Proceedings of Symposia in Applied Mathematics
Volume 47, 1993

Wavelet Transforms and Orthonormal Wavelet Bases

INGRID DAUBECHIES

ABSTRACT. We introduce the wavelet transform and discuss its motivation as a time-frequency localization tool. We review the different types of wavelet transform, with a special emphasis on orthonormal wavelet bases and their properties. We finish by a short discussion of their shortcomings.

"Wavelets" or "wavelet transforms" are a tool for decomposing functions in various applications, several of which are presented in this short course. The functions to be analyzed can be solutions of a differential equation with shocks, or integral kernels of singular integral operators, or 1 or 2-dimensional signals, as in sound (speech or music), time series or images. The wavelet transform can be viewed as a synthesis over the last fifteen years of ideas from many different fields, ranging from pure mathematics to quantum physics and electrical engineering. I will give here a description of several types of wavelet transform, with a special emphasis on (orthonormal) wavelet bases.

1. Time-frequency localization: what and why?

Let $f(t)$ be a function depending on time. If we are interested in its "frequency content" or "spectrum", our first reflex is to compute its Fourier transform,

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int f(x)\, e^{-i\xi x}\, dx .$$

Just as the different harmonic components were present in $f(t)$, but impossible to read off at a glance, so the time information is present in $\hat{f}(\xi)$ but hard to

1991 *Mathematics Subject Classification.* Primary 33E20; Secondary 46E15, 41A15, 94A12.
These notes were written while the author was in the Mathematics Department at Rutgers University, partially supported by NSF grant 4-20875.
Parts of this paper are reprinted with permission from *Ten Lectures on Wavelets.*

read off (it is all hidden in the phase of $\hat{f}(\xi)$). Often we would like to have a frequency decomposition of f locally in time, similar to music notation, which tells the musician which note (= frequency information) to play when (= time information). This is what is achieved by so-called time frequency representations. The wavelet transform of f can be viewed as such a time frequency representation. There exist other, older and very useful time frequency representations. The most widely used is the windowed Fourier transform. Here the function f is first "windowed" by multiplying it by a fixed $g(t)$ (the window function); this effectively restricts f to an interval (with smoothed edges) (see Figure 1). Then

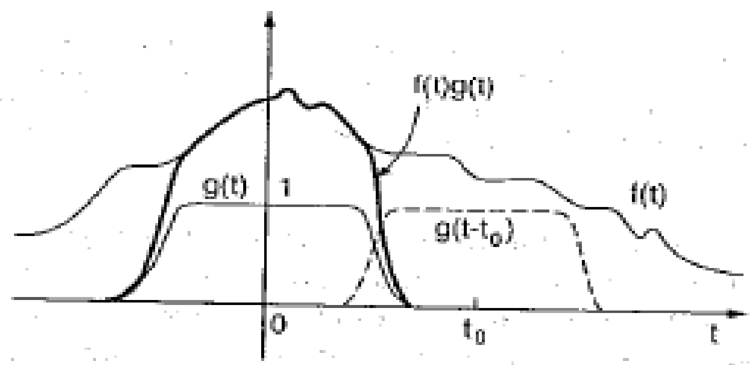

FIGURE 1.

the Fourier coefficients of this product are computed. This process is repeated with shifted versions of g, i.e. $g(t-nt_0)$, $n \in \mathbb{Z}$, leading to a family of windowed Fourier coefficients,

$$S_{m,n}(f) = \int f(s)\, g(s-nt_0)\, e^{im\omega_0 s}\, ds \,, \tag{1.1}$$

with $m, n \in \mathbb{Z}$. These can also be viewed as the inner products (in $L^2(\mathbb{R})$) of f with the

$$g_{mn}(t) = e^{-im\omega_0 s} g(t-nt_0) \tag{1.2}$$

(we assume g is real). Each g_{mn} consists of an envelope function, shifted by nt_0, and then "filled in" with oscillations (see Figure 2); the index n gives us the time localization of g_{mn}, the index m its frequency.

The wavelet transform is similar to the windowed Fourier transform in that it also computes inner products of f with a sequence of functions $\psi_{m,n}$, with m indicating frequency localization, and n time localization,

$$W_{m,n}(f) = \int f(s)\, \overline{\psi_{m,n}(t)}\, dt \,, \tag{1.3}$$

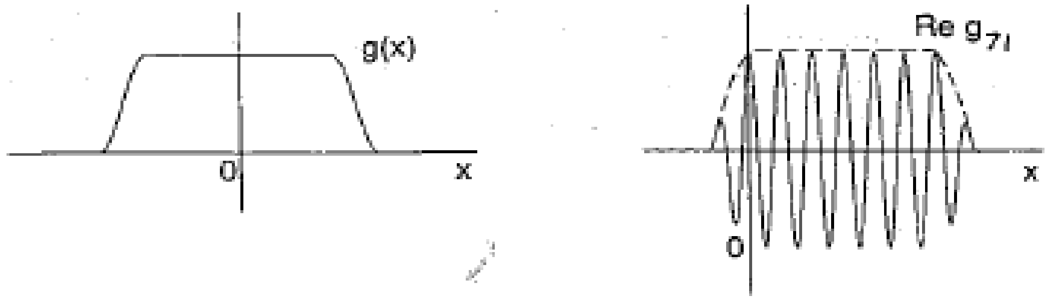

FIGURE 2.

but the $\psi_{m,n}$ are generated in different way. The basic wavelet ψ is typically well concentrated in time and in frequency, and has integral zero

$$\int \psi(t)\, dt = 0 , \tag{1.4}$$

which means it has at least some oscillations. The $\psi_{m,n}$ are then generated by dilations and translations:

$$\psi_{m,n}(t) = a_0^{-m/2} \psi(a_0^{-m} t - n b_0) , \tag{1.5}$$

where $a_0 > 1$ and $b_0 > 0$ are fixed parameters (similar to the ω_0, t_0 in (1.1), and m, n range over all of $\mathbb{Z}$). Changing m in (1.5) amounts to packing the oscillations of ψ into a smaller ($m > 0$) or larger ($m < 0$) width, i.e. to wavelets with higher or lower frequency ranges; for fixed m, the $\psi_{m,n}$ are then translates of $\psi_{m,0}$ by $n a_0^m b_0$, i.e. the wavelets are translated by amounts proportional to their width. A few typical wavelets are illustrated in Fig. 3. It is clear that high frequency

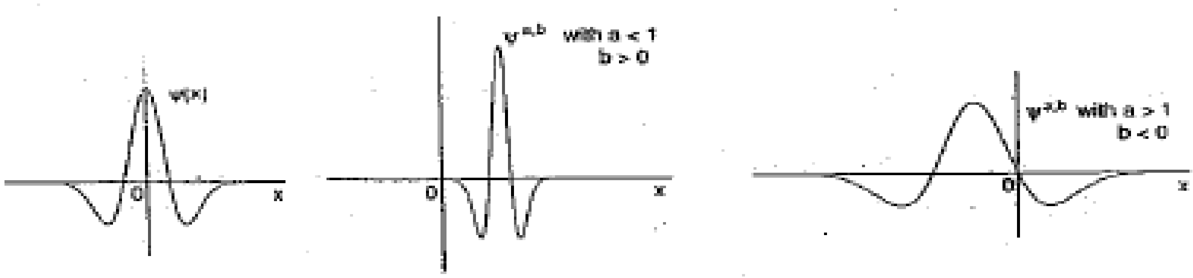

FIGURE 3.

wavelets are narrow, low frequency wavelets wide. This is the main difference between the wavelet transform and the short term Fourier transform: the $g_{m,n}$ of the short time Fourier transform all have the same width. It is therefore to be expected (and borne out in reality) that the wavelet transform is particularly well adapted to functions, signals or operators with highly concentrated high frequency components superposed on longer lived low frequency components.

2. Different types of wavelet transform.

2.1. The continuous wavelet transform. Here the dilation and translation parameters a, b vary continuously over $\mathbb{R}$. That is, we define (in 1 dimension; higher dimensional versions are straightforward)

$$\psi^{a,b}(x) = a^{-1/2}\psi\left(\frac{x-b}{a}\right) , \tag{2.1}$$

with $a, b \in \mathbb{R}$, $a > 0$. Then

$$(Wf)(a,b) = \langle f, \psi^{a,b}\rangle = \int f(x)\, a^{-1/2}\, \overline{\psi\left(\frac{x-b}{a}\right)} = \int \hat{f}(\xi)\, a^{1/2}\, \overline{\hat{\psi}(a\xi)}\, e^{ib\xi} , \tag{2.2}$$

and

$$\begin{aligned} &\int_0^\infty \int_{-\infty}^\infty \langle f, \psi^{a,b}\rangle\langle \psi^{a,b}, g\rangle \, db\, \frac{da}{a^2} \\ &= 2\pi \int_0^\infty \int_{-\infty}^\infty \hat{f}(\xi)\, \overline{\hat{g}(\xi)}\, |\hat{\psi}(a\xi)|^2 \, d\xi\, \frac{da}{a} \\ &= 2\pi C_\psi \, \langle f, g\rangle , \end{aligned} \tag{2.3}$$

provided that

$$\int_0^\infty \xi^{-1}|\hat{\psi}(\xi)|^2 \, d\xi = \int_{-\infty}^0 |\xi|^{-1}\, |\hat{\psi}(\xi)|^2 \, d\xi \; =: \; C_\psi < \infty . \tag{2.4}$$

Condition (2.4) implies that $\int_{-\infty}^\infty |\xi|^{-1}|\hat{\psi}(\xi)|^2 \, d\xi < \infty$, which (for reasonable ψ) amounts to the same as our earlier requirement $\int \psi(x)\, dx = 0$. Another ingredient in (2.4) is a symmetry of concentration in $|\hat{\psi}(\xi)|^2$, with respect to the measure $|\xi|^{-1}\, d\xi$, on positive and negative frequency axes. This requirement is automatically satisfied if ψ is real. On the other hand, if (2.3) is only required for real f, g, where, since $\hat{f}(-\xi) = [\hat{f}(\xi)]^*$, the positive frequency behavior completely determines the negative frequency analog, then one can find formulations in which the symmetry for $\xi \leftrightarrow -\xi$ in (2.4) is no longer necessary. Similarly, if one allows negative a in (2.1) and (2.2) then (2.4) collapses to $C_\psi := \int_{-\infty}^\infty |\xi|^{-1}\, |\hat{\psi}(\xi)|^2 < \infty$ (see Daubechies (1992)).

Formula (2.3) can also be rewritten as

$$f(x) = \frac{1}{2\pi C_\psi} \int_{-\infty}^\infty \int_0^\infty \langle f, \psi^{a,b}\rangle \psi^{a,b}(x)\, \frac{da\, db}{a^2} , \tag{2.5}$$

with weak convergence in L^2-sense. In fact, for reasonable ψ, (2.5) converges in many more topologies; in particular, it converges pointwise in any point x where f is continuous (see Holschneider and Tchamitchian (1990)).

Note that (2.5) can be read in two different ways: it tells us, once we know the $\langle f, \psi^{a,b} \rangle$, how to reconstruct f from these wavelet coefficients; it also gives a recipe for writing any arbitrary f as a superposition of $\psi^{a,b}$.

Formula (2.5) has in fact been known for quite a while: it is already implicit in Calderón (1964) as a useful mathematical tool (with completely different notations), and it appeared as the "reproducing identity for the $ax+b$-group" in Aslaksen and Klauder (1968). A similar and even older reproducing identity exists for the continuous windowed Fourier transform. (For an extensive discussion of these and other reproducing identities, see Klauder and Skagerstam (1985).)

It may seem puzzling that, according to (2.5), we can write any f, even if $\int f(x)dx > 0$, as a superposition of $\psi^{a,b}$, each of which has zero integral. The solution to this paradox is that (2.5) converges in L^2, or pointwise, but not in L^1. In fact, for any finite a_1, R, and any nonzero a_0, the functions

$$f_{a_0,a_1;R}(x) = \frac{1}{2\pi C_\psi} \int_{-R}^{R} \int_{a_0}^{a_1} \langle f, \psi^{a,b} \rangle \psi^{a,b}(x) \, \frac{da\, db}{a^2}$$

will have zero integral; for a_0 close to 0 and a_1, R very large, their graph will be very close to that of f, except that they will have large, shallow, negative "pools" in regions where f is small, leading to small pointwise or L^2 differences, but sufficient to ensure $\int f_{a_0,a_1;R}(x)\, dx = 0$. (See Figure 4.)

The continuous wavelet transform is useful when one wants to recognize or extract features. Scaling or translating f leads to a shift of the $(Wf)(a,b)$ in a and b, so that the whole analysis can be made to be scale and translation invariant, a desirable property in some applications. Of course, it can be cumbersome to have to deal with the very redundant $(Wf)(a,b)$: after all, we have changed a 1-dimensional function f into the 2-dimensional Wf; pictures of Wf may give insight into the different components of f, but this is only a first stage. Several groups, mostly in Marseille (France) have developed mathematical tools for extracting the "bare bones" from $Wf(a,b)$ and use these to describe f; an extensive review article is Delprat et al (1992).

2.2. The discrete but redundant wavelet transform: frames. The wavelet family (1.5) and the wavelet transform (1.3) can be viewed as discretized versions of the continuous wavelet transform, with a, b restricted to $a = a_0^m$; $b = nb_0a_0^m$.

In the discrete case, there does not exist, in general, a "resolution of the identity" formula analogous to (2.5) for the continuous case. Reconstruction of f from the $W_{m,n}(f)$, if at all possible, must therefore be done by some other means. The following questions naturally arise:

(1) Is it possible to characterize f completely by knowing the $W_{m,n}(f)$?

(2) Is it possible to reconstruct f in a numerically stable way from the $W_{m,n}(f)$?

These questions concern the recovery of f from its wavelet transform. We can also consider the dual problem, the possibility of expanding f into wavelets,

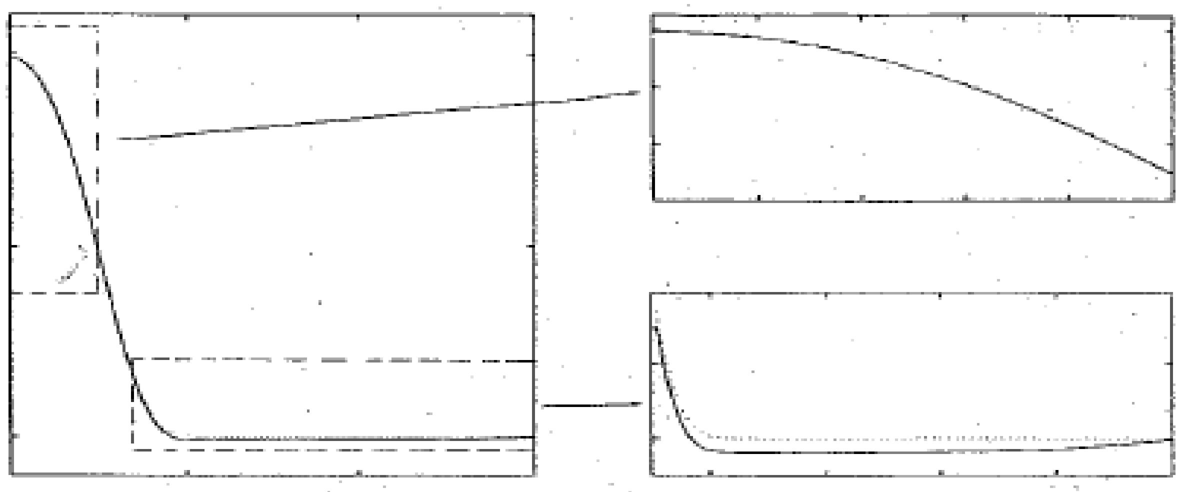

FIGURE 4. A Gaussian (dotted line) and its reconstruction (solid line) with cutoffs in a and b (see text). Only the right half is plotted; to make the effect more visible, two blowups are provided as well. The reconstruction has wide shallow "pools" on the side so that its integral is zero, even though it is close to the Gaussian at every point. Note that the horizontal scale is sinh x rather than x, giving a linear scale near 0 but an exponential scale further on.

which then leads to the dual questions:

(1′) Can any function be written as a superposition of $\psi_{m,n}$?

(2′) Is there a numerically stable algorithm to compute the coefficients for such an expansion?

As in the continuous case, these discrete wavelet transforms often provide a very redundant description of the original function. This redundancy can be exploited (it is, for instance, possible to compute the wavelet transform only approximately, while still obtaining reconstruction of f with good precision), or eliminated to reduce the transform to its bare essentials (such as in the image compression work of S. Mallat and S. Zhong (1992)). It is in this discrete form that the wavelet transform is closest to the "ϕ-transform" of Frazier and Jawerth (1988).

The choice of the wavelet ψ used in the continuous wavelet transform or in frames of discretely labelled families of wavelets is essentially only restricted by the requirement that C_ψ, as defined by (2.4), is finite. For practical reasons, one usually chooses ψ so that it is well concentrated in both the time and the frequency domain. For any such ψ, one can then find threshold values such that if a_0, b_0 are chosen below these thresholds, then all the questions above can be answered by "yes", and one can construct explicit algorithms. (For a much more extensive discussion, see Daubechies (1992).) All this still leaves a lot of freedom. Giving up a lot of this freedom allows one to build (orthonormal) bases of wavelets.

2.3. Orthonormal wavelet bases: the Haar basis as an example. For some very special choices of ψ and a_0, b_0, the $\psi_{m,n}$ constitute an orthonormal basis for $L^2(\mathbb{R})$. In particular, if we choose $a_0 = 2$, $b_0 = 1$, then there exist ψ, with good time-frequency localization properties, such that the

$$\psi_{m,n}(x) = 2^{-m/2}\,\psi(2^{-m}x - n) \tag{2.6}$$

constitute an orthonormal basis for $L^2(\mathbb{R})$. (Other choices for a_0 are possible, but we shall restrict ourselves to $a_0 = 2$ here.) The oldest example of a function ψ for which the $\psi_{m,n}$ defined by (2.6) constitute an orthonormal basis for $L^2(\mathbb{R})$ is the Haar function,

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

The Haar basis has been known since Haar (1910). Note that the Haar function does not have good time-frequency localization: its Fourier transform $\hat{\psi}(\xi)$ decays like $|\xi|^{-1}$ for $\xi \to \infty$. Nevertheless we shall use it here for illustration purposes. What follows is a proof that the Haar family does indeed constitute an orthonormal basis. This proof is different from the one in most textbooks; in fact it will use multiresolution analysis as a tool.

In order to prove that the $\psi_{m,n}(x)$ constitute an orthonormal basis, we need to establish that

(1) the $\psi_{m,n}$ are orthonormal

(2) any L^2-function f can be approximated, up to arbitrarily small precision, by a finite linear combination of the $\psi_{m,n}$.

Orthonormality is easy to establish. Since support$(\psi_{m,n}) = [2^m n, 2^m(n+1)]$, it follows that two Haar wavelets of the same scale (same value of m) never overlap, so that $\langle \psi_{m,n}, \psi_{m,n'} \rangle = \delta_{n,n'}$. Overlapping supports are possible if the two wavelets have different sizes, as in Figure 5. It is easy to check, however, that if $m < m'$, then support$(\psi_{m,n})$ lies wholly within a region where $\psi_{m',n'}$ is constant (as on the figure). It follows that the inner product of $\psi_{m,n}$ and $\psi_{m',n'}$ is then proportional to the integral of ψ itself, which is zero.

We concentrate now on how well an arbitrary function f can be approximated by linear combinations of Haar wavelets. Any f in $L^2(\mathbb{R})$ can be arbitrarily well approximated by a function with compact support which is piecewise constant on the $[\ell 2^{-j}, (\ell+1)2^{-j}[$ (it suffices to take the support and j large enough). We can therefore restrict ourselves to such piecewise constant functions only:

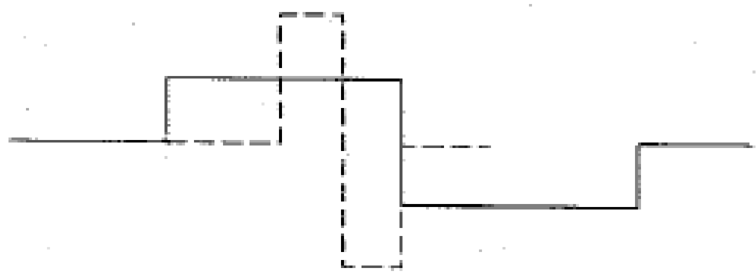

FIGURE 5. Two Haar wavelets; the support of the "narrower" wavelet is completely contained in an interval where the "wider" wavelet is constant.

assume f to be supported on $[-2^{J_1}, 2^{J_1}]$, and to be piecewise constant on the $[\ell 2^{-J_0}, (\ell+1)2^{-J_0}[$, where J_1 and J_0 can both be arbitrarily large (see Figure 6). Let us denote the constant value of $f^0 = f$ on $[\ell 2^{-J_0}, (\ell+1)2^{-J_0}[$ by f^0_ℓ. We now represent f^0 as a sum of two pieces, $f^0 = f^1 + \delta^1$, where f^1 is an approximation to f^0 which is piecewise constant over intervals twice as large as originally, i.e. $f^1|_{[k2^{-J_0+1},(k+1)2^{-J_0+1}[} \equiv constant = f^1_k$. The values f^1_k are given by the averages of the two corresponding constant values for f^0, $f^1_k = \frac{1}{2}(f^0_{2k} + f^0_{2k+1})$ (see Figure 6). The function δ^1 is piecewise constant with the same stepwidth as f^0; one immediately has

$$\delta^1_{2\ell} = f^0_{2\ell} - f^1_\ell = \frac{1}{2}(f^0_{2\ell} - f^0_{2\ell+1})$$

and

$$\delta^1_{2\ell+1} = f^0_{2\ell+1} - f^1_\ell = \frac{1}{2}(f^0_{2\ell+1} - f^0_{2\ell}) = -\delta^1_{2\ell} .$$

It follows that δ^1 is a linear combination of scaled and translated Haar functions:

$$\delta^1 = \sum_{\ell=-2^{J_1+J_0-1}+1}^{2^{J_1+J_0-1}} \delta^1_{2\ell}\psi(2^{J_0-1}x - \ell) .$$

We have therefore written f as

$$f = f^0 = f^1 + \sum_\ell c_{-J_0+1,\ell}\, \psi_{-J_0+1,\ell} ,$$

where f^1 is of the same type as f^0, but with stepwidth twice as large. We can apply the same trick to f^1, so that

$$f^1 = f^2 + \sum_\ell c_{-J_0+2,\ell}\, \psi_{-J_0+2,\ell} ,$$

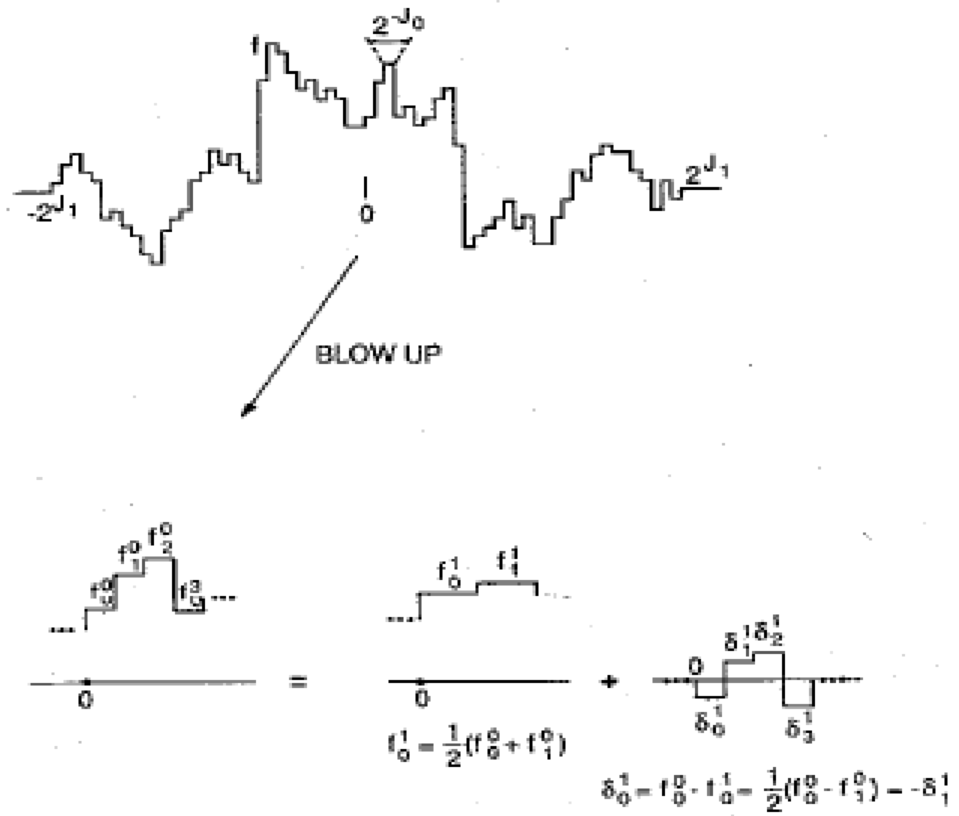

FIGURE 6. (a) A function f with support $[-2^{J_1}, 2^{J_1}]$, piecewise constant on the $[k2^{-J_0}, (k+1)2^{-J_0}[$. (b) A blowup of a portion of f. On every pair of intervals, f is replaced by its average ($\longrightarrow f^1$); the difference between f and f^1 is δ^1, a linear combination of Haar wavelets.

with f^2 still supported on $[-2^{J_1}, 2^{J_1}]$, but piecewise constant on the even larger intervals $[k2^{-J_0+2}, (k+1)2^{-J_0+2}[$. We can keep going like this, until we have

$$f = f^{J_0+J_1} + \sum_{m=-J_0+1}^{J_1} \sum_{\ell} c_{m,\ell}\, \psi_{m,\ell}\,.$$

Here $f^{J_0+J_1}$ consists of two constant pieces (see Figure 7), with $f^{J_0+J_1}|_{[0,2^{J_1}[} \equiv f_0^{J_0+J_1}$ equal to the average of f over $[0, 2^{J_1}[$, and $f^{J_0+J_1}|_{[-2^{J_1},0[} \equiv f_{-1}^{J_0+J_1}$ the average of f over $[-2^{J_1}, 0[$.

Even though we have "filled out" the whole support of f, we can still keep going with our averaging trick: nothing stops us from widening our horizon from 2^{J_1} to 2^{J_1+1}, and writing $f^{J_1+J_2} = f^{J_1+J_2+1} + \delta^{J_1+J_2+1}$, where

$$f^{J_1+J_2+1}|_{[0,2^{J_1+1}[} \equiv \frac{1}{2} f_0^{J_1+J_2}, \quad f^{J_1+J_2+1}|_{[-2^{J_1+1},0[} \equiv \frac{1}{2} f_{-1}^{J_1+J_2}$$

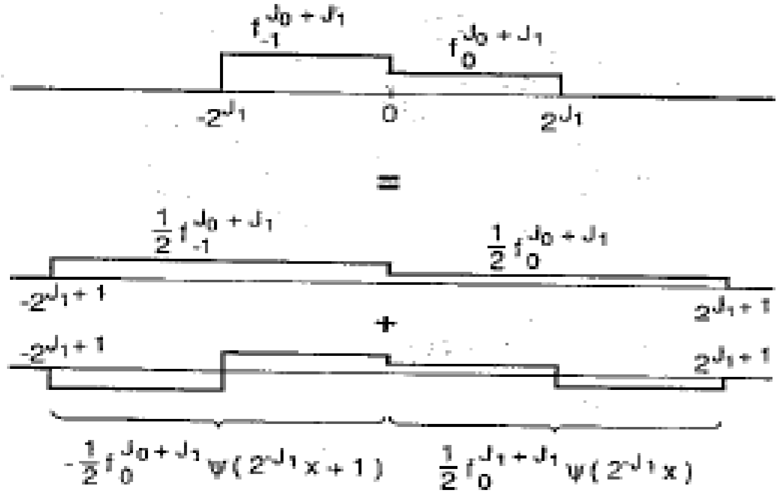

FIGURE 7. The averages of f on $[0,\ 2^{J_1}]$ and $[-2^{J_1},\ 0]$ can be "smeared" out over the bigger intervals $[0,\ 2^{J_1+1}]$, $[-2^{J_1+1},\ 0]$; the difference is a linear combination of very stretched out Haar functions.

and

$$\delta^{J_1+J_2} = \frac{1}{2} f_0^{J_1+J_2} \psi(2^{-J_1-1}x) - \frac{1}{2} f_{-1}^{J_1+J_2} \psi(2^{-J_1-1}x+1)$$

(see Figure 7). This can again be repeated, leading to

$$f = f^{J_0+J_1+K} + \sum_{m=-J_0+1}^{J_1+K} \sum_{\ell} c_{m,\ell}\, \psi_{m,\ell}\ ,$$

where $\text{support}(f^{J_0+J_1+K}) = [-2^{J_1+K}, 2^{J_1+K}]$, and

$$f^{J_0+J_1+K}|_{[0,2^{J_1+K}]} = 2^{-K} f_0^{J_0+J_1}\ ,\ f^{J_0+J_1+K}|_{[-2^{J_1+K},0]} = 2^{-K} f_{-1}^{J_0+J_1}\ .$$

It follows immediately that

$$\left\| f - \sum_{m=-J_0+1}^{J_1+K} \sum_{\ell} c_{m,\ell}\, \psi_{m,\ell} \right\|_{L^2}^2 = \| f^{J_0+J_1+K} \|_{L^2}^2$$

$$= 2^{-K/2} \cdot 2^{J_1/2} \left[|f_0^{J_0+J_1}|^2 + |f_{-1}^{J_0+J_1}|^2 \right]^{1/2} ,$$

which can be made arbitrarily small by taking sufficiently large K. As claimed, f can therefore be approximated to arbitrary precision by a finite linear combination of Haar wavelets!

The argument we just saw has implicitly used a "multiresolution" approach: we have written successive coarser and coarser approximations to f (the f^j, averaging f over larger and larger intervals), and at every step we have written the difference between the approximation with resolution 2^{j-1}, and the next coarser level, with resolution 2^j, as a linear combination of the $\psi_{j,k}$.

The Haar basis is a "good" basis for $L^p(\mathbb{R})$, $1 < p < \infty$ (i.e. it is an unconditional basis; see §7). It is however not a suitable basis for smoother function spaces, such as the Sobolev spaces. In the next section, we shall see how the multiresolution approach can be made to work for other, smoother wavelet bases, which then are unconditional bases for a much wider range of functional spaces.

3. Multiresolution analysis.

A multiresolution analysis consists of a sequence of successive approximation spaces V_j. More precisely, the closed subspaces V_j satisfy

$$\ldots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots \tag{3.1}$$

with

$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} \;=\; L^2(\mathbb{R})\,, \tag{3.2}$$

$$\bigcap_{j\in\mathbb{Z}} V_j \;=\; \{0\}\,. \tag{3.3}$$

If we denote by P_j the orthogonal projection operator onto V_j, then (3.2) ensures that $\lim_{j\to-\infty} P_j f = f$ for all $f \in L^2(\mathbb{R})$. There exist many ladders of spaces satisfying (3.1)–(3.3) which have nothing to do with "multiresolution"; the multiresolution aspect is a consequence of the additional requirement

$$f \in V_j \iff f(2^j\cdot) \in V_0\,. \tag{3.4}$$

That is, all the spaces are scaled versions of the central space V_0. An example of spaces V_j satisfying (3.1)–(3.4) is

$$V_j = \{f \in L^2(\mathbb{R});\ \forall k \in \mathbb{Z}:\ f|_{[2^j k,\, 2^j(k+1)[} = \text{constant}\}\,.$$

We shall call this example the Haar multiresolution analysis. It corresponds with our argument in §2.3; see also below. Figure 8 shows what the projection of some f on the Haar spaces V_0, V_{-1} might look like. This example also exhibits another feature that we require from a multiresolution analysis: invariance of V_0 under integer translations,

$$f \in V_0 \Rightarrow f(\cdot - n) \in V_0, \text{ for all } n \in \mathbb{Z}\,. \tag{3.5}$$

Because of (3.4) this implies that if $f \in V_j$, then $f(\cdot - 2^j n) \in V_j$ for all $n \in \mathbb{Z}$. Finally, we require also that there exists $\phi \in V_0$ so that

$$\{\phi_{0,n};\ n \in \mathbb{Z}\} \text{ is an orthonormal basis in } V_0 \tag{3.6}$$

where, for all $j, n \in \mathbb{Z}$, $\phi_{j,n}(x) = 2^{-j/2}\,\phi(2^{-j}x - n)$. Together, (3.6) and (3.4) imply that $\{\phi_{j,n};\ n \in \mathbb{Z}\}$ is an orthonormal basis for V_j, for all $j \in \mathbb{Z}$. This last

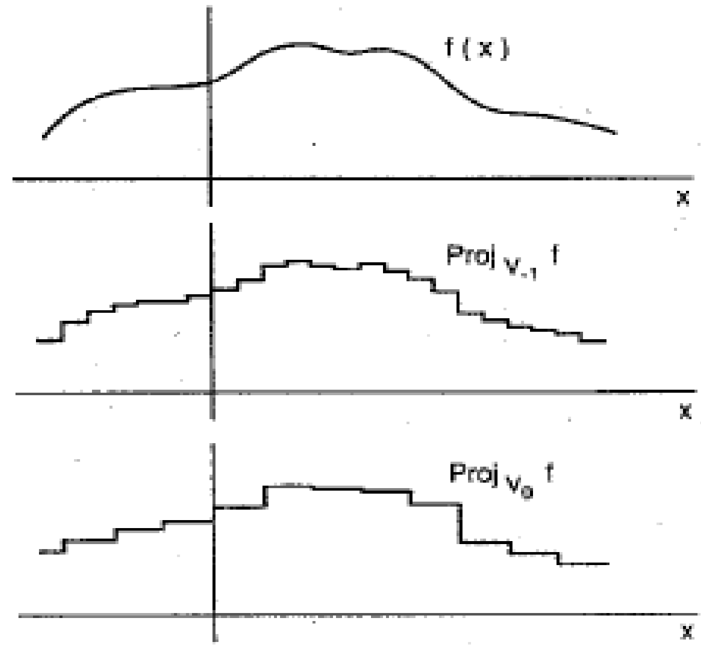

FIGURE 8. A function f and its projections onto V_{-1} and V_0.

requirement (3.6) seems a bit more "contrived" than the other ones; we shall see below that it can be relaxed considerably. In the Haar example, a possible choice for ϕ is the indicator function for $[0,1]$, $\phi(x) = 1$ if $0 \leq x \leq 1$, $\phi(x) = 0$ otherwise. We shall often call ϕ the "scaling function" of the multiresolution analysis.

The basic tenet of multiresolution analysis is that whenever a collection of closed subspaces satisfies (3.1)–(3.6), then there exists an orthonormal wavelet basis $\{\psi_{j,k};\ j,k \in \mathbb{Z}\}$ of $L^2(\mathbb{R})$, $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$, such that, for all f in $L^2(\mathbb{R})$,

$$P_{j-1}f = P_j f + \sum_{k\in\mathbb{Z}} \langle f,\ \psi_{j,k}\rangle\ \psi_{j,k}\,. \tag{3.7}$$

(P_j is the orthogonal projection onto V_j.) The wavelet ψ can moreover be constructed explicitly. Let us see how.

For every $j \in \mathbb{Z}$, define W_j to be the orthogonal complement of V_j in V_{j-1}. We have

$$V_{j-1} = V_j \oplus W_j\ , \tag{3.8}$$

and

$$W_j \perp W_{j'} \text{ if } j \neq j'\ . \tag{3.9}$$

(If $j > j'$, e.g., then $W_j \subset V_{j'} \perp W_{j'}$.) It follows that, for $j < J$

$$V_j = V_J \oplus \bigoplus_{k=0}^{J-j-1} W_{J-k} , \tag{3.10}$$

where all these subspaces are orthogonal. By virtue of (3.2) and (3.3) this implies

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j , \tag{3.11}$$

a decomposition of $L^2(\mathbb{R})$ into mutually orthogonal subspaces. Furthermore, the W_j spaces inherit the scaling property (3.4) from the V_j.

$$f \in W_j \Longleftrightarrow f(2^j\cdot) \in W_0 . \tag{3.12}$$

Formula (3.7) is equivalent to saying that, for fixed j, $\{\psi_{j,k};\ k \in \mathbb{Z}\}$ constitutes an orthonormal basis for W_j. Because of (3.11) and (3.2), (3.3) this then automatically implies that the whole collection $\{\psi_{j,k};\ j,k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$. On the other hand, (3.12) ensures that if $\{\psi_{0,k};\ k \in \mathbb{Z}\}$ is an orthonormal basis for W_0, then $\{\psi_{j,k};\ k \in \mathbb{Z}\}$ will likewise be an orthonormal basis for W_j, for any $j \in \mathbb{Z}$. Our task thus reduces to finding $\psi \in W_0$ such that the $\psi(\cdot - k)$ constitute an orthonormal basis for W_0.

To construct this ψ, let us write out some interesting properties of ϕ and W_0.

1. Since $\phi \in V_0 \subset V_{-1}$, and the $\phi_{-1,n}$ are an orthonormal basis in V_{-1}, we have

$$\phi = \sum_n h_n\, \phi_{-1,n} , \tag{3.13}$$

with

$$h_n = \langle \phi,\ \phi_{-1,n}\rangle, \text{ and } \sum_{n\in\mathbb{Z}} |h_n|^2 = 1 . \tag{3.14}$$

We can rewrite (3.13) as either

$$\phi(x) = \sqrt{2} \sum_n h_n\, \phi(2x - n) \tag{3.15}$$

or

$$\hat\phi(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n\, e^{-in\xi/2}\, \hat\phi(\xi/2) , \tag{3.16}$$

where convergence in either sum holds in L^2-sense. Formula (3.16) can be rewritten as

$$\hat\phi(\xi) = m_0(\xi/2)\, \hat\phi(\xi/2) , \tag{3.17}$$

where

$$m_0(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n \, e^{-in\xi} \, . \tag{3.18}$$

Equality in (3.17) holds pointwise almost everywhere. As (3.18) shows, m_0 is a 2π-periodic function in $L^2([0, 2\pi])$.

2. The orthonormality of the $\phi(\cdot - k)$ leads to special properties for m_0. We have

$$\begin{aligned} \delta_{k,0} &= \int dx \; \phi(x) \; \overline{\phi(x-k)} = \int d\xi \; |\hat{\phi}(\xi)|^2 \; e^{ik\xi} \\ &= \int_0^{2\pi} d\xi \; e^{ik\xi} \sum_{\ell \in \mathbb{Z}} |\hat{\phi}(\xi + 2\pi\ell)|^2 \, , \end{aligned}$$

implying

$$\sum_\ell \; |\hat{\phi}(\xi + 2\pi\ell)|^2 = (2\pi)^{-1} \qquad \text{a.e.} \tag{3.19}$$

Substituting (3.17) leads to ($\zeta = \xi/2$)

$$\sum_\ell \; |m_0(\zeta + \pi\ell)|^2 \; |\hat{\phi}(\zeta + \pi\ell)|^2 = (2\pi)^{-1} \; ;$$

splitting the sum into even and odd ℓ, using the periodicity of m_0 and applying (3.19) once more gives

$$|m_0(\zeta)|^2 \; + \; |m_0(\zeta + \pi)|^2 = 1 \qquad \text{a.e.} \tag{3.20}$$

3. Let us now characterize W_0: $f \in W_0$ is equivalent to $f \in V_{-1}$ and $f \perp V_0$. Since $f \in V_{-1}$, we have

$$f = \sum_n \; f_n \; \phi_{-1,n} \; ,$$

with $f_n = \langle f, \; \phi_{-1,n} \rangle$. This implies

$$\hat{f}(\xi) = \frac{1}{\sqrt{2}} \sum_n \; f_n \; e^{-in\xi/2} \; \hat{\phi}(\xi/2) = m_f(\xi/2) \; \hat{\phi}(\xi/2) \; , \tag{3.21}$$

where

$$m_f(\xi) = \frac{1}{\sqrt{2}} \sum_n \; f_n \; e^{-in\xi} \tag{3.22}$$

is a 2π-periodic function in $L^2([0, 2\pi])$; convergence in (3.22) holds pointwise a.e. The constraint $f \perp V_0$ implies $f \perp \phi_{0,k}$ for all k, i.e.

$$\int d\xi \; \hat{f}(\xi) \; \overline{\hat{\phi}(\xi)} \; e^{ik\xi} \; = \; 0$$

or

$$\int_0^{2\pi} d\xi\, e^{ik\xi} \sum_\ell \hat{f}(\xi + 2\pi\ell)\, \overline{\hat{\phi}(\xi + 2\pi\ell)} = 0 ,$$

hence

$$\sum_\ell \hat{f}(\xi + 2\pi\ell)\, \overline{\hat{\phi}(\xi + 2\pi\ell)} = 0 , \tag{3.23}$$

where the series in (3.23) converges absolutely in $L^1([-\pi,\pi])$. Substituting (3.17) and (3.21), regrouping the sums for odd and even ℓ (which we are allowed to do, because of the absolute convergence), and using (3.19) leads to

$$m_f(\zeta)\, \overline{m_0(\zeta)} + m_f(\zeta+\pi)\, \overline{m_0(\zeta+\pi)} = 0 \qquad \text{a.e.} \tag{3.24}$$

Since $\overline{m_0(\zeta)}$ and $\overline{m_0(\zeta+\pi)}$ cannot vanish together on a set of nonzero measure (because of (3.20)), this implies the existence of a 2π-periodic function $\lambda(\zeta)$ so that

$$m_f(\zeta) = \lambda(\zeta)\, \overline{m_0(\zeta+\pi)} \qquad \text{a.e.} \tag{3.25}$$

and

$$\lambda(\zeta) + \lambda(\zeta+\pi) = 0 . \qquad \text{a.e.} \tag{3.26}$$

This last equation can be recast as

$$\lambda(\zeta) = e^{i\zeta}\, \nu(2\zeta) , \tag{3.27}$$

where ν is 2π-periodic. Substituting (3.27) and (3.25) into (3.21) gives

$$\hat{f}(\xi) = e^{i\xi/2}\, \overline{m_0(\xi/2+\pi)}\, \nu(\xi)\, \hat{\phi}(\xi/2) , \tag{3.28}$$

where ν is 2π-periodic.

4. The general form (3.28) for the Fourier transform of $f \in W_0$ suggests that we take

$$\hat{\psi}(\xi) = e^{i\xi/2}\, \overline{m_0(\xi/2+\pi)}\, \hat{\phi}(\xi/2) \tag{3.29}$$

as a candidate for our wavelet. Disregarding convergence questions, (3.28) can indeed be written as

$$\hat{f}(\xi) = \left(\sum_k \nu_k\, e^{-ik\xi} \right) \hat{\psi}(\xi)$$

or

$$f = \sum_k \nu_k\, \psi(\cdot - k) ,$$

so that the $\psi(\cdot - n)$ are a good candidate for a basis of W_0. We need to verify that the $\psi_{0,k}$ are indeed an orthonormal basis for W_0. First of all, the properties of m_0 and $\hat{\phi}$ ensure that (3.29) defines indeed an L^2-function $\in V_{-1}$ and $\perp V_0$

(by the analysis above), so that $\psi \in W_0$. Orthonormality of the $\psi_{0,k}$ is easy to check:

$$\int dx\, \psi(x)\, \overline{\psi(x-k)} = \int d\xi\, e^{ik\xi}\, |\hat{\psi}(\xi)|^2 = \int_0^{2\pi} d\xi\, e^{ik\xi} \sum_{\ell} |\hat{\psi}(\xi + 2\pi\ell)|^2 .$$

Now

$$\begin{aligned}\sum_{\ell} |\hat{\psi}(\xi + 2\pi\ell)|^2 &= \sum_{\ell} |m_0(\xi/2 + \pi\ell + \pi)|^2\, |\hat{\phi}(\xi/2 + \pi\ell)|^2 \\ &= |m_0(\xi/2+\pi)|^2 \sum_n |\hat{\phi}(\xi/2 + 2\pi n)|^2 \\ &\quad + |m_0(\xi/2)|^2 \sum_n |\hat{\phi}(\xi/2 + \pi + 2\pi n)|^2 \\ &= (2\pi)^{-1} \left[|m_0(\xi/2)|^2 + |m_0(\xi/2+\pi)|^2\right] \qquad \text{a.e. (by (3.19))} \\ &= (2\pi)^{-1} \qquad \text{a.e. (by (3.20))} .\end{aligned}$$

Hence $\int dx\, \psi(x)\, \overline{\psi(x-k)} = \delta_{k0}$. In order to check that the $\psi_{0,k}$ are indeed a basis for all of W_0, it then suffices to check that any $f \in W_0$ can be written as

$$f = \sum_n \gamma_n\, \psi_{0,n} ,$$

with $\sum_n |\gamma_n|^2 < \infty$, or

$$\hat{f}(\xi) = \gamma(\xi)\, \hat{\psi}(\xi) , \tag{3.30}$$

with γ 2π-periodic and $\in L^2([0, 2\pi])$. But this is nothing but (3.28), where it is easy to check that ν is indeed square integrable. We have therefore proved the assertion at the start of this section: there is an orthonormal wavelet basis $\{\psi_{j,k};\ j,k \in \mathbb{Z}\}$ associated with any multiresolution analysis, and we even have a recipe for the construction of ψ:

$$\begin{aligned}\psi(x) &= \sum_n (-1)^n\, h_{-n+1}\, \phi_{-1,n} \\ &= \sqrt{2} \sum_n (-1)^n\, h_{-n+1}\, \phi(2x - n) ,\end{aligned} \tag{3.31}$$

where ϕ is the scaling function of the multiresolution analysis. (Note that (3.31) corresponds to (3.29), except for a change of sign, and a shift by 1 in x, neither of which affect the result.)

Not every orthonormal wavelet basis derives from a multiresolution analysis. There exist "pathological" counterexamples in which ψ has very bad decay. (See Mallat (1989) or Daubechies (1992)). Recently, it was proved in P. Auscher (1992) and in Lemarié-Rieusset (1992) that if ψ has a modicum of decay and smoothness, then it necessarily stems from a multiresolution analysis. Lemarié-Rieusset (1991) contains an earlier proof for compactly supported ψ. More details on these results can also be found in the chapter by P. G. Lemarié-Rieusset in this volume.

To conclude this section, let us see what the recipe (3.31) gives for the Haar multiresolution analysis. In that case $\phi(x) = 1$ for $0 \le x < 1$, 0 otherwise, hence

$$h_n = \sqrt{2} \int dx\, \phi(x)\, \overline{\phi(2x-n)} = \begin{cases} 1/\sqrt{2} & \text{if } n = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

Consequently $\psi = \frac{1}{\sqrt{2}}\, \phi_{-1,0} - \frac{1}{\sqrt{2}}\, \phi_{-1,1}$ or

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

and we recover indeed the Haar basis.

Of course, the real interest of this formalism lies in the other examples that can be built with it. The whole framework was developed by S. Mallat (1989) and Y. Meyer. The first construction of smooth wavelet bases (Stromberg (1982), which unfortunately went largely unnoticed at the time, Meyer (1985), Lemarié (1988) and Battle (1987)) did not use multiresolution analysis, and seemed much more ad hoc and miraculous. Interestingly enough, Mallat's background in vision analysis played a role in the development of multiresolution analysis (see Daubechies (1988) for a discussion of the connection): an interesting example of feedback from a very applied field to theory.

4. Spline wavelets.

Let us try the constructions in §3 for other multiresolution analysis ladders. One can choose e.g. a ladder of spline spaces, very popular in approximation theory.

$$V_j = \{f \text{ in } L^2(\mathbb{R});\quad f \in C^{\ell-1} \text{ and } f\,|_{[2^j k,\, 2^j(k+1)]} \text{ is a polynomial of order } \ell, \text{ for all } k \in \mathbb{Z}\}\ .$$

These are splines of order ℓ, with equispaced knots. The requirements (3.1)–(3.5) are obviously satisfied, but (3.6) is a bit more tricky. The usual B-spline function, i.e. the ℓ-th convolution of ϕ_{Haar} with itself, has the property that it and its integer translates generate all of V_0, but they are not orthonormal. For $\ell = 1$, for instance, we get the tent function $\phi(x) = 1 - |x|$ for $|x| \le 1$, $\phi(x) = 0$ otherwise, and obviously $\phi(x)$ is not orthogonal to $\phi(x-1)$. This can be fixed easily however; we can relax (3.6) and replace it by the requirement that the $\phi(x-n)$ constitute a Riesz basis for V_0, i.e. that they span V_0 and that for

$f = \sum_n c_n \phi_{0n} \in V_0$, the norms $\sum_n |c_n|^2$ and $\|f\|^2$ are equivalent, in the sense that

$$A|c_n|^2 \le \left\| \sum_n c_n \phi_{0n} \right\|^2 \le B \sum_n |c_n|^2 \,, \tag{4.1}$$

with $A > 0$, $B < \infty$ and independent of f.

Because

$$\begin{aligned} \left\| \sum_n c_n \phi_{0n} \right\|^2 &= \int d\xi \left| \sum_n c_n e^{in\xi} \right|^2 |\hat{\phi}(\xi)|^2 \\ &= \int_0^{2\pi} d\xi \left| \sum_n c_n e^{in\xi} \right|^2 \sum_{\ell \in \mathbb{Z}} |\hat{\phi}(\xi + 2\pi\ell)|^2 \,, \end{aligned}$$

(4.1) is equivalent with

$$0 < \frac{A}{2\pi} \le \sum_{\ell \in \mathbb{Z}} |\hat{\phi}(\xi + 2\pi\ell)|^2 \le \frac{B}{2\pi} < \infty \,, \tag{4.2}$$

a requirement that is satisfied by the tent function (as well as the higher order B-splines). We can therefore define $\tilde{\phi}$ by

$$\hat{\tilde{\phi}}(\xi) = \frac{\hat{\phi}(\xi)}{\left[2\pi \sum_\ell |\hat{\phi}(\xi + 2\pi\ell)|^2 \right]^{1/2}} \,; \tag{4.3}$$

because of the stability conditions (4.2) one easily checks that $\tilde{\phi} \in V_0$, and that the $\tilde{\phi}_{0n}$ span V_0 again, as the ϕ_{0n} did. Moreover $\sum_\ell |\hat{\tilde{\phi}}(\xi + 2\pi\ell)|^2 = (2\pi)^{-1}$, so that the $\tilde{\phi}_{0n}$ are orthonormal. One can then repeat the recipe of §2:

$$\begin{aligned} h_n &= \langle \tilde{\phi}, \tilde{\phi}_{-1,n} \rangle \\ \psi(x) &= \sqrt{2} \sum_n (-1)^n \, h_{-n+1} \, \tilde{\phi}(2x - n) \,, \end{aligned}$$

and the resulting $\psi_{j,k}$ will constitute an orthonormal basis associated with the given multiresolution analysis. Figures 9 and 10 show the functions $\tilde{\phi}$ and ψ for respectively linear and quadratic splines. These orthonormal spline bases were first constructed, independently and by completely different ad hoc methods by P. G. Lemarié (1989) and G. Battle (1988), before the advent of multiresolution analysis. Note that even though the original B-splines have compact support (of width $\ell + 1$ for splines of order ℓ), the orthogonalization trick (4.3) destroys this property; the resulting $\tilde{\phi}$ and ψ are supported on the whole line (with exponential decay).

The very first orthonormal basis of smooth wavelets, constructed by Stromberg (1982), also consists of spline functions; in terms of multiresolution analysis, the difference with the Battle-Lemarié wavelets is that another choice than (3.29)

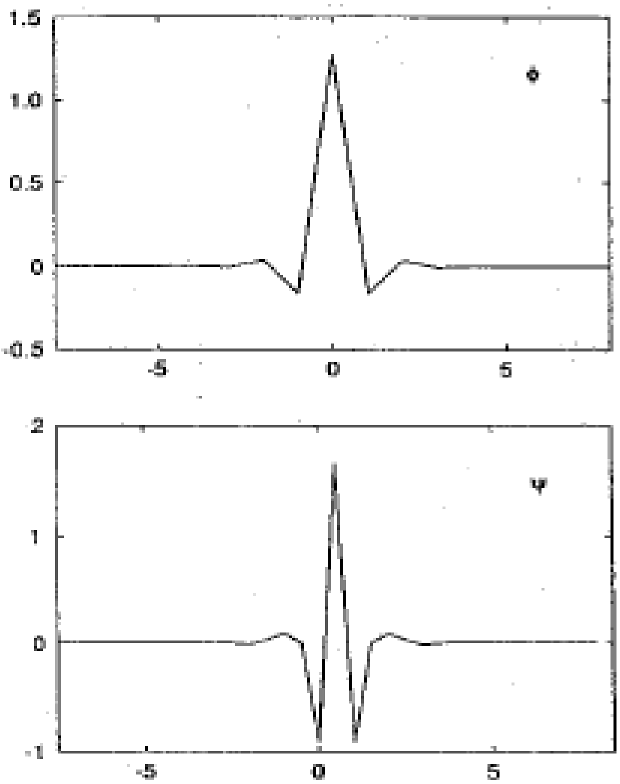

FIGURE 9. Scaling function and wavelet for the linear Battle-Lemarié spline basis.

is made: multiplying (3.29) by any 2π-periodic function of modulus 1 leads to another acceptable candidate for ψ.

Instead of wanting to reduce the spline multiresolution ladder with their very natural but nonorthogonal B-spline basis in every V_j to the case in §3, with orthonormal ϕ_{0n}, one can also try to stick to the B-splines, and characterize W_j, and find ψ, directly.

In this case, one still has

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

but

$$\int \phi(x)\, \overline{\phi(x-m)}\, dx = \gamma_m \neq \delta_{m0} .$$

Note that γ_m can also be written as

$$\gamma_m = \int |\hat{\phi}(\xi)|^2 e^{im\xi}\, d\xi = \int_0^{2\pi} e^{im\xi} \left(\sum_\ell |\hat{\phi}(\xi + 2\pi\ell)|^2 \right) d\xi . \tag{4.4}$$

Let us define $c(\xi) = \sum_m c_m e^{-im\xi}$, $\gamma(\xi) = \sum_m \gamma_m e^{-im\xi}$. Because of (4.4),

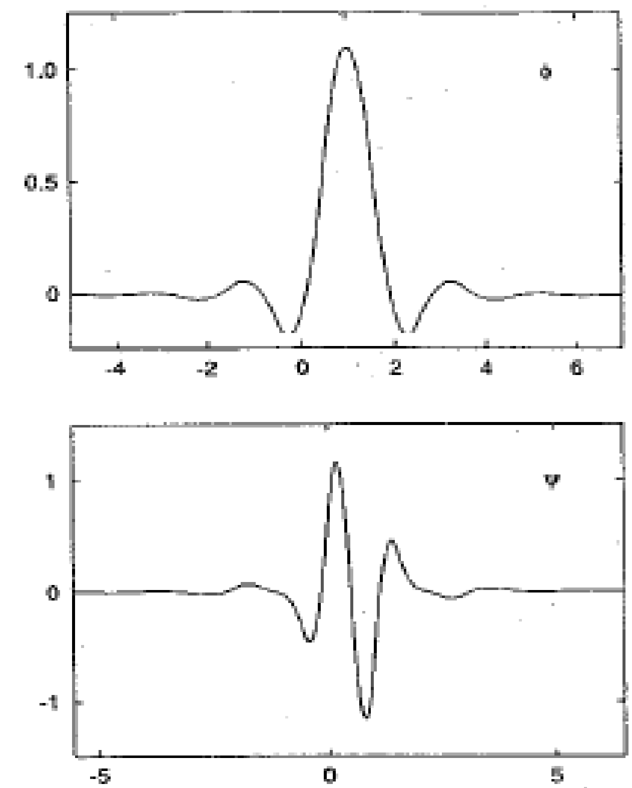

FIGURE 10. Scaling function and wavelet for the quadratic Battle-Lemarié spline basis.

$\gamma(\xi) = \frac{1}{2\pi} \sum_\ell |\hat{\phi}(\xi + 2\pi\ell)|^2$. Saying now that $\psi \in V_{-1} \cap (V_0)^\perp$ amounts to

$$\psi(x) = \sum_n d_n \phi(2x - n) , \tag{4.5}$$

and

$$0 = \int \psi(x) \, \overline{\phi(x-k)} = \sum_{n,m} d_n \, c_m \, \gamma_{2k-m+n}, \text{ for all } k \in \mathbb{Z} . \tag{4.6}$$

With the notation $d(\xi) = \sum_n d_n e^{-in\xi}$, (4.6) can be rewritten as

$$d(\xi) \, \overline{c(\xi)} \, \overline{\gamma(\xi)} + d(\xi + \pi) \, \overline{c(\xi + \pi)} \, \overline{\gamma(\xi + \pi)} = 0 ,$$

leading to the candidate solution

$$d(\xi) = -e^{i\xi} \, \overline{c(\xi + \pi)} \, \overline{\gamma(\xi + \pi)} ,$$

or

$$d_n = \sum_m (-1)^n c_{-m} \, \gamma_{-n+1+m} . \tag{4.7}$$

Since only finitely many c_n, g_n are nonzero, the same is true for the d_n.

Substituting (4.7) into (4.5) leads to a compactly supported function ψ such that the $\psi(x-k)$ constitute a Riesz basis for W_0; it then easily follows that the $\psi_{j,k}$ constitute a Riesz basis for all of $L^2(\mathbb{R})$. The dual Riesz basis in W_0 of the $\psi(x-k)$ is given by $\tilde{\psi}(x-k)$, with

$$\hat{\tilde{\psi}}(\xi) = \frac{\hat{\psi}(\xi)}{2\pi \sum_\ell |\hat{\psi}(\xi + 2\ell\pi)|^2} ;$$

$\tilde{\psi}$ is not compactly supported, but has exponential decay. The $\tilde{\psi}_{j,k}$ can easily be checked to be the dual Riesz basis for the $\psi_{j,k}$.

This alternative construction was proposed independently by Unser, Aldroubi and Eden (1990) and Chui and Wang (1991); there exist explicit formulas for ψ in terms of spline functions (Chui (1992)). Note that if one chooses to orthonormalize the $\psi(x-k)$, the result is the Battle-Lemarié wavelet again.

It is possible to choose for ψ a compactly supported spline function, and to find another compactly supported function $\tilde{\psi}$ such that the $\psi_{j,k}$, $\tilde{\psi}_{j,k}$ constitute dual Riesz bases. In this case however, one loses the orthonormality between the different j levels (unlike the construction above), and $\tilde{\psi}$ is not a spline function. For details, see Cohen, Daubechies and Feauveau (1992) or Daubechies (1992).

5. Fast algorithms.

Multiresolution analysis leads naturally to a hierarchical and fast scheme for the computation of the wavelet coefficients of a given function. Suppose that we have computed, or are given, the inner products of f with the $\phi_{j,k}$ at some given, fine scale. By rescaling our "units" (or rescaling f) we can assume that the label of this fine scale is $j = 0$. It is then easy to compute the $\langle f, \psi_{j,k} \rangle$ for $j \geq 1$. First of all, we have (see (3.31))

$$\psi = \sum_n g_n \, \phi_{-1,n} ,$$

where $g_n = \langle \psi, \phi_{-1,n} \rangle = (-1)^n \, h_{-n+1}$. Consequently,

$$\begin{aligned} \psi_{j,k}(x) &= 2^{-j/2} \, \psi(2^{-j}x - k) \\ &= 2^{-j/2} \sum_n g_n \, 2^{1/2} \, \phi(2^{-j+1}x - 2k - n) \\ &= \sum_n g_n \, \phi_{j-1,2k+n}(x) \\ &= \sum_n g_{n-2k} \, \phi_{j-1,n}(x) . \end{aligned} \tag{5.1}$$

(To simplify matters, we assume we are in the orthonormal case. All this can be generalized to the nonorthonormal but dual bases $\psi_{j,k}$, $\tilde{\psi}_{j,k}$ presented at the end of §4.)

It follows that

$$\langle f,\ \psi_{1,k}\rangle = \sum_n \overline{g_{n-2k}}\ \langle f,\ \phi_{0,n}\rangle ,$$

i.e. the $\langle f,\ \psi_{1,k}\rangle$ are obtained by convolving the sequence $(\langle f,\ \phi_{0,n}\rangle)_{n\in\mathbb{Z}}$ with $(\overline{g}_{-n})_{n\in\mathbb{Z}}$, and then retaining only the even samples. Similarly, we have

$$\langle f,\ \psi_{j,k}\rangle = \sum_n \overline{g_{n-2k}}\ \langle f,\ \phi_{j-1,n}\rangle , \tag{5.2}$$

which can be used to compute the $\langle f,\ \psi_{j,k}\rangle$ by means of the *same operation* (convolution with $\overline{g}$, decimation by factor 2) from the $\langle f,\ \phi_{j-1,k}\rangle$, if these are known. But, by (3.15),

$$\begin{aligned}\phi_{j,k}(x) &= 2^{-j/2}\ \phi(2^{-j}x-k)\\ &= \sum_n h_{n-2k}\ \phi_{j-1,n}(x) ,\end{aligned} \tag{5.3}$$

whence

$$\langle f,\ \phi_{j,k}\rangle = \sum_n \overline{h_{n-2k}}\ \langle f,\ \phi_{j-1,n}\rangle . \tag{5.4}$$

The procedure to follow is now clear: starting from the $\langle f,\ \phi_{0,n}\rangle$, we compute the $\langle f,\ \psi_{1,k}\rangle$ by (5.2), and the $\langle f,\ \phi_{1,k}\rangle$ by (5.4). We can then apply (5.2), (5.4) again to compute the $\langle f,\ \psi_{2,k}\rangle$, $\langle f,\ \phi_{2,k}\rangle$ from the $\langle f,\ \phi_{1,n}\rangle$, etc. At every step we compute not only the wavelet coefficients $\langle f,\ \psi_{j,k}\rangle$ of the corresponding j-level, but also the $\langle f,\ \phi_{j,k}\rangle$ for the same j-level, which are useful for the computation of the next level wavelet coefficients.

The whole process can also be viewed as the computation of successively coarser approximations of f, together with the difference in "information" between every two successive levels. In this view we start out with a fine-scale approximation to f, $f^0 = P_0 f$ (recall that P_j is the orthogonal projection onto V_j; we shall denote the orthogonal projection onto W_j by Q_j), and we decompose $f^0 \in V_0 = V_1 \oplus W_1$ into $f^0 = f^1 + \delta^1$, where $f^1 = P_1 f^0 = P_1 f$ is the next coarser approximation of f in the multiresolution analysis, and $\delta^1 = f^0 - f^1 = Q_1 f^0 = Q_1 f$ is what is "lost" in the transition $f^0 \to f^1$. In each of these V_j, W_j spaces we have the orthonormal bases $(\phi_{j,k})_{k\in\mathbb{Z}}$, $(\psi_{j,k})_{k\in\mathbb{Z}}$, respectively, so that

$$f^0 = \sum_n c_n^0\ \phi_{0,n},\ f^1 = \sum_n c_n^1\ \phi_{1,n},\ \delta^1 = \sum_n d_n^1\ \psi_{1,n} .$$

Formulas (5.2), (5.4) give the effect on the coefficients of the orthogonal basis transformation $(\phi_{0,n})_{n\in\mathbb{Z}} \to (\phi_{1,n},\psi_{1,n})_{n\in\mathbb{Z}}$ in V_0:

$$c_k^1 = \sum_n \overline{h_{n-2k}}\ c_n^0,\ d_k^1 = \sum_n \overline{g_{n-2k}}\ c_n^0 . \tag{5.5}$$

With the notations $a = (a_n)_{n\in\mathbb{Z}}$, $\overline{a} = (\overline{a_{-n}})_{n\in\mathbb{Z}}$ and $(Ab)_k = \sum_n a_{2k-n}\, b_n$, we can rewrite this as

$$c^1 = \overline{H}\, c^0 \ , \ d^1 = \overline{G}\, c^0 \ .$$

The coarser approximation $f^1 \in V_1 = V_2 \oplus W_2$ can again be decomposed into $f^1 = f^2 + \delta^2$, $f^2 \in V_2$, $\delta^2 \in W_2$, with

$$f^2 = \sum_n c_n^2\, \phi_{2,n}\ \delta^2 = \sum_n d_n^2\, \psi_{2,n} \ .$$

We have again

$$c^2 = \overline{H}\, c^1,\ d^2 = \overline{G}\, c^1 \ .$$

Schematically, all this can be represented as in Figure 11. In practice, we will

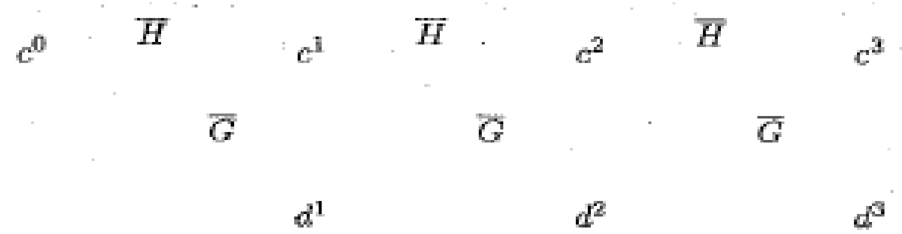

FIGURE 11. Schematic representation of (5.5)

stop after a finite number of levels, which means we have rewritten the information in $(\langle f,\ \phi_{0,n}\rangle)_{n\in\mathbb{Z}} = c^0$ as $d^1, d^2, d^3, \ldots, d^J$ and a final coarse approximation c^J, i.e. $(\langle f,\ \psi_{j,k}\rangle)_{k\in\mathbb{Z},\ j=1,\ldots,J}$ and $(\langle f,\ \phi_{J,k}\rangle)_{k\in\mathbb{Z}}$. Since all we have done is a succession of orthogonal basis transformations, the inverse operation is given by the adjoint matrices. Explicitly,

$$\begin{aligned} f^{j-1} &= f^j + \delta^j \\ &= \sum_k c_k^j\, \phi_{j,k} + \sum_k d_k^j\, \psi_{j,k} \ , \end{aligned}$$

hence

$$\begin{aligned} c_n^{j-1} &= \langle f^{j-1},\ \phi_{j-1,n}\rangle \\ &= \sum_k c_k^j\, \langle \phi_{j,k},\ \phi_{j-1,n}\rangle + \sum_k d_k^j\, \langle \psi_{j,k},\ \phi_{j-1,n}\rangle \\ &= \sum_k \left[h_{n-2k}\, c_k^j + g_{n-2k}\, d_k^j\right] \qquad (\text{use } (5.1), (5.3)) \ . \end{aligned} \tag{5.6}$$

An important aspect of the whole decomposition is that it is a fast algorithm. Let us go back to the Haar basis for a moment. If we start with N data points c_n^0, then we have to compute $N/2$ averages c_n^1, and $N/2$ differences d_n^1; from the $N/2$ different c_n^1 we compute $N/4$ averages c_n^2 and $N/4$ differences d_n^2, etc. The total number of computations is therefore $2\left(\frac{N}{2} + \frac{N}{4} + \ldots\right) = 2N$. For more sophisticated wavelet bases, the "averages" and "differences" involve more than just two numbers, but the same argument holds. If every "generalized average or difference" involves K coefficients of the previous level (rather than

2 as in the Haar case), then the total number of computations is $2KN$ (with KN multiplications, KN additions; this can be reduced further if the h_n have additional structure).

The orthonormal spline bases we saw in §4 have infinitely supported $\tilde{\phi}$ and ψ, resulting in infinitely many nonvanishing h_n. In practice, one needs to truncate to a finite number (otherwise we will hardly have a fast algorithm!). Since $\tilde{\phi}$, and therefore the h_n, have exponential decay, this truncating can in principle be done very easily; in practice one finds that K is rather large. This is one motivation to look at other multiresolution analysis ladders, where the emphasis is on the construction of ϕ associated with a finite number of h_n rather than on the choice of natural spaces V_j.

It should be noted that the fast algorithms associated with an orthonormal wavelet basis are also known, in electrical engineering, as a subband filtering scheme with exact reconstruction. Such schemes were constructed in EE by Smith and Barnwell (1986), Mintzner (1985) and Vetterli (1986), independently of, and in fact before, wavelets.

6. Orthonormal bases of compactly supported wavelet bases.

The easiest way to ensure compact support for the wavelet ψ is to choose the scaling function ϕ with compact support (in its orthogonalized version). It then follows from the definition of the h_n,

$$h_n = \sqrt{2} \int dx\ \phi(x)\ \overline{\phi(2x-n)}\ ,$$

that only finitely many h_n are nonzero, so that ψ reduces to a finite linear combination of compactly supported functions (see (3.31)), and therefore automatically has compact support itself.

For compactly supported ϕ the 2π-periodic function m_0,

$$m_0(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n\ e^{-in\xi}\ ,$$

becomes a trigonometric polynomial. As shown in §4, orthonormality of the $\phi_{0,n}$ implies

$$|m_0(\xi)|^2 + |m_0(\xi+\pi)|^2 = 1\ , \tag{6.1}$$

where we have dropped the "almost everywhere" because m_0 is necessarily continuous, so that (6.1) has to hold for all ξ if it holds a.e.

We are also interested in making ψ and ϕ reasonably regular. When we were working with spline spaces, we automatically controlled the regularity of ϕ and ψ. In this different setting, things are not as automatic. First of all, we have a necessary condition:

THEOREM 6.1. Suppose $f \in L^2(\mathbb{R})$ satisfies

$$\langle f_{j,k}, f_{j',k'} \rangle = \delta_{jj'} \delta_{kk'} ,$$

with $f_{j,k}(x) = 2^{-j/2} f(2^{-j}x - k)$. Suppose that f has compact support and that $f \in C^m$, with $f^{(\ell)}$ bounded for $\ell \leq m$. Then

$$\int dx\, x^\ell\, \bar{f}(x) = 0 \ for\ \ell = 0, 1, \ldots, m . \tag{6.2}$$

The idea of the proof is very simple. Choose j, k, j', k' so that $f_{j,k}$ is rather spread out, and $f_{j',k'}$ very much concentrated. On the tiny support of $f_{j',k'}$ the slice of $f_{j,k}$ "seen" by $f_{j',k'}$ can be replaced by its Taylor series, with as many terms as are well-defined. Since, however, $\int dx\, \overline{f_{j,k}(x)}\, f_{j',k'}(x) = 0$, this implies that the integral of the product of f and a polynomial of order m is zero. We can then vary the locations of $f_{j',k'}$, as given by k'. For each location the argument can be repeated, leading to a whole family of different polynomials of order m which all give zero integral when multiplied with f. This leads to the desired moment condition. For a true proof, see Daubechies (1992).

Since (see §4) $\hat{\psi}(\xi) = e^{-i\xi/2}\, \overline{m_0(\xi/2 + \pi)}\, \hat{\phi}(\xi/2)$, with $\hat{\phi}(0) = 1$, and since (6.2) is equivalent with $\frac{d^\ell}{d\xi^\ell}\hat{\psi}\,|_{\xi=0} = 0$ for $\ell = 0, 1, \ldots, m$, it follows that $\psi \in C^m$ implies that m_0 has a zero of order $m+1$ in π, or $m_0(\xi) = \left(\frac{1+e^{i\xi}}{2}\right)^{m+1} \mathcal{L}(\xi)$, with $\mathcal{L}$ again a trigonometric polynomial.

In addition to (6.1), we therefore also impose

$$m_0(\xi) = \left(\frac{1 + e^{i\xi}}{2}\right)^N \mathcal{L}(\xi) , \tag{6.3}$$

for some $N > 1$.

A first question is whether such m_0 exist. Taking the modulus square of (6.3) gives

$$|m_0(\xi)|^2 = \left(\cos^2 \frac{\xi}{2}\right)^N |\mathcal{L}(\xi)|^2 ,$$

where $|\mathcal{L}(\xi)|^2$ is a polynomial in $\cos\xi$, which can therefore also be written as a polynomial in $\sin^2 \frac{\xi}{2}$, i.e.

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

with P a polynomial. Substituting this into (6.1) leads to an equation for P,

$$x^N P(1-x) + (1-x)^N P(x) = 1 . \tag{6.4}$$

Because x^N and $(1-x)^N$ are two polynomials of degree N which are relatively prime, Bezout's theorem tells us that there exists a unique polynomial P of degree $N-1$ which solves (6.4). An explicit expression for P is given by

$$P(x) = \sum_{k=0}^{N-1} \binom{N-1+k}{k} x^k ,$$

which fortunately is positive for $0 < x < 1$, so that $P\left(\sin^2 \frac{\xi}{2}\right)$ is at least a possible candidate for $|\mathcal{L}(\xi)|^2$. There also exist higher degree solutions P to (6.4); they can be written as

$$P(x) = \sum_{k=0}^{N-1} \binom{N-1+k}{k} x^k + x^N R\left(x - \frac{1}{2}\right)$$

where R is an odd polynomial. We shall restrict ourselves to the lowest degree solution here.

Now that we have a candidate for $|\mathcal{L}(\xi)|^2$, the next question is to find $\mathcal{L}(\xi)$ itself. This can be achieved by the following lemma of Riesz, also known as "spectral factorization",

LEMMA 6.2. Let A be a positive trigonometric polynomial invariant under the substitution $\xi \to -\xi$; A is necessarily of the form

$$A(\xi) = \sum_{m=0}^{M} a_m \cos m\xi, \text{ with } a_m \in \mathbb{R} .$$

Then there exists a trigonometric polynomial B of order M, i.e.

$$B(\xi) = \sum_{m=0}^{M} b_m e^{im\xi}, \text{ with } b_m \in \mathbb{R} ,$$

such that $|B(\xi)|^2 = A(\xi)$.

The proof (which we skip here; details for this derivation can be found in many textbooks; they are also given in Daubechies (1988) or Daubechies (1992)) is constructive, so that we have a recipe for $\mathcal{L}(\xi)$ from $P(x)$.

All this leads us to a family of candidates $m_{0,N}$, with N the order of the zero at π, as in (6.2). Next we need to see how this determines ϕ and ψ. This is easy: since we expect $\phi \in L^1$, with $\int \phi(x)dx = 1$, $\hat{\phi}$ is continuous, with $\hat{\phi}(0) = \frac{1}{\sqrt{2\pi}}$, so that $\hat{\phi}(\xi) = m_0(\xi/2)\, \hat{\phi}(\xi/2)$ can be iterated, leading to

$$\begin{aligned}\hat{\phi}(\xi) &= \lim_{J\to\infty} \left[\prod_{j=1}^{J} m_0(2^{-j}\xi)\right] \hat{\phi}(2^{-J}\xi) \\ &= (2\pi)^{-1/2} \prod_{j=1}^{\infty} m_0(2^{-j}\xi) , \qquad (6.5)\end{aligned}$$

where the infinite product converges because m_0 is a trigonometric polynomial with $m_0(0) = 1$, so that

$$|m_0(\xi) - 1| \le C|\xi| \text{ for } |\xi| \le 1 .$$

It is rather straightforward to show (for details see Daubechies (1992)) that the infinite product (6.5) is an entire function of exponential type; more precisely, if

$$m_0(\xi) = \sum_{n=n_1}^{n_2} \alpha_n e^{-in\xi} ,$$

then

$$|\hat{\phi}(\xi)| \le C_1(1+|\xi|)^{M_1}\ e^{N_1|\mathrm{Im}\ \xi|} \quad \text{if Im } \xi \ge 0$$

$$|\hat{\phi}(\xi)| \le C_2(1+|\xi|)^{M_2}\ e^{N_2|\mathrm{Im}\ \xi|} \quad \text{if Im } \xi \le 0 ,$$

implying that ϕ is a distribution with support in $[N_1, N_2]$.

On the other hand, ϕ is also in L^2. We have indeed

$$\begin{aligned} \int |\hat{\phi}(\xi)|^2 d\xi &= \lim_{J\to\infty} \int_{|\xi|\le 2^J\pi} |\hat{\phi}(\xi)|^2\, d\xi \\ (6.6) \qquad &\le \lim_{J\to\infty} (2\pi)^{-1} \int_{|\xi|\le 2^J\pi} \prod_{j=1}^{J} |m_0(2^{-j}\xi)|^2\, d\xi \end{aligned}$$

(because $|m_0| \le 1$ by (6.1));

now

$$\begin{aligned} &\int_{|\xi|\le 2^J\pi} \prod_{j=1}^{J} |m_0(2^{-j}\xi)|^2\, d\xi \\ &= \int_0^{2^{J+1}\pi} \prod_{j=1}^{J} |m_0(2^{-j}\xi)|^2\, d\xi \qquad \text{(because of periodicity)} \\ &= \int_0^{2^J\pi} \left[\prod_{j=1}^{J-1} |m_0(2^{-j}\xi)|^2\right] \left[|m_0(2^{-J}\xi)|^2 + |m_0(2^{-J}\xi+\pi)|^2\right]\, d\xi \\ &= \int_0^{2^J\pi} \prod_{j=1}^{J-1} |m_0(2^{-j}\xi)|^2 = \ldots = \int_0^{4\pi} \left|m_0\left(\frac{\xi}{2}\right)\right|^2\, d\xi \\ &= \int_0^{2\pi} \left[\left|m_0\left(\frac{\xi}{2}\right)\right|^2 + \left|m_0\left(\frac{\xi}{2}+\pi\right)\right|^2\right]\, d\xi = 2\pi , \end{aligned}$$

so that (6.6) implies $\int |\hat{\phi}(\xi)|^2 d\xi \le 1$. It follows that ϕ, ψ are compactly supported L^2-functions, and things are looking good. There is one tricky step still, however:

all this is not sufficient to ensure that the $\phi(x-n)$ are orthonormal, nor even independent. A counterexample is

$$\begin{aligned} m_0(\xi) &= \left(\frac{1+e^{-i\xi}}{2}\right)(1-e^{-i\xi}+e^{-2i\xi}) \\ &= \frac{1+e^{-3i\xi}}{2} = e^{-3i\xi/2}\cos\frac{3\xi}{2} . \end{aligned}$$

This satisfies (6.1), as well as $m_0(0)=1$. Substituting it into (6.5) leads to

$$\hat{\phi}(\xi) = (2\pi)^{-1/2}\, e^{-3i\xi/2}\, \frac{\sin 3\xi/2}{3\xi/2}$$

or

$$\phi(x) = \begin{cases} 1/3 & 0 \le x \le 3 \\ 0 & \text{otherwise} . \end{cases}$$

This is not a "good" scaling function: the $\phi_{0,n}(x) = \phi(x-n)$ are not orthonormal, even though m_0 satisfies (6.1). Another way of looking at this is to see that (3.19) is not satisfied:

$$\sum_\ell |\hat{\phi}(\xi+2\pi\ell)|^2 = (2\pi)^{-1}\left[\frac{1}{3}+\frac{4}{9}\cos\ \xi+\frac{2}{9}\cos 2\xi\right] .$$

Note that this means that $\sum_\ell |\hat{\phi}(\xi+2\pi\ell)|^2 = 0$ for $\xi = \frac{2\pi}{3}$, so that even (4.2) is not satisfied: the $\phi_{0,n}$ are not even a Riesz basis for the space they span.

In order to avoid this kind of mishap, we have to impose extra conditions on m_0 to make sure that ϕ generates a true multiresolution analysis. These conditions ensure that

$$\sum_\ell |\hat{\phi}(\xi+2\pi\ell)|^2 = (2\pi)^{-1} \tag{6.7}$$

for all ξ. It turns out that this is the crucial condition: once (6.7) is satisfied, everything else follows automatically, and the $\psi_{j,k}$ constitute an orthonormal wavelet basis.

There are several ways of formulating necessary and sufficient conditions on m_0 ensuring that (6.7) holds, mostly due to Cohen (1990) and Lawton (1990); a detailed discussion is given in Daubechies (1992; sections 6.2, 6.3). A sufficient (but not necessary) condition implying (6.7) is (Mallat (1989)):

$$\min_{|\xi|\le\pi/2} |m_0(\xi)| > 0 .$$

Since this is satisfied for the $m_{0,N}$ we constructed above, everything is safe: for each N we have functions ϕ_N, ψ_N, of supportwidth $2N-1$, and the $2^{-j/2}\ \psi_N(2^{-j}x-k)$, $j,k\in\mathbb{Z}$, constitute an orthonormal basis for $L^2(\mathbb{R})$. Figure 12 shows a few examples for $N=2,3,5$.

How smooth are these functions? Clearly they are not as smooth as we might have hoped: even though we have zeros for m_0 at π of order resp. 2, 3, 5, the resulting ϕ are obviously not C^1, C^2 or C^4. Nevertheless they have higher regularity than the Haar basis (which was after all our goal), and their regularity increases with N. In fact, asymptotically, $\phi_N \in C^{\mu N}$ (for large N), with $\mu \simeq .2019$ (see Daubechies (1992; chapter 7)); ψ_N has the same regularity as ϕ_N.

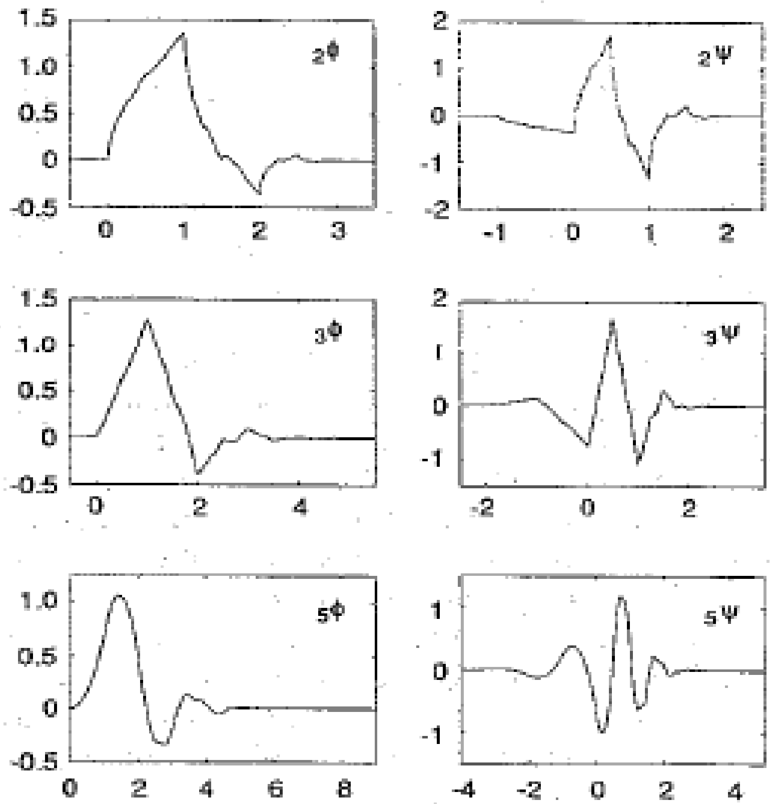

FIGURE 12.

7. Characterization of other function spaces than $L^2(\mathbb{R})$.

One of the interesting features of smooth wavelet basis is that they provide not only orthonormal bases for $L^2(\mathbb{R})$ but also unconditional bases for many other function spaces.

Let us first review the concept of "unconditional basis". A sequence of vectors $e_1, \ldots, e_n, \ldots$ in a (complex) separable Banach space E is a Schauder basis if, given any $x \in E$, we can find unique $\mu_n \in \mathbb{C}$ so that

$$\lim_{N \to \infty} \left\| x - \sum_{n=1}^{N} \mu_n e_n \right\| = 0 . \tag{7.1}$$

The basis is called "unconditional" if in addition, given any sequence $(\mu_n)_{n \in N}$ in $\mathbb{C}$, there is a criterium, using only the absolute values $|\mu_n|$, to decide whether or not $\sum_{n=1}^{N} \mu_n e_n$ converges to some x in E, as $N \to \infty$. Another equivalent way of stating this is the following: whenever $\sum_{n=1}^{N} \mu_n e_n$ is in E (in the sense that there exists $x \in E$ so that (7.1) holds), then $\sum_{n=1}^{\infty} \epsilon_n \mu_n e_n \in E$ as well, for any arbitrary choice of the $\epsilon_n = \pm 1$.

The Fourier basis $e_n(x) = e^{2\pi i n x}$, $n \in \mathbb{Z}$, for instance, is an unconditional basis for $L^2([0,1])$ (since it is an orthonormal basis for $L^2([0,1])$, but it is not an unconditional basis for any $L^p([0,1])$ for any $p \neq 1$. One can check for instance that the two series

$$\sum_{n=2}^{\infty} n^{-1/4} e^{2\pi i n x} \text{ and } \sum_{n=2}^{\infty} n^{-1/4} e^{i\sqrt{n}} e^{2\pi i n x}$$

both have their worst singularity at $x = 0$; the first one behaves like $|x|^{-3/4}$ for $|x| \to 0$, the second like $|\log x|$ for $x > 0$, $x \to 0$ and like $|x|^{-2}$ for $x < 0$, $x \to 0$ (see Zygmund (1968)). The first is therefore in $L^{7/6}([0,1])$, while the second isn't. Yet the absolute values of their coefficients are the same! No such problem in L^p-spaces exists if one uses the Haar basis. Restricting the Haar basis to only $[0,1]$, i.e. taking $\{\psi_{\mathrm{Haar};j,k};\ j \leq 0,\ 0 \leq k < 2^{|j|}\}$ and adding to this the constant function 1 on $[0,1]$ gives an unconditional basis for all $L^p([0,1])$ spaces with $1 < p < \infty$ (for $p = 1$ or $p = \infty$, $L^p([0,1])$ does not have an unconditional basis). For smoother function spaces, the discontinuous Haar functions are useless.

This is where smooth wavelet bases are useful. Because of their "logarithmic" treatment of the frequency components, similar to what happens in the Littlewood-Paley approach, they are "good" (i.e. unconditional) bases for L^p-spaces. Because they have good decay and smoothness properties, they are also "good" for function spaces with smoothness requirements, such as the Hölder, Sobolev or Besov spaces. There is no time in this lecture to discuss any of this in detail; let me just give a list of how one can characterize $f \in E$ by means of a criterium on only the $|\langle \psi_{j,k},\ f\rangle|$, for several function spaces E.

1. L^p-spaces

$$f \in L^p(\mathbb{R}) \iff \left[\sum_{j,k} |\langle f,\ \psi_{j,k}\rangle|^2\ |\psi_{j,k}(x)|^2\right]^{1/2} \in L^p(\mathbb{R})$$
$$\iff \left[\sum_{j,k} |\langle f,\ \psi_{j,k}\rangle|^2\ 2^{-j} \chi_{[2^j k, 2^j(k+1)]}(x)\right]^{1/2} \in L^p(\mathbb{R})\ .$$

2. Sobolev spaces

$$f \in W^s(\mathbb{R}) \quad = \quad \left\{f;\ \int (1 + |\xi|^2)^s\ |\hat{f}(\xi)|^2\, d\xi < \infty\right\}$$
$$\iff \sum_{j,k} |\langle f,\ \psi_{j,k}\rangle|^2\ (1 + 2^{-2js}) < \infty\ .$$

3. Hölder spaces.
For $s = n + \alpha$, with $n \in \mathbb{N}$ and $0 < \alpha < 1$, we define

$$C^s(\mathbb{R}) = \left\{ f \in L^\infty(\mathbb{R}) \cap C^n(\mathbb{R});\ \sup_{x,h} \frac{|f^{(n)}(x+h) - f^{(n)}(x)|}{|h|^\alpha} < \infty \right\} .$$

If ψ itself is in $C^r(\mathbb{R})$, with $r > s$ (hence the importance of the smoothness of ϕ, ψ!), then

$$f \in C^s(\mathbb{R}) \Longleftrightarrow \begin{cases} |\langle f, \phi_{0,k}\rangle| \le C & \text{for all } k \in \mathbb{Z} . \\ \text{and} & \\ |\langle f, \psi_{j,k}\rangle| \le C\, 2^{j(s+1/2)} & \text{for all } k \in \mathbb{Z}, \text{ all } j \le 0 . \end{cases}$$

Similar characterizations exist for all the Besov spaces (except those corresponding to L^1 or L^∞ conditions), for the Wiener "bump algebra", for the Hardy space H^1 of Stein and Weiss, for BMO, for the Zygmund class, etc. See Meyer (1990) for a thorough discussion.

Another important aspect of wavelet decompositions is that they are *local*. This can be exploited to characterize local smoothness properties of a function. Again, this can be done by looking at only the absolute values $|\langle f, \psi_{j,k}\rangle|$. For numerically stable computations of local Hölder exponents, it is however often more useful to consider redundant wavelet transforms (see Figure 9.2 in Daubechies (1992), and Mallat and Hwang (1992)).

8. Beyond wavelets.

Wavelets and wavelet transforms have proved useful in a variety of applications which exploit their smoothness, their good concentration in space, their scaling properties, and especially the fact that there exist fast algorithms. Some of these applications will be explained in more detail in this short course. In several of these applications, refinements of the constructions above are needed, such as multidimensional wavelet bases (a first construction of multidimensional wavelets is in Lemarié and Meyer (1986); see also Meyer (1992)) or wavelet bases adapted to an interval (Cohen, Daubechies and Vial (1992)).

There are of course also many applications where wavelets are not the best time frequency tool. Among these we find situations where the Fourier transform is the ideal tool, but also many cases where something intermediary is needed, with ideally a time-frequency analysis adapted to the signal, zooming in on transients (i.e. short-lived high frequency phenomena) with a wavelet-like approach whenever transients are present, but settling for a more Fourier-transform type decomposition for steadily oscillating components. Such more varied approaches can be achieved by means of a generalization of wavelets, called wavelet packets, or by the localized sine transform, an elegant and adaptive variant on the windowed Fourier transform. Both will be discussed in the following chapters.

References

1. E. W. Aslaksen and J. R. Klauder (1968), "Unitary representations of the affine group", *J. Math. Phys.* **9**, pp. 206–211; see also "Continuous representation theory using the affine group", *J. Math. Phys.* **10** (1969), pp. 2267-2275.
2. P. Auscher (1992), "Solution of two problems on wavelets", preprint, submitted to *Annals of Mathematics*.
3. G. Battle (1987), "A block spin construction of ondelettes. Part I: Lemarié functions", *Comm. Math. Phys.* **110**, pp. 601–615.
4. A. P. Calderón (1964), "Intermediate spaces and interpolation, the complex method", *Studia Math.* **24**, pp. 113–190.
5. C. K. Chui (1992), *An Introduction to Wavelets*, Academic Press, New York.
6. C. K. Chui and J. Z. Wang (1991), "A cardinal spline approach to wavelets", *Proc. Amer. Math. Soc.* **113**, pp. 785–793 and "On compactly supported spline wavelets and a duality principle", *Trans. Amer. Math. Soc.* **330** (1992) pp. 903–915.
7. A. Cohen (1990), "Ondelettes, analyses multirésolutions et filtres miroir en quadrature", *Ann. Inst. H. Poincaré, Analyse non linéaire* **7**, pp. 439–459.
8. A. Cohen, I. Daubechies and J. C. Feauveau (1990), "Biorthogonal bases of compactly supported wavelets", *Comm. Pure Appl. Math.* **45**, pp. 485–500.
9. A. Cohen, I. Daubechies and P. Vial (1992), "Wavelets on the interval and fast wavelet transforms", preprint, submitted to *Applied and Computational Harmonic Analysis*.
10. I. Daubechies (1988), "Time-frequency localization operators: a geometric phase space approach," *IEEE Trans. Inf. Th.* **34**, pp. 605–612.
11. I. Daubechies (1992), *Ten lectures on wavelets*, CBMS Lecture Notes nr. 61., SIAM, Philadelphia.
12. N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian and B. Torrésani (1992), "Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies", *IEEE Trans. Inf. Th.* **38**, pp. 644–664.
13. M. Frazier and B. Jawerth (1988), "The ϕ-transform and applications to distribution spaces", in *Function Spaces and Application*, M. Cwikel et al., eds., Lecture Notes in Mathematics 1302, pp. 233–246 (Springer, Berlin); see also "A discrete transform and decompositions of distribution spaces", *J. Func. Anal.*, (1990).
14. M. Holschneider and Ph. Tchamitchian (1990), "Régularité locale de la fonction 'non-différentiable' de Riemann", pp. 102–124 in Lemarié (1990).
15. J. R. Klauder and B.-S. Skagerstam (1985), *Coherent states*, (World Scientific, Signapore).
16. W. Lawton (1990), "Tight frames of compactly supported wavelets", *J. Math. Phys.* **31**, pp. 1898–1901; see also "Necessary and sufficient conditions for constructing orthonormal wavelet bases", *J. Math. Phys.* **32**, (1991) pp. 57–61.
17. P. G. Lemarié (1988), "Une nouvelle base d'ondelettes de $L^2(\mathbb{R}^n)$", *J. de Math. Pures et Appl.* **67**, pp. 227–236.
18. P. G. Lemarié and Y. Meyer (1986), "Ondelettes et bases hilbertiennes", *Rev. Mat. Iberoamericana* **2**, pp. 1–18.
19. P. G. Lemarié-Rieusset (1991), "Sur l'existence des analyses multirésolutions en
20. P. G. Lemarié and Y. Meyer (1986), "Ondelettes et bases hilbertiennes", *Rev. Mat. Iberoamericana* **2**, pp. 1–18. théorie des ondelettes", preprint, to appear in *Rev. Mat. Iberoamericana*.
21. P. G. Lemarié-Rieusset (1992), "Projecteurs invariants, matrice de dilation, ondelettes et analyses multirésolutions", preprint, submitted to *Rev. Mat. Iberoamericana*.
22. S. Mallat (1989), "Multiresolution approximation and wavelets", *Trans. Am. Math. Soc.* **315**, pp. 69-88.
23. S. Mallat and S. Zhong (1992), "Characterization of signals from multiscale edges," *IEEE Trans. PAMI* **14**.
24. Y. Meyer (1985), "Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs", Séminaire Bourbaki, 1985-1986, nr. 662.
25. Y. Meyer (1990), *Ondelettes ét opérateurs, I: Ondelettes, II: Opérateurs de Calderón-Zygmund, III: Opérateurs multilinéaires*, Hermann, Paris. An English translation is being

prepared by the Cambridge University Press 1992.

26. F. Mintzer (1985), "Filters for distortion-free two-band multirate filter banks", *IEEE Trans. Acoust. Speech Signal Process.* **33**, pp. 626-630.
27. M. J. T. Smith and T. P. Barnwell III (1986), "Exact reconstruction techniques for tree-structured subband coders", *IEEE Trans. ASSP* **34**, pp. 434–441; the basic results were already presented at the *IEEE Int. Conf. ASSP*, March 1984, San Diego.
28. J. O. Stromberg (1982), "A modified Franklin system and higher order spline systems on $\mathbb{R}^n$ as unconditional bases for Hardy spaces", Conf. in honor of A. Zygmund, Vol. II, pp. 475–493, ed. W. Beckner et al. (Wadsworth math. series).
29. M. Unser, A. Aldroubi and M. Eden (1990), "A family of polynomial spline wavelet transform", NCRR report 153/90, Nat. Inst. Health.
30. M. Vetterli (1986), "Filter banks allowing perfect reconstruction", *Signal Proc.* **10**, pp. 219–244; these results were already presented as "Splitting a signal into subsampled channels allowing perfect reconstruction", *IASTED Conf. on Applied Sig. Proc. and Dig. Filt.*, June 1985, Paris.
31. A. Zygmund (1968), *Trigonometric Series*, 2nd ed., Cambridge University Press, Cambridge.

AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey, 07974
E-mail: ingrid@research.att.com