1}{2\pi} = \sum_k |\hat{\phi}(\xi + 2k\pi)|^2 = \sum_k |\hat{\phi}(2\xi + 2k\pi)|^2 = \sum_k |H(\xi + k\pi)|^2 |\hat{\phi}(\xi + k\pi)|^2.$$

Because H is 2π-periodic,

$$\begin{aligned}
\frac{1}{2\pi} &= \sum_{k \text{ even}} + \sum_{k \text{ odd}} \\
&= |H(\xi)|^2 \sum_k |\hat{\phi}(\xi + 2k\pi)|^2 + |H(\xi + \pi)|^2 \sum_k |\hat{\phi}(\xi + (2k+1)\pi)|^2 \\
&= |H(\xi)|^2 \sum_k |\hat{\phi}(\xi + 2k\pi)|^2 + |H(\xi + \pi)|^2 \sum_k |\hat{\phi}(\xi + \pi + 2k\pi)|^2.
\end{aligned}$$

Hence (9.2) follows. □

Note that the hat function ($c_0 = c_2 = \frac{1}{2}$, $c_1 = 1$) does not satisfy this relation (check it).

Corollary 9.3 *In terms of the* c_k *(9.2) is transformed into*

$$\boxed{\sum_n c_{n-2k}\overline{c}_{n-2\ell} = 2\delta_{k\ell}.} \tag{9.3}$$

Proof. Note that (9.2) means that

$$\frac{1}{4}\sum_{k,l} c_k\overline{c}_l e^{-i(k-l)\xi} + \frac{1}{4}\sum_{k,l} c_k\overline{c}_l(-1)^{k-l}e^{-i(k-l)\xi} = 1.$$

Hence the odd terms drop out and we get

$$\sum_{k-l \text{ even}} c_k\overline{c}_l e^{-i(k-l)\xi} = \sum_t C_t e^{-2it\xi} = 2$$

with

$$C_t = \sum_{k-l=2t} c_k\overline{c}_l.$$

This has to be true for all ξ, so that $C_t = 2\delta_{t0}$. Thus $\sum_n c_n\overline{c}_{n-2t} = 2\delta_{t0}$. This is of course equivalent with the statement to be proved. The complex conjugate drops out if the c_k are real. $\square$

Because $\hat{\phi}(0) = H(0)\hat{\phi}(0)$ and $\hat{\phi}(0) \neq 0$, $H(0) = 1$ and because $|H(\pi)|^2 = 1 - |H(0)|^2 = 0$, we get

$$\boxed{H(\pi) = 0.} \tag{9.4}$$

Corollary 9.4 *In terms of the c_k the previous relation (9.4) becomes*

$$\boxed{\sum_n(-1)^n c_n = 0 \quad \textit{or} \quad 1 = \sum_k c_{2k} = \sum_k c_{2k+1}.} \tag{9.5}$$

Proof. This is just by filling in

$$H(\pi) = \frac{1}{2}\sum_n(-1)^n c_n = 0.$$

$\square$

Note: The orthogonality of ϕ to its integer translates implies (9.2) and (9.4) and hence also (9.3) and (9.5), but the converse is not true. The conditions (9.2) and (9.4) are necessary for this orthogonality but not sufficient. For example the choice $c_0 = c_3 = 1$, $c_1 = c_2 = 0$ defines a ϕ which is not orthogonal to its translates, yet these coefficients satisfy (9.3) and (9.5) as can be easily checked.

We show another interesting property:

Lemma 9.5 *We have*

$$\boxed{\sum_k \phi(x-k) = \sum_k \phi(k) = 1.}$$

Proof. Set $w(x) = \sum_k \phi(x-k)$ then using the dilation equation we find

$$\begin{aligned}
w(x) &= \sum_k\sum_n c_n\phi(2x-2k-n) \\
&= \sum_k(\sum_{n \text{ even}} c_n\phi(2x-(2k+n)) + \sum_{n \text{ odd}} c_n(2x-(2k+n))) \\
&= \sum_k(\sum_\ell c_{2\ell}\phi(2x-2(k+\ell)) + \sum_\ell c_{2\ell+1}\phi(2x-2(k+\ell)-1)) \\
&= \sum_t(\sum_\ell c_{2\ell}\phi(2x-2t) + \sum_\ell c_{2\ell+1}\phi(2x-2t-1)) \\
&= \sum_t \phi(2x-2t)(\sum_\ell c_{2\ell}) + \sum_t \phi(2x-2t-1)(\sum_\ell c_{2\ell+1}) \\
&= \sum_t \phi(2x-t) = w(2x)
\end{aligned}$$

Take the Fourier transform of $w(x) = w(2x)$ to get

$$\hat{w}(\xi) = 2\hat{w}(2\xi).$$

This is a dilation equation with $c_0 = 2$ and has a solution $\hat{w} = \delta$ the Dirac delta function (and $\hat{w} = 0$, but this is excluded because $\int \phi(x)dx = 1$). This means that w has to be constant. Hence $\sum_k \phi(x-k) = c$, a constant and thus independent of x, so it equals $\sum_k \phi(k)$. Now integrate $\sum_k \phi(x-k) = c$ over $[0,1]$ then

$$c = \sum_k \int_0^1 \phi(x-k)dx = \sum_k \int_{x-k}^{x-k+1} \phi(x)dx = \int_{\mathbb{R}} \phi(x)dx = 1.$$

Hence $c = 1$. □

10 Existence of the wavelet

Now let us prove the existence of the mother (wavelet) function.

We know that any $f \in V_0$ can be written as

$$f(x) = \sum_k a_k \phi(x-k)$$

with $(a_k) \in \ell^2$ since the series should converge in L_2. Taking Fourier transforms, this gives

$$\hat{f}(\xi) = \sum_k a_k e^{-ik\xi}\hat{\phi}(\xi) = A_f(\xi)\hat{\phi}(\xi), \quad A_f(\xi) = \sum_k a_k e^{-ik\xi}.$$

Clearly $A_f(\xi) \in L^2_{2\pi}$, i.e. $A_f(\xi)$ is 2π-periodic and

$$\|A_f(\xi)\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |A_f(\xi)|^2 d\xi = \|(a_k)\|^2 < \infty.$$

Thus $f \in V_0 \Leftrightarrow \hat{f} = A_f\hat{\phi}$ with $A_f \in L^2_{2\pi}$.

Theorem 10.1 *There exists a function $\psi \in W_0$ such that $\{\psi(x-k)\}_{k\in\mathbb{Z}}$ forms an orthonormal basis for W_0.*

Proof. If V is a subspace of L_2, then $\hat{V}$ will denote the subspace of L_2 containing all the Fourier transforms of V.
By the Fourier isomorphism:

$$V_{-1} \oplus W_{-1} = V_0 \Leftrightarrow \hat{V}_{-1} \oplus \hat{W}_{-1} = \hat{V}_0.$$

We know that

$$\begin{aligned} \hat{V}_0 &= \{A\hat{\phi} : A \in L^2_{2\pi}\} \\ \hat{V}_{-1} &= \{A(2\cdot)\hat{\phi}(2\cdot) : A \in L^2_{2\pi}\}. \end{aligned}$$

By $\hat{\phi}(2\xi) = H(\xi)\hat{\phi}(\xi)$, we get

$$\hat{V}_{-1} = \{A(2\cdot)H(\cdot)\hat{\phi}(\cdot) : A \in L^2_{2\pi}\}. \tag{10.1}$$

Define the operator (we show below that it is unitary)

$$S : \hat{V}_0 \to L^2_{2\pi} : A\hat{\phi} \mapsto A.$$

Note $S\hat{V}_0 = L^2_{2\pi}$. Instead of computing $\hat{W}_{-1}$ directly, we compute $S(\hat{W}_{-1})$ first, i.e. the orthogonal complement of $S(\hat{V}_{-1})$ in $L^2_{2\pi}$.

Lemma 10.2 *S is a unitary operator.*

Proof. It holds for any $f \in V_0$ that $\hat{f} = A\hat{\phi}$ and that

$$\begin{aligned}\|\hat{f}\|^2 &= \int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi = \int_{\mathbb{R}} |A_f(\xi)|^2 |\hat{\phi}(\xi)|^2 d\xi = \sum_k \int_{k2\pi}^{(k+1)2\pi} |A_f(\xi)|^2 |\hat{\phi}(\xi)|^2 d\xi \\ &= \int_0^{2\pi} |A_f(\xi)|^2 (\sum_k |\hat{\phi}(\xi + 2k\pi)|^2) d\xi = \frac{1}{2\pi} \int_0^{2\pi} |A_f(\xi)|^2 d\xi = \|A_f\|^2.\end{aligned}$$

□

From (10.1) :

$$S(\hat{V}_{-1}) = \{A(2\cdot)H(\cdot) : A \in L^2_{2\pi}\}$$

Let $F \in L^2_{2\pi}$ be in the orthogonal complement of $S(\hat{V}_{-1})$, then

$$\int_0^{2\pi} A(2\xi)H(\xi)\overline{F}(\xi)d\xi = 0 \quad , \quad \forall A \in L^2_{2\pi}.$$

Thus

$$\int_0^{\pi} A(2\xi)[H(\xi)\overline{F}(\xi) + H(\xi+\pi)\overline{F}(\xi+\pi)]d\xi = 0 \quad , \quad \forall A \in L^2_{2\pi}$$

which implies

$$H(\xi)\overline{F}(\xi) + H(\xi+\pi)\overline{F}(\xi+\pi) = 0 \quad , \quad \forall \xi \in \mathbb{R}.$$

This means that in $\mathbb{C}^2$ the vector $\vec{h} = [H(\xi)\ H(\xi+\pi)]$ is orthogonal to the vector $\vec{f} = [F(\xi)\ F(\xi+\pi)]$:

$$\vec{h}\vec{f}^H = 0, \qquad \vec{h}\vec{h}^H = 1,$$

^{H} means complex conjugate transpose. It is clear that

$$\begin{aligned}F(\xi) &= \overline{H}(\xi+\pi) \\ F(\xi+\pi) &= -\overline{H}(\xi)\end{aligned}$$

is a solution. More generally any solution is of the form

$$\begin{aligned}F(\xi) &= -\beta(\xi)\overline{H}(\xi+\pi) \\ F(\xi+\pi) &= \beta(\xi)\overline{H}(\xi).\end{aligned}$$

For convenience, we choose $\beta(\xi) = \alpha(\xi)e^{-i\xi}$, because then

$$\begin{aligned}F(\xi) &= -\alpha(\xi)e^{-i\xi}\overline{H}(\xi+\pi) \\ F(\xi+\pi) &= \alpha(\xi)e^{-i\xi}\overline{H}(\xi)\end{aligned}$$

implies that $\alpha(\xi)$ is π-periodic. Such a function can be written in terms of $e^{i2k\xi}$. Thus we may choose

$$f_k(\xi) = -\sqrt{2}e^{-i\xi}\overline{H}(\xi+\pi)e^{2ki\xi} \quad , \quad k \in \mathbb{Z}$$

as a set of functions in $S(\hat{W}_{-1})$ that generates the whole space. These functions form an orthonormal basis because $(f_k, f_\ell) = (f_{k-\ell}, f_0)$ and using $|H(\xi+\pi)|^2 + |H(\xi)|^2 = 1$ and noting that $e^{2ik\xi}$ is π-periodic, we get

$$\begin{aligned}
(f_k, f_0) &= \frac{1}{\pi}\int_0^{2\pi} e^{2ik\xi}|H(\xi+\pi)|^2 d\xi \\
&= \frac{1}{\pi}\int_0^{\pi} e^{2ik\xi}[|H(\xi+\pi)|^2 + |H(\xi)|^2] d\xi \\
&= \frac{1}{\pi}\int_0^{\pi} e^{2ik\xi} d\xi \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{ik\eta} d\eta \\
&= \delta_{k0}
\end{aligned}$$

Taking the S^{-1} transform we find that

$$\hat{\psi}_{-1,k}(\xi) = -\sqrt{2}e^{-i\xi}\overline{H}(\xi+\pi)\hat{\phi}(\xi)e^{2ki\xi} \quad , \quad k \in \mathbb{Z}$$

is an orthonormal basis for $\hat{W}_{-1}$.

Choosing a function $\psi \in W_0$ with Fourier transform $\hat{\psi}$ satisfying

$$\hat{\psi}(2\xi) = -e^{-i\xi}\overline{H}(\xi+\pi)\hat{\phi}(\xi)$$

we just found that $\sqrt{2}\hat{\psi}(2\xi)e^{2ki\xi}$ forms an orthonormal basis for $\hat{W}_{-1}$. Taking the inverse Fourier transform reveals that $\psi_{-1,k}(x) = \frac{1}{\sqrt{2}}\psi(\frac{x}{2} - k)$ forms an orthonormal basis for W_{-1}. Moreover, after rescaling, we find that $\psi_{0k}(x) = \psi(x-k)$ is an orthonormal basis for W_0.

This concludes the proof of Theorem 10.1. $\square$

Because $e^{i\xi}\hat{\psi}(2\xi)$ is the Fourier transform of $\frac{1}{2}\psi(\frac{x+1}{2})$ and

$$\overline{H}(\xi+\pi)\hat{\phi}(\xi) = \frac{1}{2}\sum_k (-1)^{k+1}\overline{c}_k e^{ik\xi}\hat{\phi}(\xi)$$

we get, after taking Fourier transforms

$$\frac{1}{2}\psi(\frac{x+1}{2}) = \frac{1}{2}\sum_k (-1)^{k+1}\overline{c}_k \phi(x+k)$$

or

$$\boxed{\psi(x) = \sum_k (-1)^k \overline{c}_{1-k}\phi(2x-k).} \tag{10.2}$$

It is not difficult to accept that

$$\{2^{n/2}\psi(2^n x - k) \quad , \quad k \in \mathbb{Z}\}$$

gives an orthonormal basis for W_n, and after taking the limit $L^2 = \bigoplus_n W_n$, we find that

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

forms an orthonormal wavelet basis for L_2.

As an exercise one can prove the following properties. If $H(\xi) = \frac{1}{2}\sum_n c_n e^{-in\xi}$ and $G(\xi) = \frac{1}{2}\sum_n d_n e^{-in\xi}$, $d_n = (-1)^n \overline{c}_{1-n}$, then

$$\hat{\psi}(2\xi) = G(\xi)\hat{\phi}(\xi) \quad \text{and} \quad G(\xi) = -e^{i\xi}\overline{H}(\xi+\pi).$$

This implies that

$$M = \begin{bmatrix} H(\xi) & H(\xi+\pi) \\ G(\xi) & G(\xi+\pi) \end{bmatrix} = \begin{bmatrix} H(\xi) & -e^{-i\xi}\overline{G}(\xi) \\ G(\xi) & -e^{-i\xi}\overline{H}(\xi) \end{bmatrix}$$

satisfies $MM^H = I$, thus that $|H(\xi)|^2 + |G(\xi)|^2 = 1$ and $G(\xi)\overline{H}(\xi) = H(\xi)\overline{G}(\xi) = 0$.

11 Interpretation of the condition on the c_k

The condition $\sum_n c_n \bar{c}_{n-2k} = 2\delta_{k0}$ was derived from the orthogonality of the $\phi(x-k)$ and it implies the orthogonality of the $\psi(x-k)$. One may check that for all the simple examples we have seen whether this condition is satisfied. The box function and D_4 are the only ones which satisfy them.

The box function was the first known (orthogonal) scaling father function. Excluding the delta function, we find that none of the other examples (except D_4) satisfies the above condition and hence none of them is guaranteed to generate wavelets orthogonal to their translates.

The condition $\sum_n (-1)^n c_n = 0$ ensured that $H(\pi) = 0$. This is a special case of a set of more general conditions which require $H(\pi)$ to have a zero of order $p-1$. This would give

$$\sum_n (-1)^n n^k c_n = 0 \quad , \quad k = 0, 1, \ldots, p-1.$$

One can show that for the box function $p = 1$, for the hat function $p = 2$ and for D_4, $p = 2$. The quadratic spline has $p = 3$, the cubic spline $p = 4$.

It can be shown that this condition is related with the degree of the polynomials that can be represented exactly with the corresponding (wavelet) basis functions. We give the following theorem without proof (see [29]).

Theorem 11.1 *If $H^{(k)}(\pi) = 0, \quad k = 0, 1, \ldots, p-1$, then*

1. *The polynomials $1, x, \ldots, x^{p-1}$ are linear combinations of the translates $\phi(x-k)$.*
2. *Smooth functions can be approximated with error $O(h^p)$ by combinations at scale $h = 2^{-j}$:*

$$\exists C, \exists a_k : \|f - \sum_k a_k \phi(2^j x - k)\| \le C 2^{-jp} \|f^{(p)}\|, \quad f \textit{ smooth}$$

3. *The first p moments of $\psi(x)$ are zero:*

$$\int x^m \psi(x) dx = 0 \quad , \quad m = 0, \ldots, p-1$$

4. *The wavelet coefficients $c_j = \int f(x)\psi(2^j x)dx$ decay like $c_j \le C2^{-jp}$.*

The previous observations thus show that the splines are best in approximating, but ... they are not orthogonal! The wavelet D_4 is as good as the linear spline and it is orthogonal. Of course a sine/cosine system is also good in approximation and it is orthogonal, but they do not have compact support as D_4 has. Therefore D_4 is in some sense the simplest genuine wavelet one can imagine.

12 Wavelet expansion and filtering

We want to come to algorithms for wavelet decomposition and wavelet reconstruction. What do we mean by that? Let f_n be in V_n. Because $V_n = V_{n-1} \oplus W_{n-1}$, we can decompose f_n uniquely as

$$f_n = f_{n-1} + g_{n-1} \quad \text{with} \quad f_{n-1} \in V_{n-1}, \quad g_{n-1} \in W_{n-1}.$$

If we repeat this, then

$$f_n = g_{n-1} + g_{n-2} + \cdots + g_{n-m} + f_{n-m}, \quad f_j \in V_j, \quad g_j \in W_j.$$

The integer m is large enough when f_{n-m} is sufficiently "blurred".

$$
\begin{tikzcd}
 & q_{n-1} & q_{n-2} & & q_{n-m+1} & q_{n-m} \\
p_n \arrow[r] \arrow[ur] & p_{n-1} \arrow[r] \arrow[ur] & p_{n-2} & \cdots \arrow[r] \arrow[ur] & p_{n-m+1} \arrow[r] \arrow[ur] & p_{n-m}
\end{tikzcd}
$$

Figure 6: The decomposition scheme.

$$
\begin{tikzcd}
q_{n-m} \arrow[dr] & q_{n-m+1} \arrow[dr] & q_{n-m+2} & {} \arrow[dr] & q_{n-1} \arrow[dr] & \\
p_{n-m} \arrow[r] & p_{n-m+1} \arrow[r] & p_{n-m+2} & \cdots \arrow[r] & p_{n-1} \arrow[r] & p_n
\end{tikzcd}
$$

Figure 7: The reconstruction scheme.

Now suppose that

$$
\begin{aligned}
f_j(x) &= \sum_k p_{jk}\phi_{jk}(x), \quad p_j = (p_{jk}) \in \ell^2 \\
g_j(x) &= \sum_k q_{jk}\psi_{jk}(x), \quad q_j = (q_{jk}) \in \ell^2
\end{aligned}
$$

The decomposition algorithm will decompose p_n into p_{n-1} and q_{n-1}, then p_{n-1} again into p_{n-2} and q_{n-2} etc. like in figure 6. When we want to reconstruct the p_n, the algorithm should perform operations represented schematically in figure 7. The purpose of this section is to find p_{n-1} and q_{n-1} from p_n (decomposition) and to recover p_n from p_{n-1} and q_{n-1} (reconstruction).

We recall

$$
\begin{aligned}
\hat{\phi}(2\xi) = H(\xi)\hat{\phi}(\xi), \qquad & H(\xi) = \frac{1}{2}\sum_k c_k e^{-ik\xi} \\
\hat{\psi}(2\xi) = G(\xi)\hat{\phi}(\xi), \qquad & G(\xi) = -e^{-i\xi}\overline{H}(\xi+\pi) = \frac{1}{2}\sum_k d_k e^{-ik\xi}
\end{aligned}
$$

with $d_k = (-1)^k \overline{c}_{1-k}$.

$$
\phi_{nk}(x) = 2^{n/2}\phi(2^n x - k), \quad \psi_{nk} = 2^{n/2}\psi(2^n x - k).
$$

The projection P_n on V_n and Q_n on W_n are given by

$$
\begin{aligned}
P_n f = \sum_k p_{nk}(f)\phi_{nk}, \qquad & p_{nk}(f) = \langle f, \phi_{nk}\rangle \\
Q_n f = \sum_k q_{nk}(f)\psi_{nk}, \qquad & q_{nk}(f) = \langle f, \psi_{nk}\rangle\,.
\end{aligned}
$$

We want to relate p_{nk} and q_{nk} to $p_{n+1,k}$. We first prove

Lemma 12.1 *We have*

$$
\phi_{nk}(x) = \frac{1}{\sqrt{2}}\sum_l c_{l-2k}\phi_{n+1,l}(x).
$$

Proof. First note that

$$
\phi(\frac{x}{2}) = \sum_k c_k \phi(x-k).
$$

Then

$$\begin{aligned}
\phi_{nk}(x) &= 2^{n/2}\phi(2^n x - k) \\
&= 2^{n/2}\sum_i c_i \phi(2^{n+1}x - 2k - i) \\
&= \frac{1}{\sqrt{2}}\sum_i c_i 2^{(n+1)/2}\phi(2^{n+1}x - (2k+i)) \\
&= \frac{1}{\sqrt{2}}\sum_l c_{l-2k}\phi_{n+1,l}(x)
\end{aligned}$$

which proves the result. □

From their definitions, it thus follows that

$$p_{nk} = \frac{1}{\sqrt{2}}\sum_l \bar{c}_{l-2k} p_{n+1,l}$$

Now define a *filter* $\mathcal{H} : \ell^2 \to \ell^2$, $a = (a_k) \mapsto \mathcal{H}a$ with

$$(\mathcal{H}a)_k = \frac{1}{\sqrt{2}}\sum_l \bar{c}_{l-2k} a_l.$$

Then, setting $p_n = (p_{nk})$, it is clear that

$$p_n = \mathcal{H}p_{n+1}.$$

We do a similar thing for the q_{nk} : Define the filter $\mathcal{G} : \ell^2 \to \ell^2$, $a \mapsto \mathcal{G}a$ with

$$(\mathcal{G}a)_k = \frac{1}{\sqrt{2}}\sum_l \bar{d}_{l-2k} a_l.$$

Lemma 12.2 *We have*

$$\psi_{nk} = \frac{1}{\sqrt{2}}\sum_l d_{l-2k}\phi_{n+1,l}.$$

Proof. This is along the same lines as the previous one

$$\begin{aligned}
\psi_{nk}(x) &= 2^{n/2}\psi(2^n x - k) \\
&= 2^{n/2}\sum_i d_i \phi(2^{n+1}x - 2k - i) \\
&= \frac{1}{\sqrt{2}}\sum_i d_i 2^{(n+1)/2}\phi(2^{n+1}x - (2k+i)) \\
&= \frac{1}{\sqrt{2}}\sum_l d_{l-2k}\phi_{n+1,l}(x)
\end{aligned}$$

□

Thus with the previous definition of $\mathcal{G}$

$$q_n = \mathcal{G}p_{n+1}.$$

Now can easily prove that

Lemma 12.3 *The following relations hold*

$$\begin{aligned} \langle \phi_{nl}, \phi_{n+1,k} \rangle &= \frac{1}{\sqrt{2}} c_{k-2l} \\ \langle \psi_{nl}, \phi_{n+1,k} \rangle &= \frac{1}{\sqrt{2}} d_{k-2l}. \end{aligned}$$

Proof. Also this one is trivial. For example from

$$\phi_{nl}(x) = \frac{1}{\sqrt{2}} \sum_t c_{t-2l} \phi_{n+1,t}(x)$$

we find that

$$\langle \phi_{nl}, \phi_{n+1,k} \rangle = \frac{1}{\sqrt{2}} c_{k-2l}.$$

□

If $\mathcal{F}$ is an operator (filter) on ℓ^2, then the *adjoint* $\mathcal{F}^* : \ell^2 \to \ell^2$ is defined by

$$\langle \mathcal{F}a, b \rangle = \langle a, \mathcal{F}^* b \rangle$$

The matrix representation of the adjoint operator is the Hermitian conjugate of the matrix representation of the operator. Hence the adjoints of $\mathcal{H}$ and $\mathcal{G}$ are given by

$$\begin{aligned} (\mathcal{H}^* a)_k &= \frac{1}{\sqrt{2}} \sum_l c_{k-2l} a_l \\ (\mathcal{G}^* a)_k &= \frac{1}{\sqrt{2}} \sum_l d_{k-2l} a_l \end{aligned}$$

We now express p_{n+1} in terms of p_n and q_n.

$$\begin{aligned} p_{n+1,k} &= \langle f, \phi_{n+1,k} \rangle \\ &= \langle P_{n+1} f, \phi_{n+1,k} \rangle \\ &= \langle P_n f + Q_n f, \phi_{n+1,k} \rangle \\ &= \left\langle \sum_l p_{nl} \phi_{nl}, \phi_{n+1,k} \right\rangle + \left\langle \sum_l q_{nl} \psi_{nl}, \phi_{n+1,l} \right\rangle \\ &= \frac{1}{\sqrt{2}} \sum_l c_{k-2l} p_{nl} + \frac{1}{\sqrt{2}} \sum_l d_{k-2l} q_{nl}. \end{aligned}$$

Thus

$$p_{n+1} = \mathcal{H}^* p_n + \mathcal{G}^* q_n.$$

Remark: The previous decomposition corresponds to a change of basis in V_{n+1} :

$$\{\phi_{n+1,k}, k \in \mathbb{Z}\} \to \{\phi_{n,k}, k \in \mathbb{Z}\} \cup \{\psi_{n,k}, k \in \mathbb{Z}\}.$$

We can use the conditions on the c_k we have found before to find that $\mathcal{H}\mathcal{H}^* = \mathcal{G}\mathcal{G}^* = \mathcal{I}$ (the identity) and $\mathcal{H}\mathcal{G}^* = \mathcal{G}\mathcal{H}^* = \mathcal{O}$ (the zero operator). Somewhat more difficult is to show that $\mathcal{H}^*\mathcal{H} + \mathcal{G}^*\mathcal{G} = \mathcal{I}$. Thus with $\mathcal{K}^* = [\mathcal{H}^* \;\; \mathcal{G}^*]$, we have $\mathcal{K}\mathcal{K}^* = \mathcal{I}$ and $\mathcal{K}^*\mathcal{K} = \mathcal{I}$.

The results of this section can be reflected in the decomposition and reconstruction scheme as in figure 8

$$
\begin{tikzcd}
 & q_n \\
p_{n+1} \arrow[ur, "\mathcal{G}"] \arrow[r, "\mathcal{H}"'] & p_n
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
q_n \arrow[dr, "\mathcal{G}^*"] & \\
p_n \arrow[r, "\mathcal{H}^*"'] & p_{n+1}
\end{tikzcd}
$$

Figure 8: The operations in the decomposition and reconstruction schemes.

13 Fast Discrete Wavelet transform

We suppose from now on that we work with orthogonal compactly supported wavelets, i.e. $c_k = 0$ for $k < 0$ and $k \geq 2N$ and that the coefficients c_k are real.
We want to invert the relation

$$p_{n+1} = [\mathcal{H}^* \ \mathcal{G}^*] \begin{bmatrix} p_n \\ q_n \end{bmatrix}$$

which is done by

$$\begin{bmatrix} \mathcal{H} \\ \mathcal{G} \end{bmatrix} p_{n+1} = \begin{bmatrix} p_n \\ q_n \end{bmatrix}$$

In general the matrices corresponding to $\mathcal{H}$ and $\mathcal{G}$ are infinite dimensional e.g. for $c_k \neq 0, k = 0, 1, 2, 3$

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} \ddots & & & & & & & \\ c_0 & c_1 & c_2 & c_3 & & & & \\ & & \boxed{c_0} & c_1 & c_2 & c_3 & & \\ & & & & c_0 & c_1 & c_2 & c_3 \\ & & & & & & \ddots & \end{bmatrix}, G = \frac{1}{\sqrt{2}} \begin{bmatrix} \ddots & & & & & & \\ \cdots & c_1 & c_0 & & & & \\ & c_3 & -c_2 & \boxed{c_1} & -c_0 & & \\ & & & c_3 & -c_2 & c_1 & -c_0 \\ & & & & & \ddots & \end{bmatrix}$$

However, for practical reasons we work with discrete data vectors of length $M = 2^K$. The operators (filters) $\mathcal{H}$, $\mathcal{G}$ will transform the data vector of length 2^K into another vector of the same length by multiplication with a $2^K \times 2^K$ matrix.

In DFT, this transform can be made fast because the transformation is represented as a product of sparse elementary matrices. Moreover, the transformation matrix is orthogonal, so that its inverse is just equal to its transpose. This matrix is made orthogonal by choosing the unit vectors as basis in the “time” domain and the sines/cosines as basis in the frequency domain.

A similar observation holds for the DWT.

For finite dimensional data, we shall have to truncate the infinite dimensional matrices. So for $M = 2^K$ we make H, G of dimension $2^{K-1} \times 2^K$. E.g., for $K = 3$ and $N = 2$

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & & & & \\ & & c_0 & c_1 & c_2 & c_3 & & \\ & & & & c_0 & c_1 & c_2 & c_3 \\ & & & & & & c_0 & c_1 \end{bmatrix}; G = \frac{1}{\sqrt{2}} \begin{bmatrix} c_3 & -c_2 & c_1 & -c_0 & & & & \\ & & c_3 & -c_2 & c_1 & -c_0 & & \\ & & & & c_3 & -c_2 & c_1 & -c_0 \\ & & & & & & c_3 & -c_2 \end{bmatrix}$$

However this will cause some edge effects (orthogonality is lost). Therefore we suppose that the data are periodically extended, which amounts to reenter cyclically the data in H that were chopped off (a similar technique is used in DFT). So we use instead of the previous H and G for $K = 3$ and $N = 2$ the matrices

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & & & & \\ & & c_0 & c_1 & c_2 & c_3 & & \\ & & & & c_0 & c_1 & c_2 & c_3 \\ c_2 & c_3 & & & & & c_0 & c_1 \end{bmatrix}; G = \frac{1}{\sqrt{2}} \begin{bmatrix} c_3 & -c_2 & c_1 & -c_0 & & & & \\ & & c_3 & -c_2 & c_1 & -c_0 & & \\ & & & & c_3 & -c_2 & c_1 & -c_0 \\ c_1 & -c_0 & & & & & c_3 & -c_2 \end{bmatrix}$$

G, H is "orthogonal" again in the sense $HH^T = I$, $GG^T = I$ and $HG^T = GH^T = 0$. Thus, with $K^T = [H^T \ \ G^T]$: $KK^T = K^TK = I$.
Suppose we have a data vector p of length 2^K. We can write it in terms of 2^K basis functions

$$p = \sum_{k=0}^{2^K-1} p_{k0}\phi_{0k}$$

with $\phi_{0k}(x) = \phi(x-k)$ with coefficients p_{k0}

$$p_{k0} = \langle p, \phi_{0k} \rangle .$$

The inner product, which used to be $\langle f, g\rangle = \int f(x)\overline{g}(x)dx$ will for the discrete case be transformed into $\langle f, g\rangle = \sum_k f(k)\overline{g}(k)$. Hence, for a data vector $p = (p_k)$ with $k = 0, 1, \ldots, 2^K - 1$, we get

$$\langle p, \phi_{0l} \rangle = \sum_{k=0}^{2^K-1} p_k\phi_{0l}(k).$$

The expansion thus has to be understood as

$$\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{2^K-1} \end{bmatrix} = \sum_{k=0}^{2^K-1} p_{k0} \begin{bmatrix} \phi_{0k}(0) \\ \phi_{0k}(1) \\ \vdots \\ \phi_{0k}(2^K-1) \end{bmatrix}$$

E.g., consider $p = [9 \ 1 \ 2 \ 0]^T$ and use for ϕ the box function i.e., $K = 2$. Note that $\phi_{0l}(k) = \delta_{kl}$ and thus $p_{k0} = \langle p, \phi_{0k}\rangle = p_k$.

$$p = 9\begin{bmatrix} \phi_{00}(0) \\ \phi_{00}(1) \\ \phi_{00}(2) \\ \phi_{00}(3) \end{bmatrix} + 1\begin{bmatrix} \phi_{01}(0) \\ \phi_{01}(1) \\ \phi_{01}(2) \\ \phi_{01}(3) \end{bmatrix} + 2\begin{bmatrix} \phi_{02}(0) \\ \phi_{02}(1) \\ \phi_{02}(2) \\ \phi_{02}(3) \end{bmatrix} + 0\begin{bmatrix} \phi_{03}(0) \\ \phi_{03}(1) \\ \phi_{03}(2) \\ \phi_{03}(3) \end{bmatrix}.$$

We can also expand it in terms of

$$\phi_{-1,k} \text{ and } \psi_{-1,k} \quad , \quad k = 0, 1, \ldots, 2^{K-1} - 1$$

with coefficients $p_1 = (p_{k1})$ and $q_1 = (q_{k1})$ given by

$$p_{k1} = \langle p, \phi_{-1,k}\rangle \quad , \quad q_{k1} = \langle p, \psi_{-1,k}\rangle .$$

We could directly compute them by evaluating the inner products. However, by our previous analysis, we can also find them as

$$\begin{bmatrix} p_1 \\ q_1 \end{bmatrix} = \begin{bmatrix} H_K \\ G_K \end{bmatrix} p_0.$$

For our previous example,

$$H_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} c_0 & c_1 & & \\ & & c_0 & c_1 \end{bmatrix}; \quad G_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} c_1 & -c_0 & & \\ & & c_1 & -c_0 \end{bmatrix}$$

so that

$$\begin{bmatrix} p_{01} \\ p_{11} \\ q_{01} \\ q_{11} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & & \\ & & 1 & 1 \\ 1 & -1 & & \\ & & 1 & -1 \end{bmatrix}\begin{bmatrix} 9 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 10 \\ 2 \\ 8 \\ 2 \end{bmatrix}$$

and we can check that this gives indeed the correct decomposition

$$\begin{aligned} p &= \frac{10}{\sqrt{2}}\begin{bmatrix} \phi_{-1,0}(0) \\ \phi_{-1,0}(1) \\ \phi_{-1,0}(2) \\ \phi_{-1,0}(3) \end{bmatrix} + \frac{2}{\sqrt{2}}\begin{bmatrix} \phi_{-1,1}(0) \\ \phi_{-1,1}(1) \\ \phi_{-1,1}(2) \\ \phi_{-1,1}(3) \end{bmatrix} + \frac{8}{\sqrt{2}}\begin{bmatrix} \psi_{-1,0}(0) \\ \psi_{-1,0}(1) \\ \psi_{-1,0}(2) \\ \psi_{-1,0}(3) \end{bmatrix} + \frac{2}{\sqrt{2}}\begin{bmatrix} \psi_{-1,1}(0) \\ \psi_{-1,1}(1) \\ \psi_{-1,1}(2) \\ \psi_{-1,1}(3) \end{bmatrix} \\ &= \frac{10}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{2}{2}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \frac{8}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \frac{2}{2}\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 9 \\ 1 \\ 2 \\ 0 \end{bmatrix}! \end{aligned}$$

The first and second term, are the components of p along $\phi_{-1,0}$ and $\phi_{-1,1}$. Together they form the part of p that is in V_{-1}. This can again be partitioned and written in terms of

$$\phi_{-2,k} \text{ and } \psi_{-2,k} \quad , \quad k = 0, 1, \ldots, 2^{K-2} - 1$$

which in our example is

$$\phi_{-2,0} \text{ and } \psi_{-2,0} \quad \rightarrow \quad \text{coefficients } q_{02} \text{ and } q_{02}.$$

This is for our example given by

$$\begin{bmatrix} p_{02} \\ q_{02} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} H_1 \\ G_1 \end{bmatrix}\begin{bmatrix} p_{01} \\ p_{11} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} c_0 & c_1 \\ c_1 & -c_0 \end{bmatrix}\begin{bmatrix} p_{01} \\ p_{11} \end{bmatrix}$$

thus, explicitly

$$\begin{bmatrix} p_{02} \\ q_{02} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 10 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

and indeed

$$\begin{bmatrix} 5 \\ 5 \\ 1 \\ 1 \end{bmatrix} = 6\begin{bmatrix} \phi_{-2,0}(0) \\ \phi_{-2,0}(1) \\ \phi_{-2,0}(2) \\ \phi_{-2,0}(3) \end{bmatrix} + 4\begin{bmatrix} \psi_{-2,0}(0) \\ \psi_{-2,0}(1) \\ \psi_{-2,0}(2) \\ \psi_{-2,0}(3) \end{bmatrix} = \frac{6}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{4}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

is equal to the sum of the first two terms in the previous decomposition. Thus we have written p as

$$\begin{aligned} p &= 6\phi_{-2,0} + 4\psi_{-2,0} + \frac{8}{\sqrt{2}}\psi_{-1,0} + \frac{2}{\sqrt{2}}\psi_{-1,1} \\ &= 6\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} + 4\begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix} + \frac{8}{\sqrt{2}}\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{bmatrix} + \frac{2}{\sqrt{2}}\begin{bmatrix} 0 \\ 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \end{aligned}$$

Note that in this simple example we didn't need the wrap around of the H and G matrix.
The two transformations together give the result $[p_{02}\ q_{02}\ q_{01}\ q_{11}]^T$ in terms of $[p_{00}\ p_{10}\ p_{20}\ p_{30}]^T$ as

$$\begin{aligned} \begin{bmatrix} p_{02} \\ q_{02} \\ q_{01} \\ q_{11} \end{bmatrix} &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & \sqrt{2} & \\ & & & \sqrt{2} \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & & \\ & & 1 & 1 \\ 1 & -1 & & \\ & & 1 & -1 \end{bmatrix}\begin{bmatrix} p_{00} \\ p_{10} \\ p_{20} \\ p_{30} \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}\begin{bmatrix} 9 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 8/\sqrt{2} \\ 2/\sqrt{2} \end{bmatrix} \end{aligned}$$

The filters (matrices) G and H are often intertwined to have a matrix like e.g.

$$T = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & & & & & & \\ c_3 & -c_2 & c_1 & -c_0 & & & & & & \\ & & c_0 & c_1 & c_2 & c_3 & & & & \\ & & c_3 & -c_2 & c_1 & -c_0 & & & & \\ & & & & & \ddots & & & & \\ & & & & & & c_0 & c_1 & c_2 & c_3 \\ & & & & & & c_3 & -c_2 & c_1 & -c_0 \\ c_2 & c_3 & & & & & & & c_0 & c_1 \\ c_1 & -c_0 & & & & & & & c_3 & -c_2 \end{bmatrix}$$

and the factors $\frac{1}{\sqrt{2}}$ are included in the coefficients.
If this matrix has to be orthogonal, then, multiplying it with its transpose should give the identity. This results in

$$\begin{cases} c_0^2 + c_1^2 + c_2^2 + c_3^2 = 1 \\ c_2c_0 + c_3c_1 = 0 \end{cases}$$

If we require in addition the approximation to be of order 2.
Then

$$\begin{cases} c_0 - c_1 + c_2 - c_3 = 0 & (H(\pi) = 0) \\ 0c_0 - c_1 + 2c_2 - 3c_3 = 0 & (H'(\pi) = 0) \end{cases}$$

A solution of these 4 equations is given by

$$\begin{aligned} c_0 &= \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad c_1 = \frac{3+\sqrt{3}}{4\sqrt{2}} \\ c_2 &= \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad c_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} \end{aligned}$$

which is Daubechies D_4.
When similarly introducing 6 coefficients $c_0, \ldots, c_5$, the orthogonality requirement gives 3 conditions, so that we can require the order to be 3 giving 3 more conditions.
A solution is given by D_6:

$$\begin{aligned} c_0 &= \frac{(1+\sqrt{10}+\sqrt{5+2\sqrt{10}})}{16\sqrt{2}}, & c_1 &= \frac{5+\sqrt{10}+3\sqrt{5+2\sqrt{10}}}{16\sqrt{2}} \\ c_2 &= \frac{10-2\sqrt{10}+2\sqrt{5+2\sqrt{10}}}{16\sqrt{2}}, & c_3 &= \frac{10-2\sqrt{10}+2\sqrt{5+2\sqrt{10}}}{16\sqrt{2}} \\ c_4 &= \frac{5+\sqrt{10}-3\sqrt{5+2\sqrt{10}}}{16\sqrt{2}}, & c_5 &= \frac{1+\sqrt{10}+\sqrt{5+2\sqrt{10}}}{16\sqrt{2}} \end{aligned}$$

One can check this as an exercise.

If the matrix is written in this intertwined form, after each transformation, one needs a permutation.

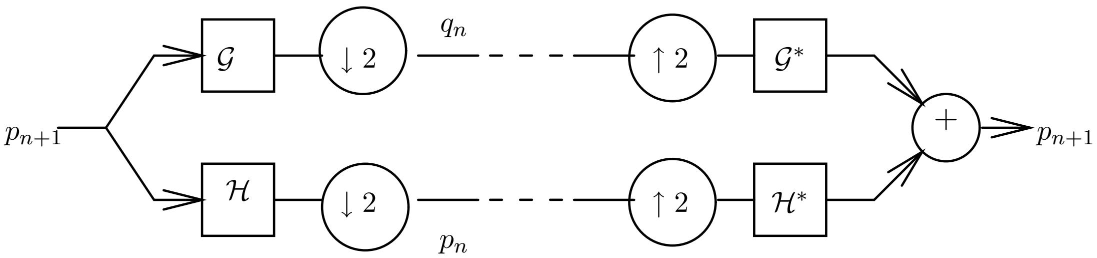
Thus e.g.

$$
\begin{array}{ccccccccccccccc}
p_{0,0} & & p_{0,1} & & p_{0,1} & & p_{0,2} & & p_{0,2} & & p_{0,3} & & p_{0,3} & \xrightarrow{T_1} & p_{0,4} \\
p_{1,0} & & q_{0,1} & & p_{1,1} & & q_{0,2} & & p_{1,2} & \xrightarrow{T_2} & q_{0,3} & \xrightarrow{\text{Perm}} & p_{1,3} & & q_{0,4} \\
p_{2,0} & & p_{1,1} & & p_{2,1} & & p_{1,2} & & p_{2,2} & & p_{1,3} & & q_{0,2} & & \\
p_{3,0} & & q_{1,1} & & p_{3,1} & \xrightarrow{T_3} & q_{1,2} & \xrightarrow{\text{Perm}} & p_{3,2} & & q_{1,3} & & q_{1,2} & & \\
p_{4,0} & & p_{2,1} & & p_{4,1} & & p_{2,2} & & q_{0,2} & & & & & & \\
p_{5,0} & & q_{2,1} & & p_{5,1} & & q_{2,2} & & q_{1,2} & & & & & & \\
p_{6,0} & & p_{3,1} & & p_{6,1} & & p_{3,2} & & q_{2,2} & & & & & & \\
p_{7,0} & \xrightarrow{T_4} & q_{3,1} & \xrightarrow{\text{Perm}} & p_{7,1} & & q_{3,2} & & q_{3,2} & & & & & & \\
p_{8,0} & & p_{4,1} & & q_{0,1} & & & & & & & & & & \\
p_{9,0} & & q_{4,1} & & q_{1,1} & & & & & & & & & & \\
p_{10,0} & & p_{5,1} & & q_{2,1} & & & & & & & & & & \\
p_{11,0} & & q_{5,1} & & q_{3,1} & & & & & & & & & & \\
p_{12,0} & & p_{6,1} & & q_{4,1} & & & & & & & & & & \\
p_{13,0} & & q_{6,1} & & q_{5,1} & & & & & & & & & & \\
p_{14,0} & & p_{7,1} & & q_{6,1} & & & & & & & & & & \\
p_{15,0} & & q_{7,1} & & q_{7,1} & & & & & & & & & &
\end{array}
$$

The p-coefficients at a certain level give the coarse (smooth, large scale) information about the data vector while the q-coefficients give the more detailed information (small scale).

FORTRAN programs for this transform can be found in [26].

In signal processing literature, this technique is known as subband coding. This means that for the analysis, the signal p_{n+1} is filtered by both the filters $\mathcal{H}$ (lowpass filter) and $\mathcal{G}$ (passband filter). The results p_n and q_n are both subsampled ($\downarrow 2$) this means that only the samples with an even index are kept, the rest is thrown away. In the reconstruction, the signals p_n and q_n are first upsampled ($\uparrow 2$) which means that a zero is introduced between 2 consecutive samples and the upsampled signals are filtered by $\mathcal{G}^*$ and $\mathcal{H}^*$ and the result is added to give p_{n+1}.



14 Truncated wavelet approximation

A most interesting aspect of DWT is that in many applications, most of the coefficients of the DWT are very small and therefore we can set them equal to zero and yet retain the most important information.

The following example is similar to the one given in [26].

Consider a function like in fig. 9, sampled at 1024 equidistant points. The function values are integers generated by a shifted and rescaled $\log |x|$ function. Taking the DWT gives 1024 coefficients corresponding to certain wavelets. We rounded the wavelet coefficients to integer values between 0 and 255. Only 7% of these coefficients is not rounded to zero and 5% is larger than 1. They are plotted on a logarithmic scale on the second graph of fig. 9. Only those wavelets whose support contains the cusp of the given function will really contribute. Thus it already gives a good approximation if we select only those 5% wavelets (coefficient + *position* has to be stored!)

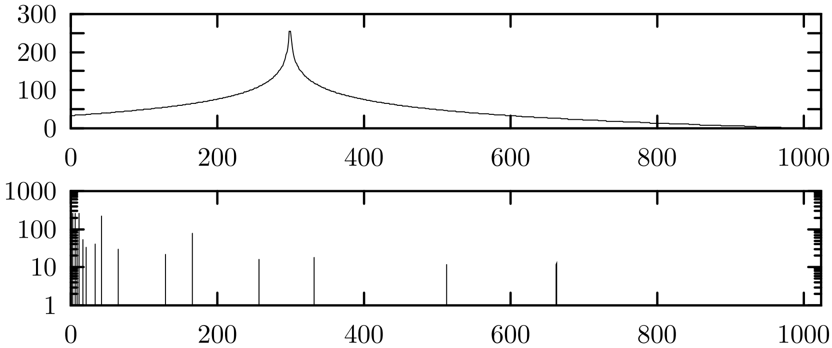


Figure 9: Function with cusp

Exercise 14.1 How many bits does one need to store the position? Conclude that in comparison with storage for the coefficients, the storage needed for position is negligible. $\Diamond$

Note that the cusp is at about one third of the interval [1,1024]. The nonzero wavelet coefficients are located at about one third of the intervals [513,1024], [257,512], [129,256], [65,128], etc. Also near the edges $\downarrow$ at 512, 256, 128,…, one gets nonzero coefficients (that is due to the edge effect).

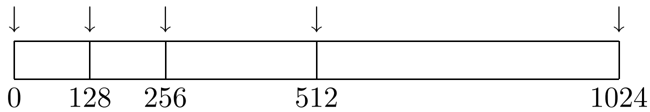


Figure 10: Edge effect

In general compact (hence unsmooth) wavelets are better for lower accuracy approximation and for functions with discontinuous (typical for edge detection in image processing) while smooth (hence noncompact) wavelets are better for high numerical accuracy.

Another example is the DWT of the sine function given in fig. 11. Note that the DWT gives

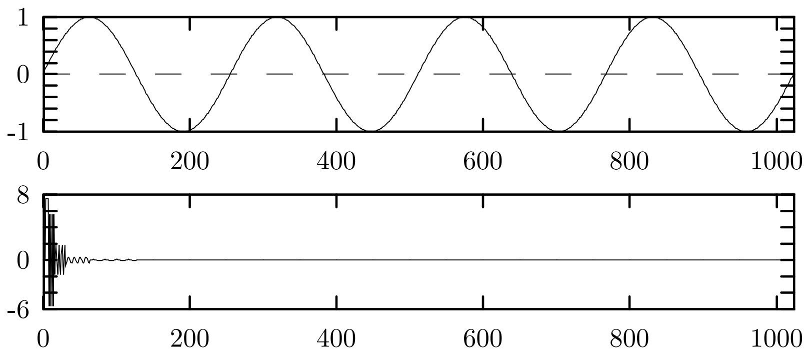


Figure 11: DWT of sine

several significant coefficients for the sine whereas the DFT would give only one nonzero coefficient. This is the price we pay for having a basis of wavelets which is local in the time as well as in the frequency domain. Note however that the number of significant coefficients is still small. The price to pay is relatively small.

15 Biorthogonal wavelets

The orthogonality condition imposed rather strong conditions on the definition of the wavelet. Therefore weaker forms exist, like e.g., biorthogonal wavelets. Then some dual functions for ϕ and ψ are defined

$$\tilde{\phi}(x) = \sum_k \tilde{c}_k \tilde{\phi}(2x-k) \quad \text{and} \quad \tilde{\psi}(x) = \sum_k \tilde{d}_k \tilde{\phi}(2x-k)$$

or

$$\hat{\tilde{\phi}}(2\xi) = \tilde{H}(\xi)\hat{\tilde{\phi}}(\xi) \quad \text{and} \quad \hat{\tilde{\psi}}(2\xi) = \tilde{G}(\xi)\hat{\tilde{\phi}}(\xi)$$

with

$$\tilde{H}(\xi) = \frac{1}{2}\sum_k \tilde{c}_k e^{-ik\xi} \quad \text{and} \quad \tilde{G}(\xi) = \frac{1}{2}\sum_k \tilde{d}_k e^{-ik\xi}.$$

It is now required that the following biorthogonality relations hold

$$\left\langle \phi_{nk}, \tilde{\phi}_{nl} \right\rangle = \delta_{kl},\ n,k,l \in \mathbb{Z} \quad \text{and} \quad \left\langle \psi_{ik}, \tilde{\psi}_{jl} \right\rangle = \delta_{ij}\delta_{kl},\ i,j,k,l \in \mathbb{Z}$$

These conditions lead to

$$\begin{bmatrix} \tilde{H}(\xi) & \tilde{H}(\xi+\pi) \\ \tilde{G}(\xi) & \tilde{G}(\xi+\pi) \end{bmatrix} \begin{bmatrix} H(\xi) & H(\xi+\pi) \\ G(\xi) & G(\xi+\pi) \end{bmatrix}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The projection operators P_j on V_j and Q_j on W_j are given by

$$P_j f(x) = \sum_k \left\langle f, \tilde{\phi}_{jk} \right\rangle \phi_{jk}(x) = \sum_k p_{jk}\phi_{jk}(x) \quad \text{and} \quad Q_j f(x) = \sum_k \left\langle f, \tilde{\psi}_{jk} \right\rangle \psi_{jk}(x) = \sum_k q_{jk}\psi_{jk}(x).$$

Analysis and reconstruction formulas are

$$p_{n+1} = [\mathcal{H}^* \ \ \mathcal{G}^*] \begin{bmatrix} p_n \\ q_n \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} p_n \\ q_n \end{bmatrix} = \begin{bmatrix} \tilde{\mathcal{H}} \\ \tilde{\mathcal{G}} \end{bmatrix} p_{n+1}.$$

16 Algorithms

The formulas we have seen in the previous section can be easily put into an algorithm. The following algorithms can be found in [18]. We suppose that all the coefficients are real and that c_k and $\tilde{c}_k$ are nonzero for $-L \le k \le L$ and that the $\tilde{d}_k$ and d_k are nonzero for $-M \le k \le M$. Moreover, suppose that $L = 2L' + 1$ and $M = 2M' + 1$ are odd. The "signal" is given as a vector of 2^K coefficients $p_{n,k}$, $k = 0, \ldots, 2^K - 1$.

The analysis is the result of the following DWT algorithm.

for $n = K - 1(-1)0$
  for $k = 0(1)2^n - 1$

$$p_{nk} = \frac{1}{\sqrt{2}} \sum_{i=-L}^{L} \tilde{c}_i p_{n+1,(i+2k) \bmod 2^{n+1}}$$

$$q_{nk} = \frac{1}{\sqrt{2}} \sum_{i=-M}^{M} \tilde{d}_i p_{n+1,(i+2k) \bmod 2^{n+1}}$$

  endfor
endfor

The inverse DWT is given by the following reconstruction algorithm

for $n = 1(1)K$
  for $k = 0(1)2^n - 1$
    if k even **then**

$$p_{nk} = \frac{1}{\sqrt{2}} \sum_{i=-L'}^{L'} c_{2i} p_{n-1,(k/2-i) \bmod 2^{n-1}} + \frac{1}{\sqrt{2}} \sum_{i=-M'}^{M'} d_{2i} q_{n-1,(k/2-i) \bmod 2^{n-1}}$$

    else k odd

$$p_{nk} = \frac{1}{\sqrt{2}} \sum_{i=-L'-1}^{L'} c_{2i+1} p_{n-1,((k-1)/2-i) \bmod 2^{n-1}} + \frac{1}{\sqrt{2}} \sum_{i=-M'-1}^{M'} d_{2i+1} q_{n-1,((k-1)/2-i) \bmod 2^{n-1}}$$

    endif
  endfor
endfor

17 Multidimensional DWT

A wavelet transform of a d-dimensional vector is most easily obtained by transforming the array sequentially on its first index (for all values of its other indices), then on the second etc. Each transformation corresponds to a multiplication with an orthogonal matrix. By associativity of the matrix product, the result is independent of the order in which the indices are chosen.
The situation is identical to multidimensional FFT. For a two-dimensional array (a square image say), this corresponds to a subdivision of the vector into regions like in fig. 12 A. This gives rise to

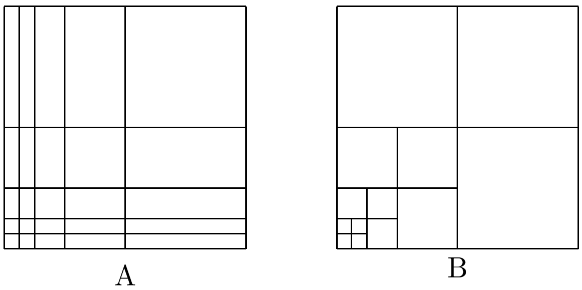


Figure 12: Different subdivisions of square

standard bases to span the spaces of the multiresolution

$$V_k = \operatorname{span}\{\phi_{kn}(x)\phi_{km}(y) : m, n \in \mathbb{Z}\}.$$

Exercise 17.1 Could you find the bases that span the orthocomplements W_k? $\Diamond$

Setting by definition

$$(f \otimes g)(x, y) = f(x)g(y),$$

we get an expansion of the form

$$f(x, y) = \sum_{i,l} \sum_{j,k} \langle f, \psi_{il} \otimes \psi_{jk} \rangle \, (\psi_{il} \otimes \psi_{jk})(x, y).$$

This corresponds to taking a 1-dimensional wavelet transform in x and y independently. Note that this gives a different resolution in x- and y-direction: In the x-direction the scaling is 2^{-i} while in the y-direction the scaling is 2^{-j}.

There is however another possibility where we subdivide the square into regions like in fig. 12 B. At each stage, only the lower left quarter is further subdivided. This case gives subspaces V_n in the multiresolution which, as before, are spanned by

$$V_k = \text{span}\{\phi_{kn}(x)\phi_{km}(y) : m, n \in \mathbb{Z}\} \quad \text{(left lower squares)}$$

but the orthocomplements are spanned by mixtures of basis functions which are now easy to describe:

$$W_k = \text{span}\{\phi_{kn}(x)\psi_{km}(y), \psi_{kn}(x)\phi_{km}(y), \psi_{kn}(x)\psi_{km}(y) : m, n \in \mathbb{Z}\}.$$

The first set are for the right bottom quarters, the second set for the left top quarters and the last one for the right top quarters.

Note that there is now only one scaling 2^{-k} for both x- and y-direction.

It is the latter approach we shall follow in the next section.

18 Image compression

In this application, we are concerned with a two dimensional multiresolution problem. This can be reconstructed using products of functions defining the 1-dimensional multiresolution in x and y. We now have

$$\begin{aligned} V_{k+1} &= V_{k+1}^{(x)} \otimes V_{k+1}^{(y)} \\ &= (V_k^{(x)} \oplus W_k^{(x)}) \otimes (V_k^{(y)} \oplus W_k^{(y)}) \\ &= (V_k^{(x)} \otimes V_k^{(y)}) \oplus (V_k^{(x)} \otimes W_k^{(y)}) \oplus (W_k^{(x)} \otimes V_k^{(y)}) \oplus (W_k^{(x)} \otimes W_k^{(y)}) \end{aligned}$$

The different resolution spaces are

$$V_k = \text{span}\{\phi_{kn}(x)\phi_{km}(y) : m, n \in \mathbb{Z}\}$$

and the orthocomplements are

$$W_k = \text{span}\{\phi_{kn}(x)\psi_{km}(y), \psi_{kn}(x)\phi_{km}(y), \psi_{kn}(x)\psi_{km}(y) : m, n \in \mathbb{Z}\}$$

The projectors are

$$P_k f = \sum_{m,n} p_{mn}^k \phi_{km}(x)\phi_{kn}(y)$$

and

$$Q_k f = \sum_{m,n} [q_{mn}^{(x)k} \psi_{km}(x)\phi_{kn}(y) + q_{mn}^{(y)k} \phi_{km}(x)\psi_{kn}(y) + q_{mn}^{(xy)k} \psi_{km}(x)\psi_{kn}(y)].$$

The coefficients are found by the recursions

$$\begin{aligned} p^{k+1} &= H_x H_y p^k \\ q^{(x)k+1} &= G_x H_y p^k \\ q^{(y)k+1} &= H_x G_y p^k \\ q^{(xy)k+1} &= G_x G_y p^k. \end{aligned}$$

An index x means that the operator acts on the first index, an index y means that the operator acts on the second index, e.g.

$$\left.\begin{array}{c}(H_x a)_{nm} = \sum_k c_{k-2n} a_{km} \\ (G_y p)_{km} = \sum_l d_{l-2m} p_{kl}\end{array}\right\} \Rightarrow (H_x G_y p)_{nm} = \sum_{k,l} c_{k-2n} d_{l-2m} p_{kl}$$

As for the 1-dimensional case, the p^{k+1} give the coarser information, while the q-vectors give the small scale information. For example, a high value in $q^{(y)k+1}$ indicates horizontal edges and $q^{(xy)k+1}$ indicates corners.

The reconstruction algorithm uses

$$p^{k-1} = H_x^* H_y^* p^k + G_x^* H_y^* q^{(x)k} + H_x^* G_y^* q^{(y)k} + G_x^* G_y^* q^{(xy)k}.$$

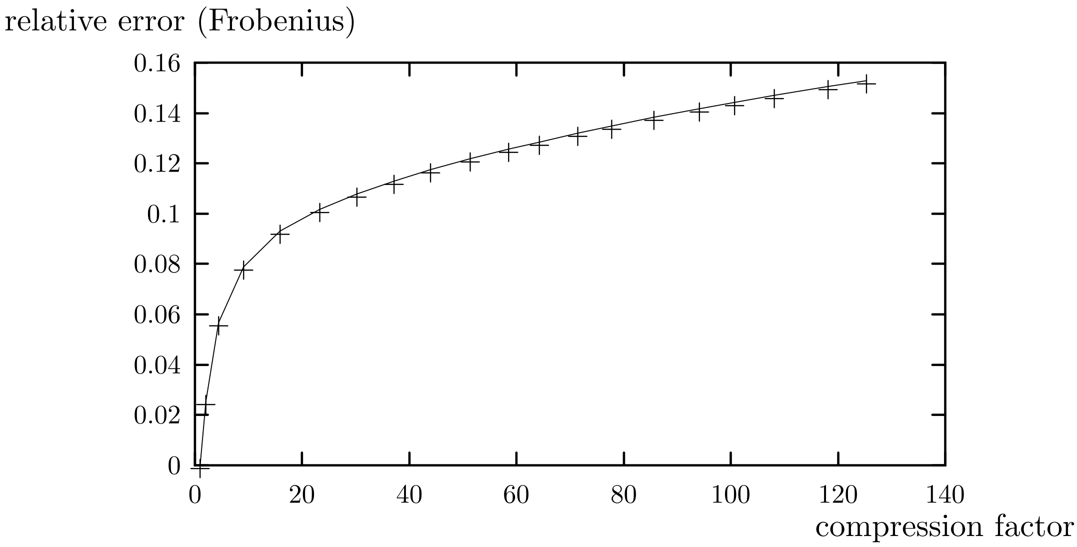
We consider a 256×256 matrix of gray scales from 0–255 (8 bits), which produces the picture of Madonna below.

We use the D_4 discrete wavelet transform in 2 dimensions.

For $p = 0(0.1)1$, we delete all the DWT coefficients that are less than $p\%$ of the maximal coefficient. It turns out that we then use only $r\%$ of the total number of coefficients, i.e., $(100-r)\%$ is set equal to zero. This corresponds to a compression factor $C = 100/r$. To get an idea of the approximation error, we compute the approximation error $E = \|P - A\|_F / \|P\|_F$ where P is the original picture and A is the approximation and $\|\cdot\|_F$ is the Frobenius norm ($\|A\|_F = (\sum_{ij} a_{ij}^2)^{1/2}$). The result is given in the table below.

p	C	r	E
0.0	1.00	100.00	0.00
0.1	4.42	22.62	0.06
0.2	15.93	6.27	0.09
0.3	30.18	3.31	0.11
0.4	43.98	2.27	0.12
0.5	58.57	1.71	0.13
0.6	71.39	1.40	0.13
0.7	85.67	1.17	0.14
0.8	100.67	0.99	0.14
0.9	118.08	0.85	0.15
1.0	135.40	0.74	0.16

In the graph below we plotted these data.

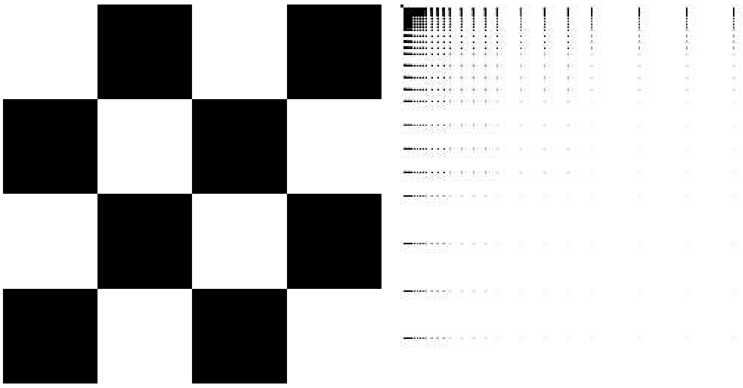


It is possible to see what compression factor can be obtained for a given tolerance in the relative error.

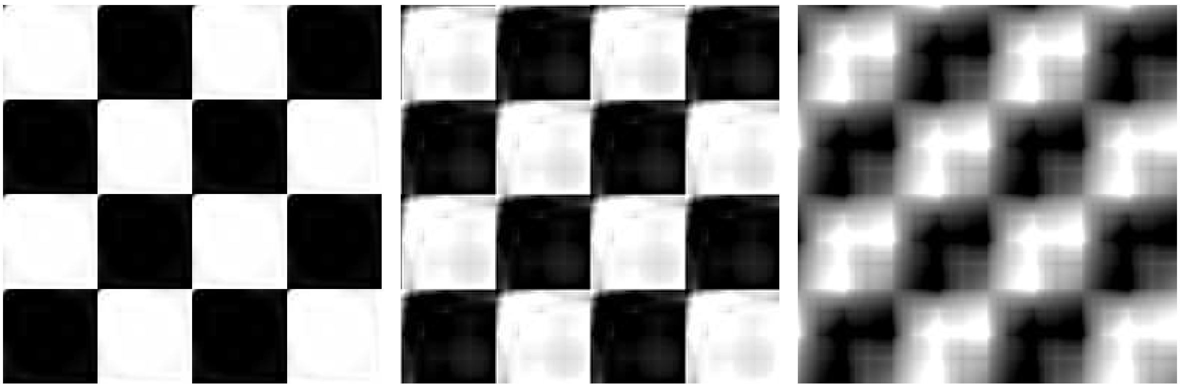
The pictures below give the approximating picture corresponding to these values of $p = 0(0.1)1$.

Below are the results of the picture where all wavelet coefficients smaller than 1 percent of the largest are deleted when the wavelet used is D_4, D_{12} and D_{20} respectively. The compression factors are 135, 146 and 145 respectively. Note that for D_{20} the picture is smoother than for D_4. This smoother situation is closer to what would be obtained by a Discrete Fourier Transform.

Another example is provided by the following checkerboard. Its DWT is given besides the original.



Below we give the reconstructed picture, on the first row for D_4 and on the second row for D_{20} when the relative percent p of deleted wavelet coefficients (as in the previous example) is respectively 1%, 5% and 20%.



19 Solution of linear systems

To solve the problem $Ax = b$, with wavelet transform, the basic idea is to think of the matrix A (supposed to be very large) as a digital image. Suppose, it compresses well under 2-dimensional wavelet transform, i.e., a large fraction of its wavelet coefficients are small (i.e. negligible), then under the wavelet transform, it becomes a sparse system. Let W be the 1-dimensional wavelet transform matrix and set

$$\tilde{A} = WAW^T \quad , \quad \tilde{b} = Wb$$

then solve the sparse system $\tilde{A}\tilde{x} = \tilde{b}$ and finally find x as $x = W^T\tilde{x}$.

An example for which this technique is particularly useful is to solve an integral equation

$$\int A(x, y)f(x)dx = g(y)$$

where $A(x, y)$ is nearly singular near the diagonal $x = y$, but is very smooth away from the diagonal. A discretization gives a matrix which may e.g. look like

$$A_{ij} = \begin{cases} -1, & i = j \\ |i - j|^{-1/2}, & \text{otherwise} \end{cases}$$

For more details see for example [3].

Acknowledgement

Some introductory textbooks and survey papers on the subject are [2, 28, 19, 23, 25, 4, 5, 24, 9, 1, 2, 11, 18, 20]. The computations were done with the FORTRAN routines described in [26]. The knowledge about wavelets is rapidly expanding. An electronic digest is available which gives all kinds of information about wavelet events. You can subscribe by sending an email to

`wavelet@math.scarolina.edu`

with “subscribe” as subject.

Preprints and long lists of references as well as back issues of the digest can be obtained from the following electronic information servers:

Ftp: `ftp.math.scarolina.edu (/pub/wavelet)`
Gopher: `gopher.math.scarolina.edu`
WWW: `http://www.math.scarolina.edu/~wavelet`

Appendices

A Fourier transforms

The Fourier transform of a function $f \in L^2(\mathbb{R})$ is given by

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ix\xi} f(x) dx$$

and the inverse Fourier transform is

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ix\xi} \hat{f}(\xi) d\xi.$$

The Plancherel formula is

$$\int_{\mathbb{R}} |f(x)|^2 dx = \int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi$$

and the Parseval relation is

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)\overline{g(x)} dx = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi = \left\langle \hat{f}, \hat{g} \right\rangle.$$

The Fourier transform is an isometric isomorphism from $L^2(\mathbb{R})$ into $L^2(\mathbb{R})$

L^2_P is the space of P-periodic functions for which

$$\int_0^P |f(x)|^2 dx < \infty.$$

The Fourier transform for functions in $L^2_{2\pi}$ is

$$\hat{f}_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} f(\theta) d\theta, \quad k \in \mathbb{Z}$$

The inverse transform is

$$f(\theta) = \sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\theta}.$$

The Fourier transform is an isometric isomorphism from $L^2_{2\pi}$ into $\ell^2(\mathbb{Z})$

$$\|f\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|^2 d\theta = \|\hat{f}\|^2 = \sum_{k \in \mathbb{Z}} |\hat{f}_k|^2$$

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\overline{g(\theta)} d\theta = \langle \hat{f}, \hat{g} \rangle = \sum_{k \in \mathbb{Z}} \hat{f}_k \overline{\hat{g}}_k.$$

B A collection of formulas

$$\phi(x) = \sum_n c_n \phi(2x - n), \qquad \psi(x) = \sum_n d_n \phi(2x - n)$$

$$\phi_{nk}(x) = 2^{n/2} \phi(2^n x - k), \qquad \psi_{nk}(x) = 2^{n/2} \psi(2^n x - k)$$

$$\hat{\phi}(2\xi) = H(\xi)\hat{\phi}(\xi), \qquad H(\xi) = \frac{1}{2} \sum_n c_n e^{-in\xi}$$

$$\hat{\psi}(2\xi) = G(\xi)\hat{\phi}(\xi), \qquad G(\xi) = \frac{1}{2}\sum_n d_n e^{-in\xi}$$

$$\hat{\phi}(\xi) = \frac{1}{\sqrt{2\pi}} \prod_{n=1}^{\infty} H(2^{-n}\xi)$$

$$\langle \phi_{nk}, \phi_{nl} \rangle = \delta_{kl}, \qquad \langle \psi_{ik}, \psi_{jl} \rangle = \delta_{ij}\delta_{kl}, \qquad \langle \phi_{nk}, \psi_{nj} \rangle = 0$$

$$K = [H(\xi) \;\; G(\xi)] \quad \Rightarrow \quad KK^* = 1, \qquad K^*K = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$$

$$\sum_n c_{2n} = \sum_n c_{2n+1} = 1, \qquad \sum_n c_n = 2, \qquad \sum_n (-1)^n c_n = 0$$

$$\sum c_{n-2k}\bar{c}_{n-2l} = 2\delta_{kl}$$

$$G(\xi) = -e^{-i\xi}\overline{H}(\xi + \pi), \qquad d_n = (-1)^n \bar{c}_{1-n}$$

$$\sum_n \phi(x - n) = 1, \qquad \sum_n |\hat{\phi}(\xi + 2n\pi)|^2 = \frac{1}{2\pi}$$

$$H(\pi) = 0, \quad H(0) = 1; \qquad G(\pi) = 1, \quad G(0) = 0$$

$$(\mathcal{H}a)_k = \frac{1}{\sqrt{2}}\sum_l \bar{c}_{l-2k}a_l, \qquad (\mathcal{G}a)_k = \frac{1}{\sqrt{2}}\sum_l \bar{d}_{l-2k}a_l,$$

$$P_n f = \sum_k p_{nk}\phi_{nk}, \qquad Q_n f = \sum_k q_{nk}\psi_{nk}$$

$$\mathcal{K}^* = [\mathcal{H}^* \;\; \mathcal{G}^*] \quad \Rightarrow \quad \mathcal{K}^*\mathcal{K} = \mathcal{I}, \qquad \mathcal{K}\mathcal{K}^* = \left[\begin{array}{cc} \mathcal{I} & \mathcal{O} \\ \mathcal{O} & \mathcal{I} \end{array}\right]$$

$$\left[\begin{array}{c} p_n \\ q_n \end{array}\right] = \mathcal{K}p_{n+1}, \qquad p_{n+1} = \mathcal{K}^*\left[\begin{array}{c} p_n \\ q_n \end{array}\right]$$

References

[1] A.N. Akansu and R.A. Haddad. *Multiresolution signal decomposition: transforms, subbands and wavelets.* Academic Press/Harcourt Brace Jovanovich, 1992.

[2] J. Benedetto and M. Frazier, editors. *Wavelets: Mathematics and Applications.* CRC Press, 2000 Corporate Blvd., N.W. Boca Raton, Florida 33431, 1993.

[3] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Comm. Pure Appl. Math.*, 44:141–183, 1991.

[4] C.K. Chui. *An Introduction to Wavelets.* Academic Press, 1992.

[5] C.K. Chui, editor. *Wavelets: A Tutorial in Theory and Applications.* Academic Press, 1992.

[6] J.M. Combes, A. Grossmann, and Ph. Tchamitchian, editors. *Wavelets: Time-Frequency Methods and Phase Space.* Inverse problems and theoretical imaging. Springer-Verlag, 1989. Proc. Internat. Conf. Marseille, France, Dec. 14-18, 1987.

[7] I. Daubechies. Discrete sets of coherent states and their use in signal analysis. In *Differential Equations and Mathematical Physiscs*, volume 1285 of *Lecture Notes in Math.*, pages 73–82. Springer-Verlag, 1987.

[8] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 41:909–996, 1988.

[9] I. Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conf. Ser. in Appl. Math.* SIAM, 1992.

[10] I. Daubechies, A. Grossmann, and Y. Meyer. Painless nonorthogonal expansions. *J. Math. Phys.*, 27(5):1271–1283, 1986.

[11] R.A. DeVore and B.J. Lucier. Wavelets. *Acta Numerica*, 1:1–56, 1992.

[12] M. Frazier and B. Jawerth. Decomposition of Besov spaces. *Indiana University Mathematics Journal*, 34(4):777–799, 1985.

[13] A. Grossman and J. Morlet. Decompostion of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 15(4):723–736, 1984.

[14] A. Grossmann, R. Kronland-Martinet, and J. Morlet. Reading and understanding continuous wavelet transforms. In J.M. Combes, A. Grossmann, and Ph. Tchamitchian, editors, *Wavelets: Time-Frequency Methods and Phase Space*, Inverse problems and theoretical imaging, pages 2–20. Springer-Verlag, 1989. Proc. Internat. Conf. Marseille, France, Dec. 14-18, 1987.

[15] A. Grossmann, J. Morlet, and T. Paul. Transforms associated to square integrable group representations I. General results. *J. Math. Phys.*, 26(10):2473–2479, 1985.

[16] H. Heijmans. Discrete wavelets and multiresolution analysis. In T.H. Koornwinder, editor, *Wavelets: an elementary treatment of theory and applications*, volume 1 of *Series in Approximations & Decompositions*. World Scientific, 1993.

[17] IEEE Transactions on Information Theory, Special issue on wavelet transform and multiresolution signal analysis, January 1992.

[18] B. Jawerth and W. Sweldens. An overview of wavelet based multiresolution analyses. *SIAM Rev.*, 36(3):377–412, 1994.

[19] T.H. Koornwinder, editor. *Wavelets: an elementary treatment of theory and applications*, volume 1 of *Series in Approximations & Decompositions*. World Scientific, 1993.

[20] P.-G. Lemarié, editor. *Les Ondelettes en 1989*, volume 1438 of *Lecture Notes in Math.* Springer-Verlag, 1990.

[21] W. Light, editor. *Advances in Numerical analysis, II: wavelets, subdivision algorithms and radial basis functions.* Claredon Press, 1991.

[22] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Patt. Anal. Machine Intell.*, 11(7):674–693, 1989.

[23] Y. Meyer. *Ondelettes et algorithmes concurrents.* Hermann, Paris, 1992.

[24] Y. Meyer, editor. *Wavelets and applications*, volume 20 of *Research Notes in Applied Mathematics.* Springer-Verlag, 1992. Proceedings Internat. Conf. Marseille, May, 1989.

[25] Y. Meyer. *Wavelets: Algorithms and Applications.* SIAM, Philadelphia, 1993. translated and revised by R.D. Ryan.

[26] W.H. Press. Wavelet transforms: a primer. Technical Report 3184, Center for Astrophysics, Department of Physics, Harvard University, 1991.

[27] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Processing Magazine*, pages 14–38, October 1991.

[28] M.B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S.G. Mallat, Y. Meyer, and L. Raphael, editors. *Wavelets and their Applications*. Jones and Bartlett, 1992.

[29] G. Strang. Wavelets and dilation equations: A brief introduction. *SIAM Rev.*, 31(4):614–627, 1989.