Wavelets and Time-Frequency Analysis

Nikolaj Hess-Nielsen* Mladen Victor Wickerhauser†

November 22, 1995

Abstract

We present a selective overview of time-frequency analysis and some of its key problems. In particular we motivate the introduction of wavelet and wavelet packet analysis. Different types of decompositions of an idealized time-frequency plane provide the basis for understanding the performance of the numerical algorithms and their corresponding interpretations within the continuous models.

As examples we show how to control the frequency spreading of wavelet packets at high frequencies using nonstationary filtering and study some properties of periodic wavelet packets. Furthermore we derive a formula to compute the time localization of a wavelet packet from its indices which is exact for linear phase filters, and show how this estimate deteriorates with deviation from linear phase.

0 Introduction

We can decompose one-dimensional signals so as to illuminate two important properties: localization in time of transient phenomena, and presence of specific frequencies. The decomposition technique is expansion in wavelet orthonormal bases, *i.e.,* into independent components which have good time-frequency localization. Features in this context are just the basis elements which contribute large amplitudes to the expansion; they are detectable from their size. Alternatively, we can look for combinations of large components, or of not-so-large components that share similar time or frequency location. The localization of the basis elements does most of our work for us; when we find a large component, we can mark the time-frequency location of its basis element to build a time-frequency picture of the analyzed signal.

The basic technical problems are: assigning positions to component functions which are nonzero over large, possibly unbounded regions; assigning frequencies to component functions other than sines and cosines; fixing a decomposition when many alternatives are available;

*Visiting Research Associate, Dept. of Math., Washington University, St.Louis, MO 63130, USA. Research supported by the Danish Natural Science Research Council. E-mail: hess@math.wustl.edu

†Associate Professor, Dept. of Math., Washington University, St.Louis, MO 63130, USA. Research supported in part by the US National Science Foundation and the US Air Force Office of Scientific Research. E-mail: victor@math.wustl.edu

and coping with the defects of rapidly-computable families of basic time-frequency functions. The last problem will occupy most of this paper, since the applicability of time-frequency analysis techniques depends on their computational efficiency. Thus it is more reasonable to correct the deficiencies in fast transforms that work almost perfectly than to wait for slow mathematically perfect transforms to catch up.

The *time-frequency plane* is a two-dimensional space useful for idealizing these two properties of transient signals. We decompose a signal into pieces called *time-frequency atoms*, then draw idealized representations of these atoms in the plane. The time and frequency measurements contain uncertaintainty, and Heisenberg's inequality prevents us from making the product of the uncertainties smaller than a fixed constant. To depict a component of a signal, therefore, we may as well use an abstract plane figures, like a rectangle or an ellipse, whose position indicates the nominal time and frequency and whose shape suggests the relative uncertainties of the two quantities. Amplitude may be indicated by shading. Our task in this article will be to describe two such abstract depictions, and to discuss the technical problem of determining the location and shape of the component figures.

Suppose that ψ is a modulated waveform of finite total energy, and suppose that both the position and momentum uncertainties of ψ are finite:

$$\triangle x(\psi) \stackrel{\text{def}}{=} \inf_{x_0} \left(\frac{\|(x-x_0)\psi(x)\|}{\|\psi(x)\|} \right) < \infty; \qquad \triangle\xi(\psi) \stackrel{\text{def}}{=} \inf_{\xi_0} \left(\frac{\|(\xi-\xi_0)\hat{\psi}(\xi)\|}{\|\hat{\psi}(\xi)\|} \right) < \infty. \quad (1)$$

The position uncertainty $\triangle x(\psi)$ is the variance of the probability density function on $\mathbf{R}$ defined by $|\psi(x)|^2/\|\psi\|^2$. Likewise, the position x_0 where the minimum variance is achieved is the mean of that density. The corresponding momentum and momentum uncertainty are the mean and variance of the probability density by $|\hat{\psi}(x)|^2/\|\hat{\psi}\|^2$.

Finite $\triangle x$ requires that on average $\psi(x)$ decays faster than $|x|^{-3/2}$ as $|x| \to \infty$. Finite $\triangle\xi$ requires that ψ is smooth, in the sense that ψ' must also have finite energy. It is useful to introduce the *Schwartz class* $\mathcal{S}$ of smooth rapidly-decaying functions (see [13], pp.10–13). These are nice functions $\psi = \psi(x)$ with the property that $x^n \frac{d^m}{dx^m}\psi(x)$ is bounded and continuous for any nonnegative integers n, m. This class is preserved by differentiation, pointwise multiplication, and the Fourier transform. Note that every function ψ belonging to $\mathcal{S}$ satisfies Equation 1.

If ψ gives the instantaneous value of a time-varying signal, then it is reasonable to speak of *time* and *frequency* rather than position and momentum, especially since both pairs of quantities are related by the Fourier transform. We will say then that ψ is *well localized in both time and frequency* if the product of its time and frequency uncertainties is small. A musical note is an example of a time-frequency atom. It may be assigned two parameters, duration and pitch, which correspond to time uncertainty and frequency. A third parameter, location in time, can be computed from the location of the note in the score, since traditional music is laid out on a grid of discrete times and frequencies. We may name these three parameters *scale*, *frequency*, and *position*, to abstract them somewhat from the musical analogy.

Heisenberg's inequality imposes a lower bound on the *Heisenberg product*: $\triangle x \triangle \xi \geq \frac{1}{4\pi} \approx 0.08$. We need not be too precise about what we mean by "small" in this context;

it is enough to have a Heisenberg product of about one. We will call such functions *time-frequency atoms*. Not every Schwartz function is a time-frequency atom, but each one may be written as a linear combination of "unit" time-frequency atoms using rapidly decreasing coefficients:

Theorem 1 *For each $\psi \in \mathcal{S}$, there is a sequence $\{\phi_n : n = 1, 2, \ldots\} \subset \mathcal{S}$ of time-frequency atoms and a sequence of numbers $\{c_n : n = 1, 2, \ldots\}$ such that:*

1. *$\psi(t) = \sum_{n=1}^{\infty} c_n \phi_n(t)$, with uniform convergence;*
2. *$\|\phi_n\| = 1$ for all $n \geq 1$;*
3. *$\triangle x(\phi_n) \triangle \xi(\phi_n) < 1$ for all $n \geq 1$;*
4. *For each $d > 0$ there is a constant M_d such that $|c_n| n^d \leq M_d < \infty$ for all $n \geq 1$.* □

This theorem is proved by constructing the Littlewood-Paley decomposition of the given function ψ. Let us call a function ψ a *time-frequency molecule* if it satisfies the four conditions of Theorem 1. Notice that the last condition implies $\{c_n\}$ is absolutely summable.

The theorem states that all functions in $\mathcal{S}$ are time-frequency molecules. It also states that time-frequency atoms are dense in the Schwartz class. Since the Schwartz class in turn is dense in many other function spaces, we see that less regular functions can be decomposed into time-frequency atoms, though in general the coefficients $\{c_n\}$ will not decay rapidly. We can now place into context the surprising discovery by Yves Meyer [8] that a single sequence of orthonormal time-frequency atoms works for all Schwartz functions, and thus for many useful function spaces:

Theorem 2 *There is a sequence $\{\phi_n : n = 1, 2, \ldots\} \subset \mathcal{S}$ of time-frequency atoms with the following properties:*

1. *$\|\phi_n\| = 1$ for all $n \geq 1$;*
2. *If $m \neq n$, then $\langle \phi_m, \phi_n \rangle = 0$;*
3. *$\triangle x(\phi_n) \triangle \xi(\phi_n) < 1$ for all $n \geq 1$.*
4. *The set $\{\phi_n : n = 1, 2, \ldots\}$ is dense in $\mathcal{S}$.*

Also, for each $\psi \in \mathcal{S}$ there is a sequence of numbers $\{c_n : n = 1, 2, \ldots\}$ such that $\psi(t) = \sum_{n=1}^{\infty} c_n \phi_n(t)$ with uniform convergence, and for each $d > 0$ there is a constant M_d such that $|c_n| n^d \leq M_d < \infty$ for all $n \geq 1$. □

Meyer's theorem permits characterizing function spaces solely in terms of the rate of decay of positive sequences [5, 9, 10], and vastly simplifies calculating properties of operators such as continuity.

In an orthogonal *adapted waveform analysis*, the user is provided with a collection of standard libraries of waveforms—called *wavelets*, *wavelet packets*, and *windowed trigonometric waveforms*—which can be combined to fit specific classes of signals. All these functions are time-frequency atoms. In addition, it is sometimes useful to consider orthogonal libraries of

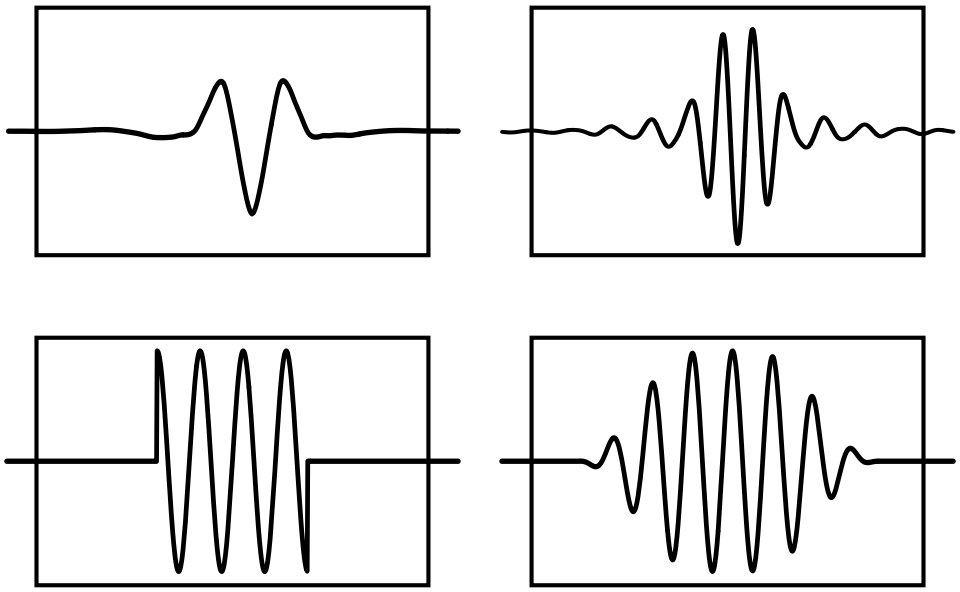

Figure 1: Example waveforms: wavelet, wavelet packet, block cosine, and local cosine functions.

functions which have large or unbounded Heisenberg product, such as *Haar–Walsh functions*, *block sines* and *block cosines*.

Examples of such waveforms are displayed in Figure 1.

Nonorthogonal examples of time-frequency atoms are easy to construct by modifying smooth bump functions. Suppose ϕ has finite Heisenberg product, e.g., take ϕ to be $O(t^{-2})$ as $|t| \to \infty$ and suppose ϕ' is continuous and $O(t^{-1})$ as $|t| \to \infty$. Then ϕ might not be in the Schwartz class $\mathcal{S}$, but it will be good enough for many practical applications. We define the *dilation*, *modulation*, and *translation* operators on ϕ by $\sigma^s\phi \stackrel{\text{def}}{=} \sigma_2^s\phi(t) = 2^{-s/2}\phi(2^{-s}t)$, $\mu_f\phi(t) = e^{2\pi i f t}\phi(t)$, and $\tau_p\phi(t) = \phi(t-p)$, respectively. If ϕ has small Heisenberg product, then the collection of dilated, modulated and translated ϕ's are also time-frequency atoms since the transformations σ, μ, τ conserve the Heisenberg product. They also conserve the energy of ϕ, so we may assume that the waveforms are all unit vectors in L^2, *i.e.*, they all have unit energy. If $\triangle x(\phi) = 1$, $\xi_0(\phi) = 0$ and $x_0(\phi) = 0$, then applying $\sigma_2^s, \mu_f, \tau_p$ moves these parameters to 2^s, f, and p, respectively.

In our analyses, we will say that the component of a function u at s, f, p is the inner product of u with the modulated waveform whose parameters are s, f, p. If the component is large, we may conclude that u has considerable energy at scale s near frequency f and position p.

We need to modify our notion of these parameters when we use real-valued time-frequency atoms, because the uncertainty results will be misleading. A real-valued function must have a Hermitean-symmetric Fourier transform: if $u(x) = \bar{u}(x)$, then $\hat{u}(-\xi) = \overline{\hat{u}(\xi)}$. Such a function must have $\xi_0(u) = 0$ no matter how much it oscillates, just because $\xi\,|\hat{u}(\xi)|^2$ is

an odd function. Another notion of frequency is needed in that case. For example, since $\int_0^\infty |\hat{u}|^2 = \int_{-\infty}^0 |u|^2 = \|u\|^2/2$, we could restrict our attention to the "positive" frequencies and use

$$\xi_0^+ \stackrel{\text{def}}{=} \frac{\left(2\int_0^\infty \xi |\hat{u}(\xi)|^2\, d\xi\right)^{1/2}}{\left(\int_0^\infty |\hat{u}(\xi)|^2\, d\xi\right)^{1/2}}. \tag{2}$$

This is equivalent to projecting the function u onto the Hardy space H^2 prior to calculating its power spectrum's center. The orthogonal projection $P : L^2 \mapsto H^2$ is defined by $\widehat{Pu}(\xi) = \mathbf{1}_{\mathbf{R}^+}(\xi)\, u(\xi)$, and the function Pu is called the *analytic signal* associated to the signal u.

If u is real-valued, then the frequency uncertainty $\triangle\xi(Pu)$, computed with $\xi_0^+ = \xi_0(Pu)$, is never larger than $\triangle\xi(u)$ computed with $\xi_0 = \xi_0(u)$. Unfortunately, P destroys decay, so that even a compactly supported u might have $\triangle x(Pu) = \infty$. The hypothesis that $\triangle x(Pu) \leq \triangle x(u) < \infty$ implies, by the Cauchy–Schwarz inequality, that both $u \in L^1$ and $Pu \in L^1$. Now if $Pu \in L^1$, then $\widehat{Pu}(\xi)$ must be continuous at $\xi = 0$, by the Riemann–Lebesgue lemma. This requires that $\hat{u}(0) = 0$.

1 The Time-Frequency Plane

We now consider an abstract two-dimensional signal representation in which time and frequency are indicated along the the horizontal and vertical axes, respectively. A waveform is represented by a rectangle in this plane with its sides parallel to the time and frequency axes, as seen in Figure 2. Let us call such a rectangle an *information cell*. The time and frequency of a cell can be read, for example, from the coordinates of its lower left corner. The uncertainty in time and the uncertainty in frequency are given by the width and height of the rectangle, respectively. Since the time and frequency positions are uncertain by the respective dimensions of the cell, it does not matter whether the nominal frequency and time position is taken from the center or from a corner of the rectangle. The product of the uncertainties is the area of the cell; it cannot be made smaller than the lower bound $1/4\pi$ given by Heisenberg's inequality.

Three waveforms whose information cells have nearly minimal area are drawn schematically in the signal plot at the bottom of Figure 2. The two at the left have small time uncertainty but big frequency uncertainty, with low and high modulation, respectively. Since they are evidently orthogonal, we have chosen to draw their information cells as disjoint rectangles. The wider waveform at the right has smaller frequency uncertainty, so its information cell is not so tall as the ones for the narrower waveforms. It also contains more energy, so its cell is darker than the preceding two. Notice that each information cell sits above its (circled) portion of the signal in this idealization.

The amplitude of a waveform can be encoded by darkening the rectangle in proportion to its waveform's energy. The idealized time-frequency plane closely resembles a musical score, and the information cells play the role of notes. However, musical notation does not indicate the pitch uncertainty by the shape of a note; for a particular instrument, this is determined by the duration of the note and the timbre of the instrument. Likewise, musical notation uses other means besides darkening the notes to indicate amplitude.

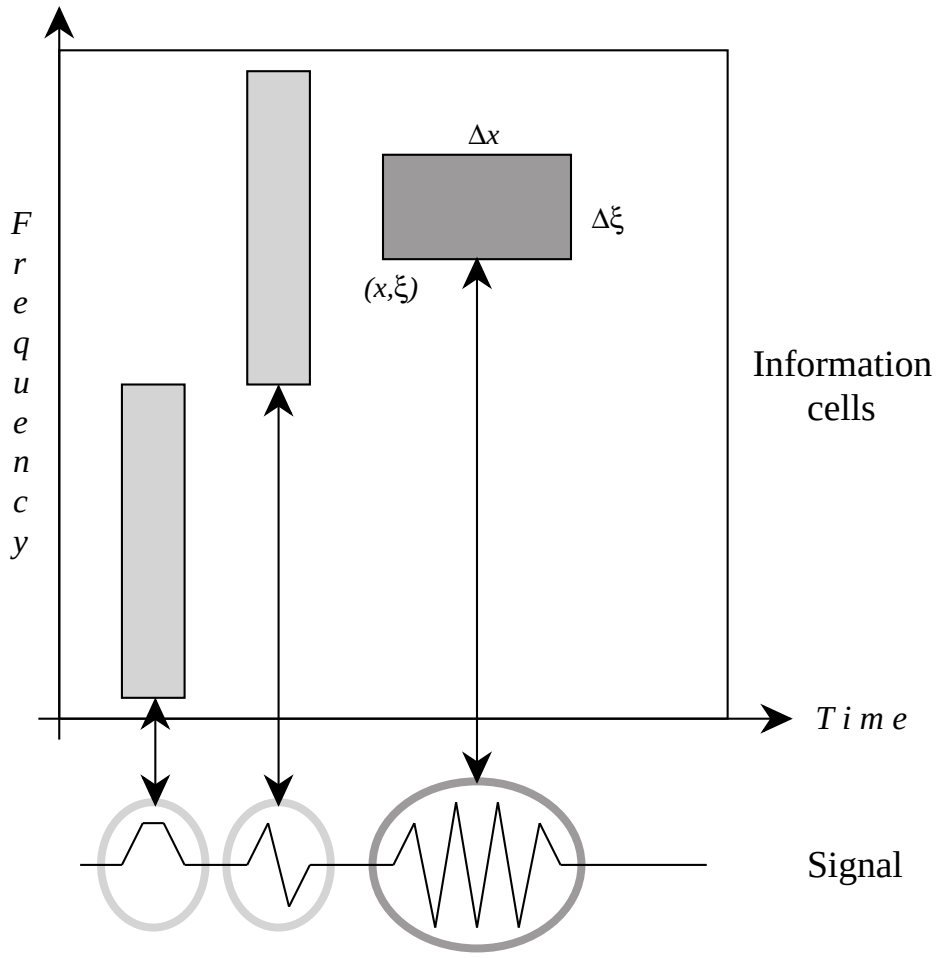

Figure 2: Information cells in the time-frequency plane.

Heisenberg's uncertainty principle for continuous waveforms implies that the area of an information cell can never be less than $1/4\pi$. Only the *Gaussian function* $g(t) = e^{-\pi t^2/2}$, suitably dilated, modulated, and translated, has the minimal information cell area. The other atoms are not too far off, though, and we will avoid the many restrictions of the Gaussian by relaxing the minimality condition. The only price we will have to pay is that a single atom might in practice require a few of the approximate atoms.

We now discuss in detail how to perform the approximate time-frequency analysis. If we have a signal of only finitely many points, then we can construct a finite version of the time-frequency plane. We will treat the signal sequence $\{a_N(k) : k = 0, 1, \ldots, N-1\}$ to be the coefficients of the function with respect to a particular N-member synthesis family Φ_N of time-frequency atoms:

$$f_N(t) = \sum_{k=0}^{N-1} a_N(k)\phi_{N,k}(t). \tag{3}$$

For any such finite signal approximation, the information cells will be confined to a finite region of the time-frequency plane, and their corners will lie on a discrete set of points determined by the sampling interval. If the signal is uniformly sampled at N points and we take the unit of length to be one sampling interval, then the width of the visible and relevant portion of the time-frequency plane is N. If $f \in L^2([0,1])$ and we use an N-dimensional approximation spanned by N time-frequency atoms which are translates of a single time-frequency atom ϕ supported in $[0,1]$, then Equation 3 specializes to the following:

$$f_N(t) = \sum_{k=0}^{N-1} a_N(k)\phi(Nt - k). \tag{4}$$

The signal may then be represented by adjacent information cells lined up at the grid points $\{\frac{k}{N} : 0 \le k < N\}$, with equal areas and with the k^{th} cell shaded to indicate its amplitude $a(k)$. The cells will be disjoint if the function ϕ is orthogonal to its translates by integers, *i.e.*, if $k \neq j \Rightarrow \int \phi(Nt-k)\phi(Nt-j)\,dt = 0$.

The Fourier exponential functions $1, e^{2\pi ix/N}, \ldots, e^{2\pi i(N-1)x/N}$ form an orthogonal basis for all such N-sampled functions. If our basic oscillating function is written as $e^{2\pi i\frac{fx}{N}}$, this means that the frequency index f ranges over the values $0, 1, \ldots, N-1$, so there are N discrete values for the frequency index and we may introduce N equally spaced points on the frequency axis to account for these. Thus the smallest region that contains all possible cells for a signal of length N must be N time units wide by N frequency units tall, for a total area of N^2 time-frequency units.

If N is even, we may use the equivalent numbering $-\frac{N}{2}, \ldots, -1, 0, 1, \ldots, \frac{N}{2} - 1$ for the frequency indices. Notice that, since we cannot distinguish the exponential of frequency f from the one at $N - f$ just by counting oscillations, there are really only $N/2$ distinguishable frequencies in the list. This, suitably rigorized, is called the *Nyquist theorem*; the maximum distinguishable or *Nyquist frequency* for this sampling rate is $N/2$ oscillations in N units, or $1/2$.

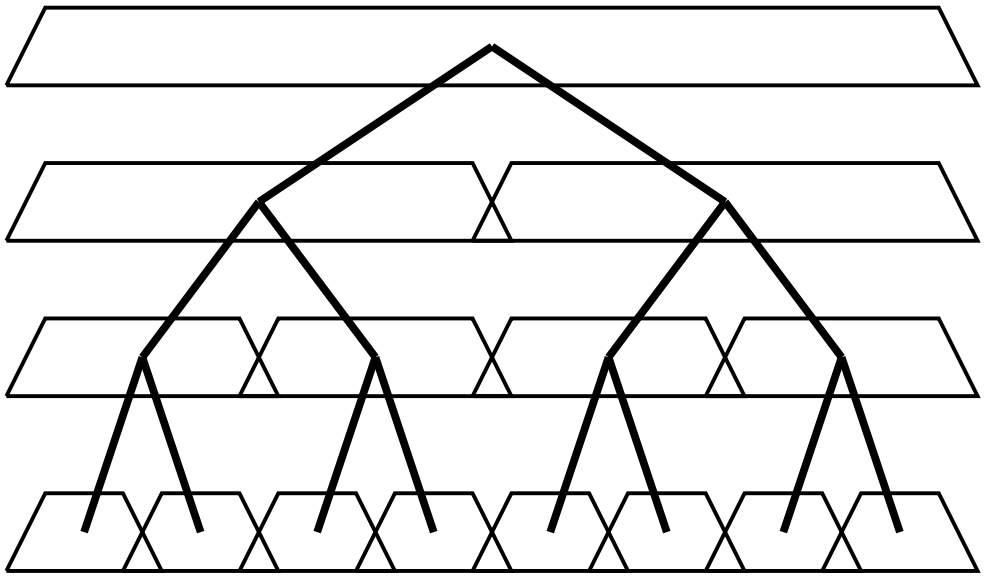

Figure 3: Splitting a signal into two pieces splits the time-frequency plane into two halves.

1.1 Splitting Algorithms

Rapidly computable expansions of a signal into time-frequency components are accomplished through recursive splitting algorithms. As depicted in Figure 3, the input signal is cut into two pieces by a pair of operators represented by the left-leading and right-leading branches departing from each intersection. This has the effect of splitting the time-frequency plane into two halves. The operators should produce orthogonal or independent parts so that the two halves do not overlap. They should also be decimating operators, so that the samples in the two parts add up to the total number of signal samples.

One way to accomplish this is to cut the signal up into windows and then compute a Fourier transform within wach window. The effect of this on the time-frequency plane is easy to visualize from Figure 3. Each trapezoidal block in that figure, produced by restricting to a portion of the signal in time, corresponds to a time-frequency rectangle of that block's width and of maximum height. The Fourier transform then cuts that rectangle into equal frequency-height pieces.

Another way to decompose a signal is to filter it into low and high frequency components using a conjugate pair of decimating quadrature filters. That operation and its inverse are depicted in Figure 4. Operation **h** denotes low-pass filtering and might be viewed as the path down the left fork; operation **g** denotes high-pass filtering and might be viewed as the path down the right fork. The pair of operators splits the time-frequency plane into a bottom half rectangle and a top half rectangle, each as tall as the corresponding trapezoid of Figure 3 is wide. Each half contains one information cell for each sample in the trapezoid, and since the filters decimate the information cells get wider as their height decreases.

The splitting algorithms and their interpretations as recursive decompositions of the time-frequency plane are well known to the engineering community. The windowed or short-time Fourier transform algorithm is used to compute and plot spectrograms, while the filter bank

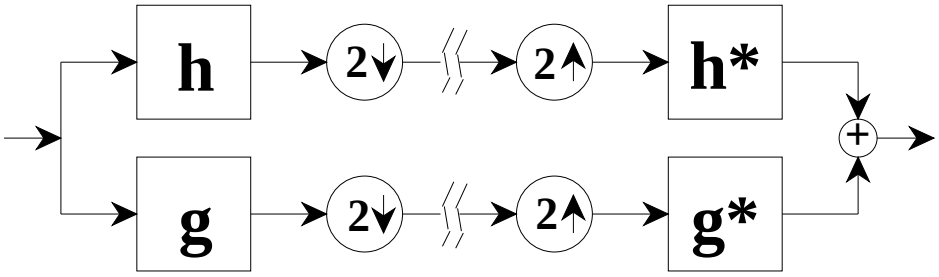

Figure 4: Low pass, high pass decimating filter bank and its inverse.

algorithm is used both in spectral analysis and in signal coding.

1.2 Bases and Tilings

A family of time-frequency atoms with uniformly bounded Heisenberg product may be represented by information cells of approximately equal area. A basis of such atoms corresponds to a cover of the plane by rectangles; an orthonormal basis may be depicted as a cover by disjoint rectangles. Certain bases have characterizations in terms of the shapes of the information cells present in their cover of the time-frequency plane. For example, the *standard basis* or *Dirac basis* consists of the cover by the tallest, thinnest patches allowed by the sampling interval and the underlying synthesis functions. The Dirac basis has optimal time localization and no frequency localization, while the Fourier basis has optimal frequency localization, but no time localization. These two bases are depicted in Figure 5.

The Fourier transform may be regarded as a rotation by 90° of the standard basis, and as a result the information cells are transposed by interchanging time and frequency. We may note that it is also possible to apply an element of the *Hermite group* [4] (also called the *angular Fourier transform*) to obtain information cells which make arbitrary angles with the time and frequency axes. This transform is a pseudodifferential operator with origins in quantum mechanics; it is formally represented by $A_t \stackrel{\text{def}}{=} \exp(-itH)$ where t is the angle from the horizontal which we wish to make with our rotated atoms, and H is the selfadjoint Hamiltonian operator obtained by quantization of the harmonic oscillator equation:

$$H = \frac{1}{2}\left(\frac{d^2}{dx^2} + x^2\right). \tag{5}$$

Then A_t is the evolution operator which produces the wave function $u(x,t) = A_t u(x,0)$ from an initial state $u(x,0)$, assuming that the function u evolves to satisfy the Schrödinger equation

$$\frac{du}{dt} + Hu = 0. \tag{6}$$

We bring all this up mainly to apprise the reader that many ideas used in the time-frequency analysis of signals have their roots in quantum mechanics and have been studied by physicists and mathematicians for several generations.

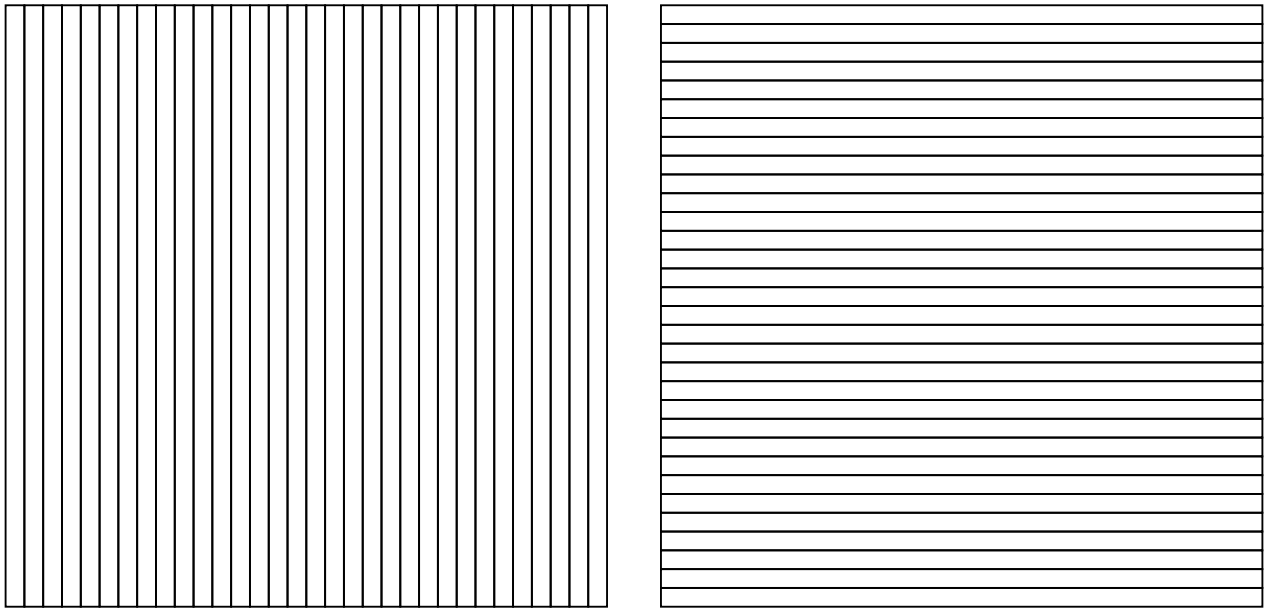

Figure 5: Dirac and Fourier bases tile the time-frequency plane.

Windowed Fourier or trigonometric transforms with a fixed window size correspond to covers with congruent information cells whose width $\triangle x$ is proportional to the window width. The ratio of frequency uncertainty to time uncertainty is the aspect ratio of the information cells, as seen in Figure 6

The wavelet basis is an octave-band decomposition of the time-frequency plane, depicted by the covering on the left in Figure 7. A wavelet packet basis gives a more general covering; the one on the right in Figure 7 is appropriate for a signal containing two almost pure tones near 1/3 and 3/4 of the Nyquist frequencies, respectively. Tilings which come from graph basis in library trees built through convolution and decimation must always contain complete rows of cells, since they first partition the vertical (frequency) axis and then fill in all the horizontal (time) positions.

An adapted local trigonometric transform tiles the plane like the left part of Figure 8. Such bases are transposes of the wavelet packet bases, since wavelet packets are related to local trigonometric functions by the Fourier transform. The tilings corresponding to graphs in a local trigonometric library tree must contain complete columns, since these first segment the horizontal (time) axis and then represent all the vertical positions (or frequencies) within each segment.

The right part of Figure 8 shows a more general tiling of the time-frequency plane, one which does not correspond to a graph in either a wavelet packet or local trigonometric library tree. Such a decomposition is achievable using the Haar quadrature filters and picking additional basis subsets besides graphs, or else by combining the local trigonometric and wavelet packet bases on a segmented signal. The number of such tilings is an order of magnitude greater than the number of graph wavelet packet bases.

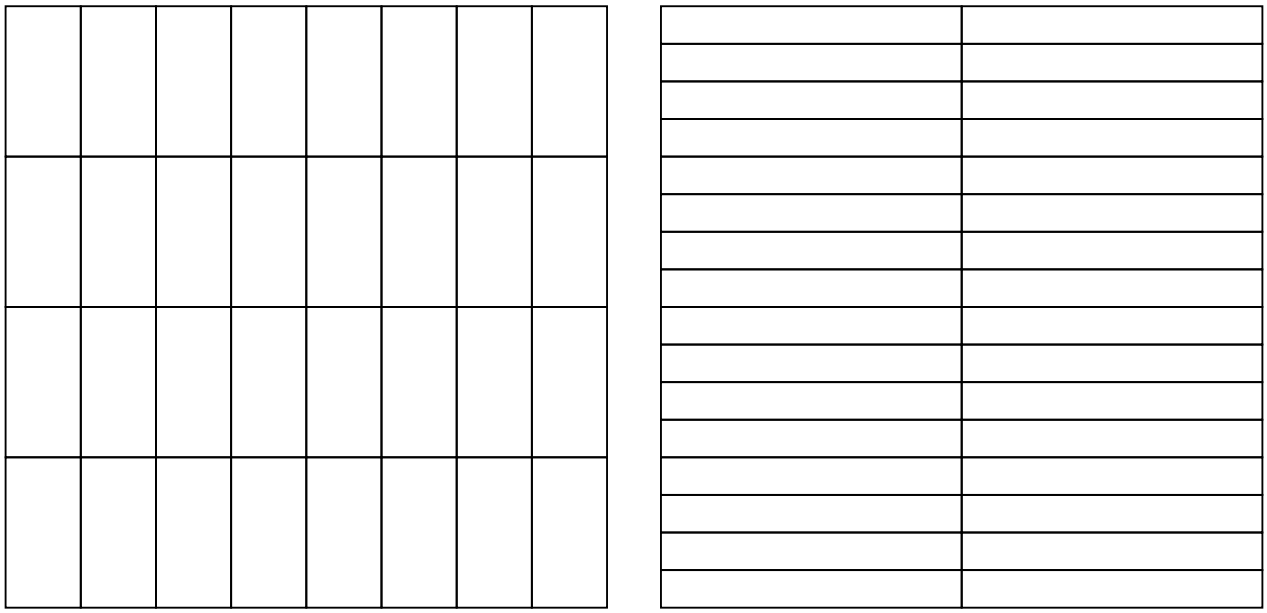

Figure 6: Windowed Fourier bases tile the time-frequency plane.

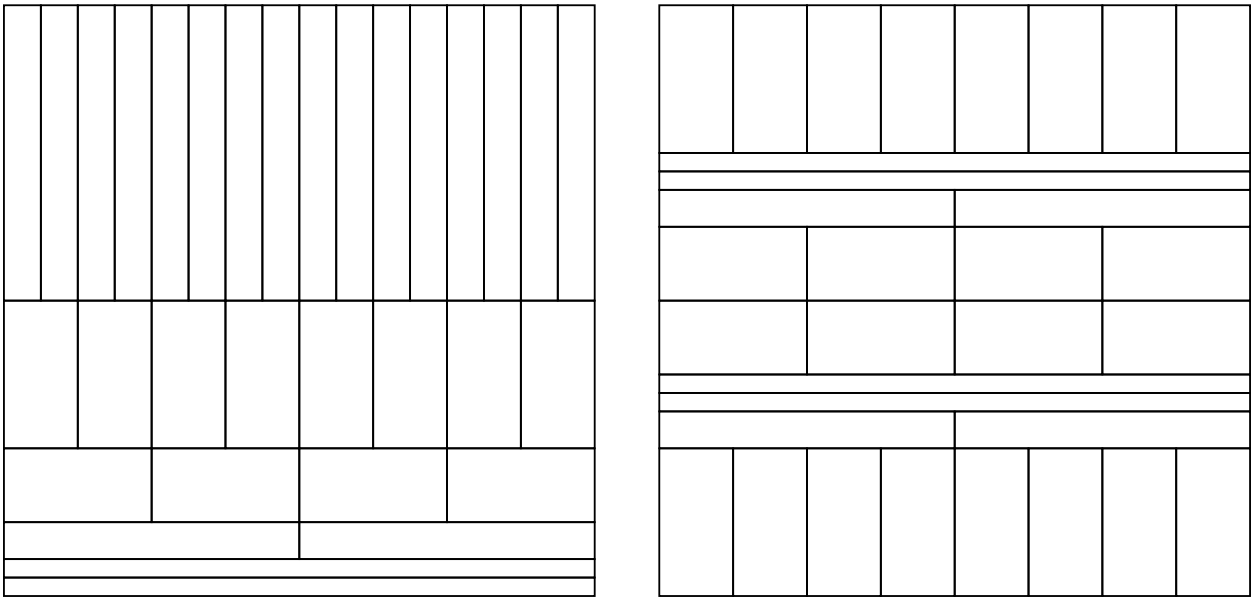

Figure 7: Wavelet and wavelet packet bases tile the time-frequency plane.

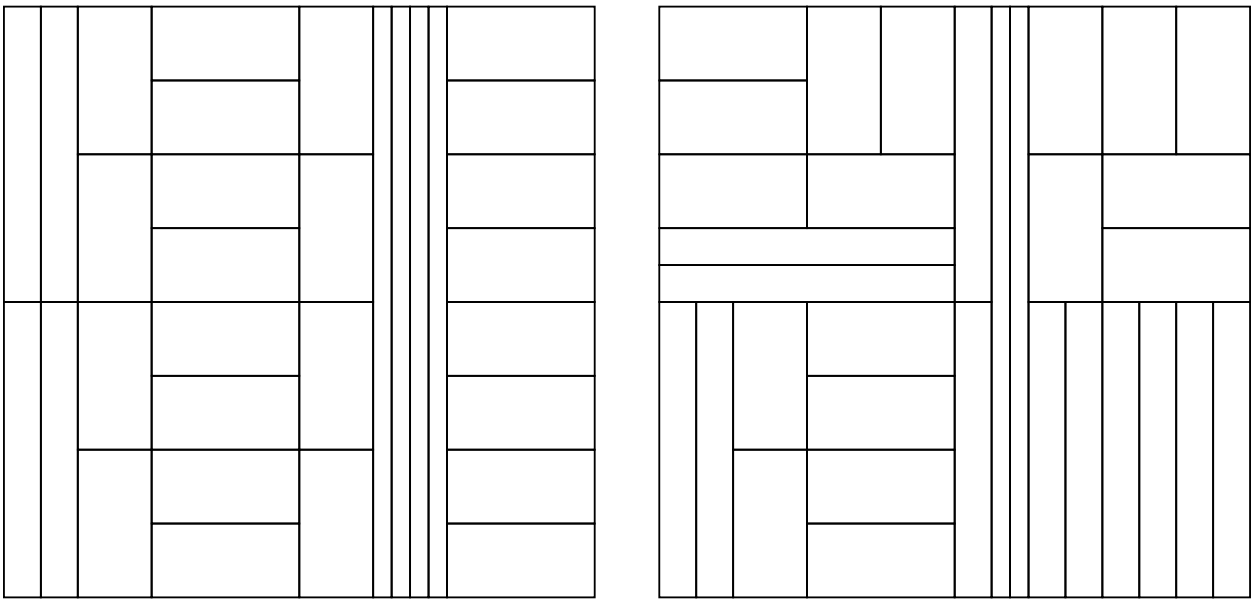

Figure 8: Adapted local trigonometric tiling and a general dyadic tiling.

1.3 Time-Frequency Analysis with Wavelet Packets

The scale, frequency, and position indices of wavelet packets can be used to draw an information cell in the time-frequency plane. We now derive the formulas for the nominal values of x_0, ξ_0, $\triangle x$, and $\triangle \xi$ for such functions.

Let ψ_{sfp} be the wavelet packet with scale index s, frequency index f, and position index p. We use conventional indexing for the filters and Paley or natural ordering for the frequency. Let us further suppose that the signal consists of $N = 2^L$ equally spaced samples, and that the library tree contains a complete wavelet packet analysis down to level L. Then we have $0 \le s \le L$, $0 \le f < 2^s$, and $0 \le p < 2^{L-s}$.

The scale parameter s gives the number of levels of decomposition below the original signal. Each application of convolution and decimation doubles the nominal width, so we set $\triangle x = 2^s$. With the usual assumption that $\triangle x \cdot \triangle \xi \approx N$, we can thus assign $\triangle \xi = 2^{L-s}$.

The frequency parameter must first be corrected by using the inverse Gray code permutation, so we compute $f' = GC^{-1}(f)$. This produces an index which is again in the range 0 to $2^s - 1$. The lower left-hand corner of the information cell should then be placed at vertical position $\triangle \xi \cdot f' = 2^{L-s} f'$.

The position parameter p needs to be shifted to correct for the (frequency-dependent) phase response of quadrature filters. This shift and its causes are discussed in Theorem 15, and Corollary 2 may be used to compute the amount by which to shift the horizontal location of the information cell:

$$p' = 2^s p + (2^s - 1)c[h] + (c[g] - c[h])f''. \tag{7}$$

Here $c[h]$ and $c[g]$ are the centers of energy of the low-pass and high-pass QMFs h and g, respectively, and f'' is the bit-reverse of f considered as an s-bit binary integer. The result

will be inaccurate by at most the deviation ϵ_h of the filter h from linear phase, as defined in Theorem 15. If the wavelet packet coefficients were computed using periodized convolution-decimation, then we should replace the equal sign in Equation 7 with congruence modulo 2^L and take p' in the range 0 to $2^L - 1$.

Since the horizontal position p' of the information cell is uncertain by $\triangle x = 2^s$, we may as well slide its lower left-hand corner horizontally left to the nearest multiple of 2^s below p'. To emphasize this uncertainty, we will use a position index p'' in the range 0 to $2^{L-s} - 1$, with each integer value representing an interval starting at an integer multiple of $\triangle x = 2^s$:

$$p'' = \lfloor p'/2^s \rfloor = \left\lfloor p + (1 - 2^{-s})c[h] + 2^{-s}(c[g] - c[h])f'' \right\rfloor . \tag{8}$$

These conventions, in the periodic case, produce a disjoint tiling which exactly fills the $N \times N = 2^L \times 2^L$ square time-frequency plane:

Theorem 3 *If $B = \{(s, f, p) \in B\}$ is the index subset of a wavelet packet graph basis for a 2^L-point signal, then the collection of rectangles*

$$\{[2^s p'', 2^s(p'' + 1)[\times[2^{L-s} f', 2^{L-s}(f' + 1)[: (s, f, p) \in B\} \tag{9}$$

is a disjoint cover of the square $[0, 2^L[\times[0, 2^L[$. □

These information cells will be located approximately where they should be to describe the time-frequency content of their wavelet packets.

1.4 Time-Frequency Analysis with Adapted Local Trigonometric Functions

For this discussion, it makes no difference whether we use local cosines or local sines. The one is obtained from the other merely by reversing the direction of time, which has no effect on the geometry of information cells. We again suppose that the signal consists of $N = 2^L$ equally spaced samples, and that the library tree contains all the local trigonometric analyses to level L, with windows of size $2^L, 2^{L-1}, \ldots, 1$. The basis functions will be indexed by the triplet (s, f, p), and we will have $0 \le s \le L$, $0 \le f < 2^{L-s}$, and $0 \le p < 2^s$.

The scale parameter s again gives the number of decompositions of the original signal window into subwindows. Each subdivision halves the nominal window width, so we set $\triangle x = 2^{L-s}$. With the usual assumption that $\triangle x \cdot \triangle \xi \approx N$, we can thus assign $\triangle \xi = 2^s$.

The position index p numbers the adjacent windows starting with zero at the left edge of the signal. Thus the information cell should be drawn over the horizontal (time) interval $I_{sp} \stackrel{\text{def}}{=} [2^{L-s}p, 2^{L-s}(p + 1)[$.

One local cosine basis for the subspace over the subinterval I_{sp} consists of the functions $\cos \pi \left(f + \frac{1}{2}\right) n/2^{L-s} = \cos \pi 2^s \left(f + \frac{1}{2}\right) n/N$, where $n \in I_{sp}$ and $0 \le f < 2^{L-s}$, multiplied by the window function subordinate to I_{sp}. These orthonormal basis functions have nominal frequencies $2^s \left(f + \frac{1}{2}\right)$, so we will draw the associated information cell alongside the interval $[2^s f, 2^s(f + 1)[$ on the vertical (or frequency) axis. Local sine bases can be depicted in a similar way.

A graph basis set of triplets produces a disjoint cover of such information cells:

Theorem 4 *If $B = \{(s,f,p) \in B\}$ is the index subset of an adapted local trigonometric graph basis for a 2^L-point signal, then the collection of rectangles*

$$\{[2^{L-s}p, 2^{L-s}(p+1)[\times[2^s f, 2^s(f+1)[: (s,f,p) \in B\} \tag{10}$$

is a disjoint cover of the square $[0, 2^L[\times[0, 2^L[$. □

These information cells will be located in the time-frequency plane at the appropriate locations for their position and frequency content. Notice how local trigonometric functions have indices s, f, p that can be directly interpreted as scale, frequency, and position: no Gray coding, bit-reversal or phase shifting is needed. In the wavelet packet case, the indices s, f, p used in the formulas to compute time-frequency components must first be adjusted to s, f', p'' in order to correspond to scale, frequency, and position. Otherwise, both types of decompositions provide complete orthonormal bases of time-frequency atoms.

1.5 Arbitrary Tilings with Haar–Walsh Functions

The Haar-Walsh wavelet packet library has a special property not shared by libraries of smoother wavelet packets. Namely, we can put all its waveforms into one-to-one correspondence with dyadic information cells which are located where they should be, but in such a way that every disjoint cover corresponds to an orthonormal basis. The total number of disjoint covers of $2^L \times 2^L$ by dyadic rectangles—*i.e.*, those whose coordinates are of the form $n2^j$ for integers n and j—is much greater than the number of graph bases for a 2^L-point signal analyzed to level L.

We define the correspondence as follows: Let ψ_{sfp} be the Haar–Walsh wavelet packet on $N = 2^L$ points with scale index $0 \le s \le L$, Paley order frequency index $0 \le f < 2^s$, and unshifted position index $0 \le p < 2^{L-s}$. We associate to it the rectangle

$$\psi_{sfp} \leftrightarrow R_{sfp} \stackrel{\text{def}}{=} [2^s p, 2^s(p+1)[\times[2^{L-s}\tilde{f}, 2^{L-s}(\tilde{f}+1)[\subset [0, N[\times[0, N[. \tag{11}$$

Here $\tilde{f} = GC^{-1}(f)$ is the inverse Gray code permutation of f, which adjusts the vertical location of R_{sfp} so it is proportional to the number of oscillations of the wavelet packet ψ_{sfp}. Since $c[h] = 1/2$ and $c[g] - c[h] = 0$ for the Haar–Walsh filters, we need not shift or bit-reverse the position index p to get the actual horizontal location of the information cell. This puts R_{sfp} where it should be to describe the location and oscillation of ψ_{sfp}. We remark that two rectangles will be disjoint in the sequency ordering of their frequency indices if and only if they are disjoint in the Paley ordering.

Theorem 5 *The Haar–Walsh wavelet packets $\{\psi_{sfp} : (s,f,p) \in B\}$ form an orthonormal basis of $\mathbf{R}^N$ if and only if the dyadic rectangles $\{R_{sfp} : (s,f,p) \in B\}$ form a disjoint cover of $[0, N[\times[0, N[$.*

Proof: Since the area of R_{sfp} is N while the area of $[0, N[\times[0, N[$ is N^2, it suffices to show that two rectangles R_{sfp} and $R_{s'f'p'}$ are disjoint if and only if ψ_{sfp} and $\psi_{s'f'p'}$ are orthogonal. Since exactly N rectangles fit into the square, and the space is N-dimensional, we must have a basis set of wavelet packets.

Let $I_{sp} = [2^s p, 2^s(p+1)[$ and $I_{s'p'} = [2^{s'}p', 2^{s'}(p'+1)[$ be the supports of ψ_{sfp} and $\psi_{s'f'p'}$, respectively, and consider the two cases:

$s = s'$: then ψ_{sfp} and $\psi_{sf'p'}$ will be orthogonal if and only if $f \neq f'$ or $f = f'$ but $p \neq p'$, one of which will be true if and only if the congruent rectangles R_{sfp} and $R_{sf'p'}$ are disjoint. In this case, both ψ_{sfp} and $\psi_{sf'p'}$ are members of the same single-level or subband basis, which is an orthonormal graph basis.

$s \neq s'$: then ψ_{sfp} and $\psi_{s'f'p'}$ will be orthogonal if and only if

$$\int_{I_{sp} \cap I_{s'p'}} \psi_{sfp}(t)\psi_{s'f'p'}(t)\,dt = 0. \tag{12}$$

But two dyadic intervals either are disjoint or one contains the other. If $I_{sp} \cap I_{s'p'} = \emptyset$, then $R_{sfp} \cap R_{s'f'p'} = \emptyset$ since the rectangles share no horizontal coordinates. If the overlap is nonempty, then we may suppose without loss that $s < s'$ and thus $I_{sp} \subset I_{s'p'}$. We can define the *nominal frequency intervals* $J_{sf} \stackrel{\text{def}}{=} [2^{L-s}\tilde{f}, 2^{L-s}(\tilde{f}+1)[$ and $J_{s'f'} = [2^{L-s'}\tilde{f}', 2^{L-s'}(\tilde{f}'+1)[$ of ψ_{sfp} and $\psi_{s'f'p'}$, respectively. Since these too are dyadic intervals, and $L - s > L - s'$, we have that either $J_{sf} \cap J_{s'f'} = \emptyset$, in which case R_{sfp} and $R_{s'f'p'}$ are disjoint, or $J_{s'f'} \subset J_{sf}$, in which case R_{sfp} and $R_{s'f'p'}$ overlap.

If J_{sf} and $J_{s'f'}$ are disjoint, then ψ_{sfp} and $\psi_{s'f'p'}$ come from different branches of the wavelet packet tree and can be embedded in a graph basis, hence must be orthogonal. Thus in all cases where R_{sfp} and $R_{s'f'p'}$ are disjoint, the functions ψsfp and $\psi_{s'f'p'}$ are orthogonal.

We now check the last remaining subcase, in which R_{sfp} and $R_{s'f'p'}$ overlap. But then, the function ψ_{sfp} is supported in one of the $2^{s'-s}$ adjacent subintervals of length 2^s contained in $I_{s'p'}$, and $\psi_{s'f'p'}$ is a direct descendent of ψ_{sfp} in the wavelet packet tree. But for the filters $h = \{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\}$ and $g = \{\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\}$, we have $\psi_{s'f'p'} = \pm 2^{(s-s')/2}\psi_{sfp}$ on each of those adjacent intervals. Thus the inner product in Equation 12 will be $2^{(s-s')/4} \neq 0$.

The normalization part of the theorem is free, since all Haar–Walsh wavelet packets have unit norm. □

Coifman and Meyer have remarked that the proof works with smooth time-frequency atoms as well, as long as we use Haar–Walsh filters to perform the frequency decompositions:

Corollary 1 *If ϕ is any time-frequency atom which is orthogonal to its integer translates, and X is the vector space spanned by $\{\phi(t-n) : 0 \leq n < 2^L\}$, then every disjoint tiling of the square $[0, 2^L[\times[0, 2^L[$ by dyadic rectangles corresponds to an orthonormal basis for X made of time-frequency atoms.* □

This corollary may be used to build plenty of smooth orthonormal bases in a smooth approximation space. We iterate longer orthogonal filters to produce smooth sampling functions ϕ of fixed scale, then use the Haar–Walsh filters to do a fixed finite number of frequency decompositions. We thus avoid using Haar–Walsh wavelet packets, which are not even time-frequency atoms since discontinuous functions have infinite frequency uncertainty. However, the Heisenberg product of these hybrid wavelet packets will blow up as the number of levels L of decomposition increases.

2 Frequency Spreading and Wavelet Packets

The fundamental idea of wavelet packet analysis is to construct a library of orthonormal bases for $L^2(\mathbf{R})$ which can be searched in real time for the best expansion with respect to a given application. The standard construction is to start from a multiresolution analysis and generate the library using the associated quadrature mirror filters. The internal structure of the MRA and the speed of the decomposition schemes make this an efficient adaptive method for simultaneous time and frequency analysis of signals. Unfortunately this standard construction of the library produces an unwanted spreading of the frequency localization of wavelet packets at high frequencies. We will here show how to control this spreading using nonstationary wavelet packet.

2.1 Nonstationary Wavelet Packets

Let (φ, ψ) be the scaling function and wavelet corresponding to a multiresolution analysis for $L^2(\mathbf{R})$ and (H, G) the associated quadrature mirror filters, see [1],[3],[9]. Furthermore let $(F_0^{(p)}, F_1^{(p)})$, $p \in \mathbf{N}$, be a family of bounded operators on $\ell^2(\mathbf{Z})$ of the form

$$(F_\varepsilon^{(p)} a)_k = \sum_{n \in \mathbf{Z}} a_n h_\varepsilon^{(p)}(n - 2k), \quad \varepsilon = 0, 1$$

with $h_1^{(p)}(n) = (-1)^n h_0^{(p)}(1 - n)$ real valued sequences in $\ell^1(\mathbf{Z})$ such that

$$\begin{aligned} F_0^{(p)*} F_0^{(p)} + F_1^{(p)*} F_1^{(p)} &= 1, \\ F_0^{(p)} F_1^{(p)*} &= 0. \end{aligned}$$

The family $\{w_n(\cdot - k)\}_{n \in \mathbf{N}, k \in \mathbf{Z}}$ of basic nonstationary wavelet packets is defined by letting $w_0 = \varphi$, $w_1 = \psi$ and then recursively for $n \in \mathbf{N}$:

$$w_{2n+\varepsilon}(t) = \sqrt{2} \sum_{k \in \mathbf{Z}} h_\varepsilon^{(p)}(k) w_n(2t - k) \quad \varepsilon = 0, 1$$

where $2^p \leq n < 2^{p+1}$. It has a simple relation to the multiresolution decomposition $L^2(\mathbf{R}) = V_0 \oplus_{j \geq 0} W_j$ corresponding to (φ, ψ): $\{w_0(\cdot - k)\}_{k \in \mathbf{Z}}$ is an orthonormal basis for V_0, $\{w_n(\cdot - k)\}_{k \in \mathbf{Z}, 2^j \leq n < 2^{j+1}}$ an orthonormal basis for W_j and in particular the family of basic nonstationary wavelet packets constitute an orthonormal basis for $L^2(\mathbf{R})$. The full library of nonstationary wavelet packet bases for $L^2(\mathbf{R})$ is obtained using the following result:

Theorem 6 *For every partition P of $\mathbf{N}$ into sets of the form $I_{jn} = \{n2^j, \ldots, (n+1)2^j - 1\}$ with $n, j \in \mathbf{N}$, the collection of functions: $\{2^{j/2} w_n(2^j \cdot - k)\}_{k \in \mathbf{Z}, I_{jn} \in P}$ form an orthonormal basis for $L^2(\mathbf{R})$.*

The input signal to the discrete wavelet packet algorithm is the projection of a signal $f \in L^2(\mathbf{R})$ onto V_N, $P_{V_N} f = \sum_{k \in \mathbf{Z}} c_k^N \varphi_{Nk}$, and the idea is to extract the information corresponding to each of the possible wavelet packet decompositions of V_N:

$$\begin{aligned} V_N &= \oplus_{j,n} \Omega_{j,n}, \\ \Omega_{j,n} &= \overline{\text{span}}_{k \in \mathbf{Z}} \{2^{\frac{j}{2}} w_n(2^j \cdot - k)\}. \end{aligned}$$

One of the possible decompositions is the wavelet decomposition, but if a better frequency resolution is needed for high frequencies, other wavelet packet decompositions can be used. Any wavelet packet decomposition of V_N into a direct orthogonal sum will only involve the functions:

$$2^{j/2}w_n(2^j \cdot -k), \quad k \in \mathbf{Z}, 0 \le j \le N, 0 \le n < 2^{N-j},$$

and the algorithm therefore needs to compute the inner product of f with these functions. Define $f_k^{n,j} = \langle f, 2^{j/2}w_n(2^j \cdot -k)\rangle$. The recipe for calculation of the nonstationary wavelet packet coefficients is given by:

Theorem 7 *Assume $k \in \mathbf{Z}, 0 \le j \le N, 0 \le n < 2^{N-j}$. For $n \ge 1$ use the binary expansion: $[n]_2 = \epsilon_1\epsilon_2\cdots\epsilon_M$ when $n = \sum_{j=1}^{M}\epsilon_j 2^{j-1}$ with $\epsilon_j = 0$ or 1 and $\epsilon_M = 1$. Then the following equations hold:*

$$\begin{aligned} f_k^{0,j} &= [(H)^{N-j}c^N]_k, \\ f_k^{n,j} &= [F_{\epsilon_1}^{(M-1)}\cdots F_{\epsilon_{M-1}}^{(1)}G(H)^{N-(j+M)}c^N]_k, \quad n \ge 1. \end{aligned}$$

Except for notational changes the proofs of the results in this section follow rather closely those of the standard case, see [2]. The discrete wavelet packet algorithm can be represented by a simple binary tree and the "best-basis algorithm" applied to pick the "best-basis" according to predetermined criteria; see [12]. Notice that the numerical complexity of this discrete wavelet packet algorithm will depend on the possible growth of filter length of the operators $(F_0^{(p)}, F_1^{(p)})$ as the level p tends to infinity. If for instance we have polynomial growth in p, then the numerical complexity of the algorithm will be $O(N[\log N]^r)$ for some power r.

2.2 Frequency Localization

One way for estimating the frequency localization of the basic wavelet packets $w_n(x)$ is to compute the variance

$$\sigma_n = \inf_{\xi_0 \in R}\int_0^\infty |\xi - \xi_0|^2 |\hat{w}_n(\xi)|^2 d\xi.$$

This choice was considered in [2]. The result obtained was that for a family of Meyer wavelets σ_n would have an average growth as n^δ for some $\delta > 0$. More precisely it was shown that there exist constants $C > 0$ and $r > 1$ such that $2^{-J}\sum_{2^J \le n < 2^{J+1}}\sqrt{1+\sigma_n} \ge C \cdot r^J$. This growth turns out to be an inavoidable consequence of the standard wavelet packet construction.

Theorem 8 *Let $\{w_n\}_{n\in\mathbf{N}}$ be the basic wavelet packets generated from a wavelet basis using its associated quadrature mirror filters. Assume that σ_n is well defined, that $|\hat{\psi}(\xi)| \ge \delta_0 > 0$ on an open set $\mathcal{I}$ and that $0 < |m_0(\xi)| < 1$ on a set $\mathcal{M}$ of positive Lebesgue measure. Then there exists constants $C > 0$ and $r > 1$ such that $2^{-J}\sum_{2^J \le n < 2^{J+1}}\sqrt{1+2\sigma_n} \ge C \cdot r^J$ and consequently σ_n has an average growth as n^δ for some $\delta > 0$.*

The conditions of the above theorem are very mild and in particular satisfied by any wavelet with a continuous Fourier transform. Assuming that δ is close to zero we may have

a slow average growth of σ_n. It is then interesting to know whether all σ_n within each scale have a comparable behaviour or if some are much worse behaved. Unfortunately this turns out to be the case:

Theorem 9 *Let $\{w_n\}_{n\in\mathbf{N}}$ be the family of basic nonstationary wavelet packets. Assume that σ_n is well defined and that $|\hat{\psi}(\xi)| \geq \delta_0 > 0$ on $[k\pi - \varepsilon; k\pi + \varepsilon]$ for some $\varepsilon > 0$ and $k \in \mathbf{Z}$. Then within each scale there exists an n such that $\sigma_n \geq 16^{-1}\varepsilon^3\delta_0^2 \cdot n^2$.*

Any reasonably localized wavelet ψ fulfills the conditions of theorem 9, since all such functions satisfy the condition $\sum_{k\in\mathbf{Z}} |\hat{\psi}(\xi+2\pi k)|^2 = 1$ with uniform convergence on compact sets, see [9]. In particular they are satisfied by the compactly supported Daubechies wavelets. For the proofs of the above theorems see [7].

The conditions of the wavelet are not fulfilled by the Shannon wavelet due to its perfect frequency localization. Indeed for Shannon wavelet packets σ_n does not depend on n. However, the theorem tells us that even if we use the sharp cutoff quadrature mirror filters of the Shannon wavelet to generate nonstationary wavelet packets from a wavelet that does fulfill the conditions, we cannot avoid having very badly behaved basic wavelet packets with respect to the measure σ_n. This indicates that the measure σ_n might not be the right choice for estimating frequency localization of wavelet packets.

When using the discrete wavelet packet algorithm the coefficients are always interpreted in the idealized time-frequency plane as if they correspond to the sharp frequency bands of Shannon wavelet packets. Once some energy of $\hat{w}_n$ is outside the corresponding frequency band it will spread away from the central frequency bands in the following iterations of the operators. The result is that for the frequency localization in the limit, energy close to but outside the central frequency band is just as unimportant as energy far away. An appropriate way to estimate the frequency localization of basic wavelet packets is therefore to choose the following definition:

Definition 1 *Let $p : \mathbf{N} \to \mathbf{N}$ be the permutation of the nonnegative integers defined by $p(\epsilon_1 \cdots \epsilon_M) = \epsilon_1^* \cdots \epsilon_M^*$, $\epsilon_i^* = \sum_{l=0}^{\infty} \epsilon_{i+l} \bmod 2$, where $[n]_2 = \epsilon_1 \cdots \epsilon_M$ denotes the binary expansion of $n = \sum_{j=1}^{M} \epsilon_j 2^{j-1}$ with $\epsilon_j = 0$ or 1. The frequency localization of a wavelet packet w_n is estimated by calculating*

$$\eta_n = \int_{-\infty}^{\infty} |\hat{w}_n(\xi)|^2 \chi_{[p(n)\pi;(p(n)+1)\pi]}(|\xi|) d\xi.$$

The wavelet packet is said to have a good frequency localization if η_n is large.

The intervals of integration equal the central frequency bands of Shannon wavelet packets, and n is the Gray-code permutation of $p(n)$. The notion "large" can be interpreted in different ways. Ideally we would like to have η_n close to 1, but in order to get a better frequency localization for high frequencies than the wavelet transform it suffices to have η_n large compared to n^{-1}.

2.3 Filter Design

The QMF's corresponding to high regularity wavelets (φ, ψ) are constructed by controlling the decay of the infinite products of the associated trigonometric polynomials. In this section we will construct a family of quadrature mirror filters directly for the purpose of generating basic nonstationary wavelet packets with a good frequency localization. Neglecting the already occured error to level p we optimize the transition $w_n \to w_{2n+\varepsilon}$, $2^p \le n < 2^{p+1}, \varepsilon = 0, 1$, in the nonstationary wavelet packet construction with respect to the frequency localization measure η by minimizing

$$\|\chi_{[-\frac{\pi}{2};\frac{\pi}{2}]} - |\frac{1}{\sqrt{2}} \sum_{k \in \mathbf{Z}} h_0^{(p)}(k) e^{ik\xi}|^2 \|_{L^1([-\pi;\pi])}$$

under the constraint that $(F_0^{(p)}, F_1^{(p)})$ are exact reconstruction quadrature mirror filters of a given length.

Consider the trigonometric polynomial $m_0(\xi) = 2^{-1/2} \sum_{k=0}^{N} h_k e^{ik\xi}$. The exact reconstruction condition on the corresponding filters (F_0, F_1) can be written using the filter coefficients as: $\sum_k h_k h_{k+2l} = \delta_{0,l}$, see [3]. In particular N must be odd. Introducing the notation $a_l = \sum_k h_k h_{k+2l+1}$ a direct calculation then shows that

$$|m_0(\xi)|^2 = \frac{1}{2} + \sum_{l=0}^{\frac{N-1}{2}} a_l \cos((2l+1)\xi).$$

The only constraints on the coefficients a_l are that they are real and that the above trigonometric polynomial is positive. A corresponding $m_0(\xi)$ with real coefficients can then be constructed using Riesz factorization, see [11].

Inserting this expression for $|m_0(\xi)|^2$ we see that we need to solve the following minimization problem: Minimize

$$\pi - 4 \sum_{l=0}^{\frac{N-1}{2}} \frac{(-1)^l a_l}{2l+1}$$

over

$$\underline{a} = (a_0, a_1, \ldots, a_{\frac{N-1}{2}}) \in \mathbf{R}^{\frac{N+1}{2}}$$

under the constraint that for all $\xi \in \mathbf{R}$,

$$\frac{1}{2} + \sum_{l=0}^{\frac{N-1}{2}} a_l \cos((2l+1)\xi) \ge 0.$$

This minimization problem is complicated due to the peculiar pointwise constraint. It turns out, however, that it is possible to give exact formulas for the minimizing trigonometric polynomials at each filter length, see [6].

To obtain the solutions we use the following result:

Lemma 1 *Let $N = 2M + P \geq 0$ where $M \in \mathbf{N}$ and $P = 0$ or -1. Consider the system of equations*

$$\begin{pmatrix} 1 \\ -\frac{1}{3} \\ \vdots \\ \frac{(-1)^N}{2N+1} \end{pmatrix} = b_0(1+P) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \sum_{l=1}^{M} t_l \begin{pmatrix} \cos(\xi_l) \\ \cos(3\xi_l) \\ \vdots \\ \cos((2N+1)\xi_l) \end{pmatrix}.$$

These systems of equations all have a unique solution $(b_0, \underline{t}, \underline{\xi})$ or $(\underline{t}, \underline{\xi})$ fulfilling $\frac{\pi}{2} < \xi_1 < \xi_2 < \cdots < \xi_M < \pi$.

For each $N \in \mathbf{N}$ we let $\underline{\xi} = (\xi_1, \ldots, \xi_M)$ be the values from the solution of the above associated system of equations and introduce

$$A^{(2M)}(\xi_1, \ldots, \xi_M) = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \cos(\xi_1) & \cos(3\xi_1) & \ldots & \cos((4M+1)\xi_1) \\ \vdots & \vdots & & \vdots \\ \cos(\xi_M) & \cos(3\xi_M) & \ldots & \cos((4M+1)\xi_M) \\ \sin(\xi_1) & 3\sin(3\xi_1) & \ldots & (4M+1)\sin((4M+1)\xi_1) \\ \vdots & \vdots & & \vdots \\ \sin(\xi_M) & 3\sin(3\xi_M) & \ldots & (4M+1)\sin((4M+1)\xi_M) \end{bmatrix},$$

$$\underline{b}^{(2M)} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \vdots \\ -\frac{1}{2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Similarly $A^{(2M-1)}(\xi_1, \ldots, \xi_M)$ is introduced as the same matrix but without the first row and last column and $\underline{b}^{(2M-1)}$ as the same vector but without the first entry.

The coefficients $\underline{a}$ of the unique minimizing trigonometric polynomial $|m_0^{(N)}(\xi)|^2 = 1/2 + \sum_{l=0}^{N} a_l \cos((2l+1)\xi)$ are then given by the simple expression

$$\underline{a} = [A^{(N)}(\underline{\xi})]^{-1} \underline{b}^{(N)}.$$

For the filter lengths 2,4,...,10 exact results can be seen in [7]. For the coefficients of the corresponding operators (F_0, F_1) see Table 1.

2.4 Control of Frequency Spreading

We want to use the nonstationary wavelet packet construction to control the frequency spreading at high frequencies. More precisely we want to construct a family of nonstationary wavelet packets such that for all $J \in \mathbf{N}$,

$$2^{-J} \sum_{2^J \leq n < 2^{J+1}} \eta_n \geq C > 0.$$

$h_0 =$	$\frac{1}{\sqrt{2}}$	$h_1 =$	$\frac{1}{\sqrt{2}}$
$h_0 =$	$\frac{1}{8}\sqrt{12+3\sqrt{10}}$	$h_1 =$	$\frac{1}{8}\sqrt{20+5\sqrt{10}}$
$h_2 =$	$\frac{1}{8}\sqrt{20-5\sqrt{10}}$	$h_3 =$	$-\frac{1}{8}\sqrt{12-3\sqrt{10}}$
$h_0 =$	0.503552845382	$h_1 =$	0.749282656009
$h_2 =$	0.358669023827	$h_3 =$	−0.146420439018
$h_4 =$	−0.155115088022	$h_5 =$	0.104244564196
$h_0 =$	0.451615929782	$h_1 =$	0.739818187277
$h_2 =$	0.422992416990	$h_3 =$	−0.102215586701
$h_4 =$	−0.200582571342	$h_5 =$	0.050352473139
$h_6 =$	0.109880116973	$h_7 =$	−0.067075413993
$h_0 =$	0.413126295869	$h_1 =$	0.726992926835
$h_2 =$	0.468173223166	$h_3 =$	−0.062992868600
$h_4 =$	−0.223538187273	$h_5 =$	0.012777578250
$h_6 =$	0.133546841599	$h_7 =$	−0.017519751720
$h_8 =$	−0.084201392174	$h_9 =$	0.047848896422

Table 1: QMF Coefficients $\underline{h}$.

If we use the same pair of FIR exact reconstruction quadrature mirror filters on each scale the above average value of η_n on scale J tends to 0 as J tends to infinity. The goal can therefore only be achieved by increasing the filter length from scale to scale. We have the following:

Theorem 10 *Consider basic wavelet packets generated from a wavelet which fulfills* $|\hat{\psi}(\xi)| \geq \delta > 0$ *on* $[a-\varepsilon; a+\varepsilon] \subseteq [\pi; 2\pi]$ *for some constant* $\varepsilon > 0$ *and point* $a \in]\pi; 2\pi[$. *Fix* $\gamma > 0$ *and a constant* $C_{\gamma,\varepsilon} > \frac{1}{\pi}(\frac{2+\gamma}{2\varepsilon\gamma})^2$, *and choose the minimizing QMF's introduced in the previous section with filter lengths* $n_p \geq C_{\gamma,\varepsilon} p^{2+\gamma} + 1$. *Then there exists a constant* $C > 0$ *such that for all* $J \in \mathbf{N}$,

$$2^{-J} \sum_{2^J \leq n < 2^{J+1}} \eta_n \geq C.$$

The proof of this theorem can be seen in [7]. It is valid not only for the minimizing family of QMF's but in fact for any family with a frequency localization better than or equal to the QMF's corresponding to the compactly supported Daubechies wavelets. It shows that the frequency spreading can be controlled by using different QMF's on each scale and allowing the filter lengths to grow. Furthermore it provides an upper bound to the necessary growth for which the resulting numerical complexity is $O(N[\log(N)]^{2+\gamma})$ for any $\gamma > 0$.

3 Periodic Wavelet Packet Analysis

Wavelet packets have been introduced as a flexible method for time-frequency analysis of signals combining the advantages of windowed Fourier and wavelet analysis. Similarly periodic wavelet packets provide an interesting alternative to Fourier series. In this section

we study the periodic wavelet packet equivalent of Fourier series and show that the basic periodic wavelet packets fulfill a fundamental translation property similar to that of sines and cosines. As an example we consider periodic wavelet packets in the optimal frequency localization limit. This limit corresponds to a family of periodic Meyer wavelet packets and provides an alternative FFT.

3.1 Periodic Wavelet Packets

By periodizing the basis functions to period 1 an MRA for $L^2(\mathbf{R})$ is transformed into an MRA for $L^2(0,1)$; see [9]. This construction extends to the wavelet packet case as well. Let $\{w_n\}_{n\in\mathbf{N}}$ be the family of (possibly) nonstationary basic wavelet packets introduced in the previous section. For $n, j \in \mathbf{N}$ and $0 \le k < 2^j$ define general periodic wavelet packets $w^{per}_{n,j,k}$ by

$$w^{per}_{n,j,k}(x) = \sum_{l\in\mathbf{Z}} 2^{j/2} w_n(2^j(x+l)-k).$$

Using this notation we have:

Theorem 11 *For every partition P of $\mathbf{N}$ into sets of the form $I_{jn} = \{n2^j, \dots, (n+1)2^j - 1\}$ with $n, j \in \mathbf{N}$, the collection of functions:*

$$\{w^{per}_{n,j,k}\}_{0\le k<2^j, I_{jn}\in P}$$

is an orthonormal basis for $L^2(0,1)$.

The input signal to the discrete periodic wavelet packet algorithm is the projection of a signal $f \in L^2(0,1)$ onto V^{per}_N, $P_{V^{per}_N} f = \sum_{0\le k<2^N} c^N_k \varphi^{per}_{Nk}$, and the idea again is to extract the information corresponding to each of the possible periodic wavelet packet decompositions of V^{per}_N:

$$V^{per}_N = \oplus_{j,n} \Omega^{per}_{j,n}, \quad \Omega^{per}_{j,n} = \mathrm{span}_{0\le k<2^j}\{w^{per}_{n,j,k}\}.$$

Define $f^{n,j}_k = \langle f, w^{per}_{n,j,k}\rangle$ and consider it to be a 2^j-periodic sequence in k. Furthermore introduce the notation $(F^{(p)}_{\varepsilon,j})_k = \sum_{s\in\mathbf{Z}} (F^{(p)}_{\varepsilon})_{k+2^j s}$, $\quad p, j \in \mathbf{N}, k \in \mathbf{Z}, \varepsilon = 0, 1$ and the correspondent operators

$$(F^{(p)}_{\varepsilon,j} a)_k = \sum_{n=0}^{2^j-1} a_n (F^{(p)}_{\varepsilon,j})_{n-2k}$$

which transforms sequences of period 2^j into sequences of period 2^{j-1}. Similarly denote by (H_j, G_j) the periodized operators of (H, G). The recipe for calculation of the periodic wavelet packet coefficients is then given by:

Theorem 12 *Assume $k \in \mathbf{Z}, 0 \le j \le N, 0 \le n < 2^{N-j}$. For $n \ge 1$ use the binary expansion $[n]_2 = \epsilon_M \epsilon_{M-1} \cdots \epsilon_1$, where $n = \sum_{j=1}^M \epsilon_j 2^{j-1}$ with $\epsilon_j = 0$ or 1 and $\epsilon_M = 1$. Then the following hold:*

$$\begin{aligned} f^{0,j}_k &= (H_{j+1} H_{j+2} \cdots H_N c^N)_k \\ f^{n,j}_k &= (F^{(M-1)}_{\epsilon_1,j+1} \cdots F^{(1)}_{\epsilon_{M-1},j+M-1} G_{j+M} H_{j+M+1} \cdots H_N c^N)_k, \quad n \ge 1. \end{aligned}$$

Except for notational changes the proofs of these results follow the nonperiodic case and the same arguments about representation as a binary tree, the best-basis search and the numerical complexity applies.

3.2 A Translation Property

We start by showing a translation result involving only a particular subset of the basic periodic wavelet packets. We have:

Lemma 2 *For $n \in \mathbf{N}$: $w^{per}_{4n+3,0,0}(x - \frac{1}{4}) = w^{per}_{4n+1,0,0}(x)$.*

Proof: Consider first the case $n \geq 1$. By definition we have

$$\begin{aligned} w^{per}_{4n+1,0,0}(x) &= 2 \sum_{k,l \in \mathbf{Z}} (F_1^{(p+1)})_k (F_0^{(p)})_l \sum_{r \in \mathbf{Z}} w_n(4x - 2k - l + 4r), \\ w^{per}_{4n+3,0,0}(x) &= 2 \sum_{k,l \in \mathbf{Z}} (F_1^{(p+1)})_k (F_1^{(p)})_l \sum_{r \in \mathbf{Z}} w_n(4x - 2k - l + 4r). \end{aligned}$$

The proof consists in rewriting these expressions using the properties of the operators $(F_0^{(p)}, F_1^{(p)})$. First we split each of the expressions into four parts depending on the value of $(2k + l) \bmod 4$. Let $w_n^{(s)}(x) = \sum_{r \in \mathbf{Z}} w_n(4x + 4r - s)$, $s = 0, \ldots, 3$, then

$$w^{per}_{4n+1,0,0}(x) = \sum_{s=0}^{3} 2 \sum_{k \in \mathbf{Z}} (F_1^{(p+1)})_k \sum_{r \in \mathbf{Z}} (F_0^{(p)})_{s+4r-2k} w_n^{(s)}(x).$$

Introducing $a_s = \sqrt{2} \sum_{r \in \mathbf{Z}} (F_0^{(p)})_{s+4r}$, $s = 0, \ldots, 3$, and using the fact that $\sum_{k \in \mathbf{Z}} (F_1^{(p+1)})_{2k} = 1/\sqrt{2} = -\sum_{k \in \mathbf{Z}} (F_1^{(p+1)})_{2k+1}$ (see [3]) we can simplify the above expression. By splitting the sums into two parts, k even and k odd, we obtain:

$$w^{per}_{4n+1,0,0}(x) = \sum_{s=0}^{3} (a_s - a_{s+2 \bmod 4} w_n^{(s)}(x).$$

Applying the same procedure to $w^{per}_{4n+3,0,0}$ and using $(F_1^{(p)})_k = (-1)^k (F_0^{(p)})_{1-k}$ we similarly obtain:

$$w^{per}_{4n+3,0,0}(x) = \sum_{s=0}^{3} (a_{(s+1) \bmod 4} - a_{(s+3) \bmod 4}) w_n^{(s)}(x).$$

Since $w_n^{(s)}(x - \frac{1}{4}) = w_n^{((s+1) \bmod 4)}(x)$, $s = 0, \ldots, 3$, the result now follows.

The case $n = 0$ looks slightly different if $F_1^{(1)}$ has been chosen different from G:

$$\begin{aligned} w_1(x) &= 2 \sum_{k,l \in \mathbf{Z}} (G)_k (H)_l w_0(4x - 2k - l), \\ w_3(x) &= 2 \sum_{k,l \in \mathbf{Z}} (F_1^{(1)})_k (G)_l w_0(4x - 2k - l). \end{aligned}$$

However, since $\sqrt{2}\sum_{k\in\mathbf{Z}}(F)_{2k} = 1 = -\sqrt{2}\sum_{k\in\mathbf{Z}}(F)_{2k+1}$ is valid for $F \in \{G, F_1^{(1)}\}$ and (H, G) like $(F_0^{(p)}, F_1^{(p)})$ fulfill $(G)_k = (-1)^k(H)_{1-k}$ the above calculations extend to this case as well. This completes the proof. □

The present ordering of the basic wavelet packets $\{w_n\}_{n\in\mathbf{N}}$ is convenient with respect to the recursive formulas in their definition. However, in many applications it is more appropriate to have them ordered according to increasing frequency. Let $G(n)$ be the Gray code permutation of n, then $\{\tilde{w}_n^{per}\}_{n\in\mathbf{N}}$ defined by

$$\tilde{w}_n^{per} = w_{G(n),0,0}^{per}$$

are the basic periodic wavelet packets ordered according to increasing frequency. The Gray code permutation also relates Paley order to sequency order for Walsh functions, and it is therefore natural that it appears in the frequency localization of wavelet packets which generalize them.

For $n \in \mathbf{N}$ use the binary expansion $2n = \sum_{l=k(n)}^{\infty} \varepsilon_l(n)2^l$, $1 \le k(n) < \infty$ with $\varepsilon_{k(n)} = 1$ and $\varepsilon_l \in \{0, 1\}$ otherwise. Define a family of constants η_n by

$$\eta_n = 1 - 2^{-(k(n)+1)} + \varepsilon_{k(n)+1}(n) \cdot (2^{-k(n)} - 1), \quad n \in \mathbf{N}.$$

We then have the following fundamental translation property:

Theorem 13 *For $n \in \mathbf{N}$: $\tilde{w}_{2n}^{per}(x) = \tilde{w}_{2n-1}^{per}(x - \eta_n)$.*

Proof: The binary expansion of $2n$ is of the form

$$[2n]_2 = x_1 \cdots x_p \, \varepsilon_{k(n)+1}(n) \, 1 \, 0 \cdots 0,$$

where each x_i denote an unspecified coefficient and where the first coefficient equal to 1 comes after $k(n) \ge 1$ zeros. Then

$$[2n-1]_2 = x_1 \cdots x_p \, \varepsilon_{k(n)+1}(n) \, 0 \, 1 \cdots 1,$$

and

$$\begin{aligned} [G(2n)]_2 &= y_1 \cdots y_p \, [(\varepsilon_{k(n)+1} + 1) \bmod 2] \, 1 \, 0 \cdots 0, \\ [G(2n-1)]_2 &= y_1 \cdots y_p \, \varepsilon_{k(n)+1} \, 1 \, 0 \cdots 0, \end{aligned}$$

where as before y_i are some unspecified coefficients and where the first coefficient equal to 1 now comes after $k(n) - 1 \ge 0$ zeros. Let a constant $M \in \mathbf{N}$ be defined by $[M]_2 = y_1 \cdots y_p$. We can then write

$$\begin{aligned} G(2n) &= 2^{k(n)-1}(4M + 1 + 2[(\varepsilon_{k(n)+1} + 1) \bmod 2]), \\ G(2n-1) &= 2^{k(n)-1}(4M + 1 + 2\varepsilon_{k(n)+1}). \end{aligned}$$

Using $w_{2^{k(n)-1}(4M+1)}(x) = w_{2^{k(n)-1}(4M+3)}(x - 2^{-(k(n)+1)})$, which follows from lemma 2 and the definition of the basic wavelet packets, the proof is completed. □

This theorem shows a strong resemblance between the usual Fourier series

$$f = c_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi n x) + \sum_{n=1}^{\infty} b_n \sin(2\pi n x)$$

and their periodic wavelet packet equivalent

$$f = \tilde{c}_0 \tilde{w}_0^{per} + \sum_{n=1}^{\infty} \tilde{a}_n \tilde{w}_{2n}^{per} + \sum_{n=1}^{\infty} \tilde{b}_n \tilde{w}_{2n-1}^{per}.$$

The even-sequency periodic wavelet packets $\tilde{w}_{2n}^{per}$ play the role of cosines and the odd-sequency periodic wavelet packets $\tilde{w}_{2n-1}^{per}$ act like sines, in the sense that for each n these functions are just translates of one another. With the added twist that the amopunt of the translate depends on n, this theorem shows how wavelet packets are generalizations of sine and cosine Walsh functions.

In the optimal frequency localization case, the basic periodic wavelet packets are in fact cosines and sines. The discrete periodic wavelet packet algorithm therefore provides an alternative FFT.

3.3 Optimal Frequency Localization

Let us introduce a family of Meyer wavelets. Choose a constant ε, $0 < \varepsilon \leq \frac{\pi}{3}$, and a function ν: $\mathbf{R} \to \mathbf{R}$ in C^k or C^∞ satisfying

$$\nu(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x \geq \frac{2\varepsilon}{\pi - \varepsilon} \end{cases}$$

with the additional property $\nu(x) + \nu(-x + 2\varepsilon/(\pi - \varepsilon)) = 1$. Define a 2π-periodic even function $m_0(\xi)$ by

$$m_0(\xi) = \cos(\frac{\pi}{2} \nu(\frac{2|\xi|}{\pi - \varepsilon} - 1)), \quad -\pi \leq \xi \leq \pi.$$

This function satisfies the QMF condition $|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1$ and furthermore $m_0(\xi) = 1$, $-\pi/3 \leq \xi \leq \pi/3$, which is sufficient for it to generate an MRA, see [1]. The corresponding scaling function φ and wavelet ψ are given by

$$\begin{aligned} \hat{\varphi}(\xi) &= \frac{1}{\sqrt{2\pi}} \prod_{l=1}^{\infty} m_0(\frac{\xi}{2^l}) = \frac{1}{\sqrt{2\pi}} \cos(\frac{\pi}{2} \nu(\frac{|\xi|}{\pi - \varepsilon} - 1)), \\ \hat{\psi}(2\xi) &= -e^{-i\xi} \overline{m_0(\xi + \pi)} \hat{\varphi}(\xi). \end{aligned}$$

These functions $(\varphi_{\nu,\varepsilon}, \psi_{\nu,\varepsilon})$ constitute the family of Meyer wavelets.

For $n \in \mathbf{N}$ use the binary expansion $2n = \sum_{l=0}^{\infty} \varepsilon_l(n) 2^l$, $\varepsilon_l(n) \in \{0, 1\}$ and introduce a family of constants $\{\kappa_n\}_{n \in \mathbf{N}}$ by

$$\kappa_n = \pi \sum_{l=0}^{\infty} [(\varepsilon_l(n) + \varepsilon_{l+1}(n)) \bmod 2](1 + \frac{2n}{2^{l+1}}).$$

We have the following result:

Theorem 14 *Let $(\varphi_{\nu,\varepsilon}, \psi_{\nu,\varepsilon})$ belong to the family of Meyer wavelets and assume that $\varepsilon \leq 2^{-N}$. Consider the periodized wavelet packets generated using the associated quadrature mirror filters. They fulfill the equations*

$$\begin{aligned} \tilde{w}_{2n}^{per}(x) &= \sqrt{2}\cos(2\pi n x - \kappa_n), \\ \tilde{w}_{2n-1}^{per}(x) &= \sqrt{2}\sin(2\pi n x - \kappa_n), \end{aligned}$$

for each n, $0 < n < 2^{N-1}$.

This result can be proven by calculating the coefficients of the Fourier series expansion $\tilde{w}_n^{per}(x) = \sqrt{2\pi}\sum_{k\in\mathbf{Z}} \hat{w}_{G(n)}(2\pi k)e^{i2\pi kx}$ using the properties of the Meyer wavelets.

4 Phase Uncertainty of Wavelet Packets

We wish to recognize features of the original signal from the coefficients produced by transformations involving QMFs, so it is necessary to keep track of which portion of the sequence contributes energy to the filtered sequence.

Suppose that F is a finitely supported filter with filter sequence $f(n)$. For any sequence $u \in \ell^2$, if $Fu(n)$ is large at some index $n \in \mathbf{Z}$, then we can conclude that $u(k)$ is large near the index $k = 2n$. Likewise, if $F^*u(n)$ is large, then there must be significant energy in $u(k)$ near $k = n/2$. We can quantify this assertion of nearness using the support of f, or more generally by computing the position of f and its uncertainty with Equation 1. When the support of f is large, the position method gives a more precise notion of where the analyzed function is concentrated.

Consider what happens when $f(n)$ is concentrated near $n = 2T$:

$$Fu(n) = \sum_{j\in\mathbf{Z}} f(j)u(2n-j) = \sum_{j\in\mathbf{Z}} f(j+2T))u(2n-j-2T). \tag{13}$$

Since $f(j+2T)$ is concentrated about $j = 0$, we can conclude by our previous reasoning that if $Fu(n)$ is large, then $u(k)$ is large when $k \approx 2n - 2T$. Similarly,

$$F^*u(n) = \sum_{j\in\mathbf{Z}} \bar{f}(2j-n)u(j) = \sum_{j\in\mathbf{Z}} \bar{f}(2j-n+2T))u(j+T). \tag{14}$$

Since $\bar{f}(2j-n+2T)$ is concentrated about $2j - n = 0$, we conclude that if $F^*u(n)$ is big then $u(j+T)$ must be big where $j \approx n/2$, which implies that $u(k)$ is big when $k \approx \frac{n}{2} + T$.

Decimation by 2 and its adjoint respectively cause the doubling and halving of the indices n to get the locations where u must be large. The translation by T or $-2T$ can be considered a "shift" induced by the filter convolution. We can precisely quantify the location of portions of a signal, measure the shift, and correct for it when interpreting the coefficients produced by applications of F and F^*. We will see that nonsymmetric filters might shift different signals by different amounts, with a variation that can be estimated by a simple expression in the filter coefficients. The details of the shift will be called the *phase response* of the filter.

4.1 Shifts for Sequences

The notion of position for a sequence may be defined by the following formula:

$$c[u] \stackrel{\text{def}}{=} \frac{1}{\|u\|^2} \sum_{k \in \mathbf{Z}} k|u(k)|^2. \tag{15}$$

This quantity, whenever it is finite, may also be called the *center of energy* of the sequence $u \in \ell^2$ to distinguish it from the function case.

The center of energy is the first moment of the probability distribution function (or *pdf*) defined by $|u(n)|^2/\|u\|^2$. We will say that the sequence u is *well-localized* if the second moment of that pdf also exists, namely if

$$\sum_{k \in \mathbf{Z}} k^2|u(k)|^2 = \|ku\|^2 < \infty. \tag{16}$$

A finite second moment insures that the first moment is also finite, by the Cauchy–Schwarz inequality:

$$\sum_{k \in \mathbf{Z}} k|u(k)|^2 = \langle ku, u\rangle \le \|ku\|\,\|u\| < \infty.$$

If $u \in \ell^2$ is a finitely supported sequence (say in the interval $[a, b]$) then $a \le c[u] \le b$.

Another way of writing $c[u]$ is in Dirac's *bra and ket* notation:

$$\|u\|^2 c[u] = \langle u|X|u\rangle \stackrel{\text{def}}{=} \langle u, Xu\rangle = \sum_{i \in \mathbf{Z}} \bar{u}(i)X(i,j)u(j), \tag{17}$$

where

$$\begin{aligned} X(i,j) &\stackrel{\text{def}}{=} i\delta(i-j) = \begin{cases} i, & \text{if } i = j, \\ 0, & \text{if } i \ne j, \end{cases} \\ &= \operatorname{diag}[\ldots, -2, -1, 0, 1, 2, 3, \ldots]. \end{aligned}$$

To simplify the formulas, we will always suppose that $\|u\| = 1$. We can also suppose that f is an orthogonal QMF, so $\sum_k \bar{f}(k)f(k+2j) = \delta(j)$. Then $FF^* = I$, F^* is an isometry and F^*F is an orthogonal projection. Since $\|F^*u\| = \|u\| = 1$, we can compute the center of energy of F^*u as $c[F^*u] = \langle F^*u|X|F^*u\rangle = \langle u|FXF^*|u\rangle$. We will call the the double sequence FXF^* between the bra and the ket the *phase response* of the adjoint convolution-decimation operator F^* defined by the filter sequence f. Namely,

$$FXF^*(i,j) = \sum_k kf(2i-k)\bar{f}(2j-k). \tag{18}$$

Now

$$FXF^*(i,j) = \sum_k ([i{+}j]+k)f([i{-}j]-k)\bar{f}([j{-}i]-k) \stackrel{\text{def}}{=} 2X(i,j) - C_f(i,j).$$

Here $2X(i,j) = (i+j)\sum_k f([i{-}j]-k)\bar{f}([j{-}i]-k) = 2i\delta(i-j)$ as above, since f is an orthogonal QMF, while

$$C_f(i,j) \stackrel{\text{def}}{=} \sum_k kf(k-[i{-}j])\bar{f}(k-[j{-}i]). \tag{19}$$

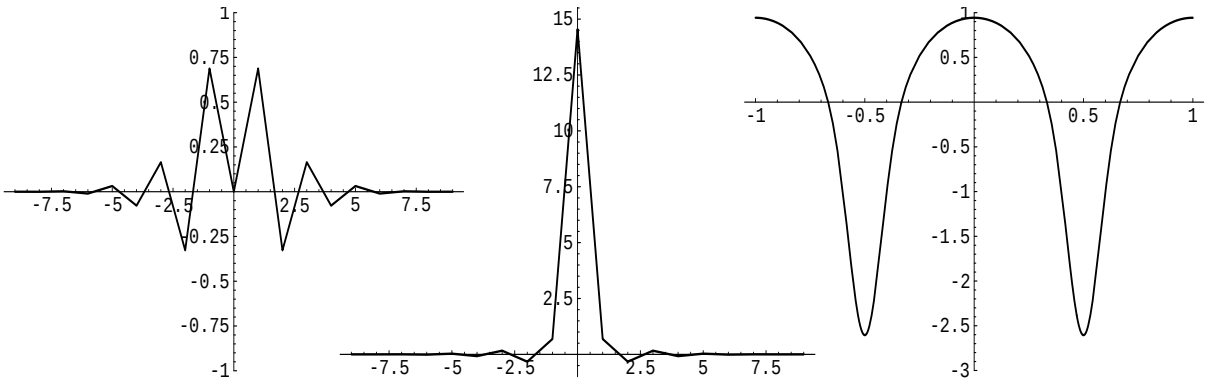

Figure 9: γ^0, γ, and $\hat{\gamma}^0$ for "Beylkin 18" high-pass QMF.

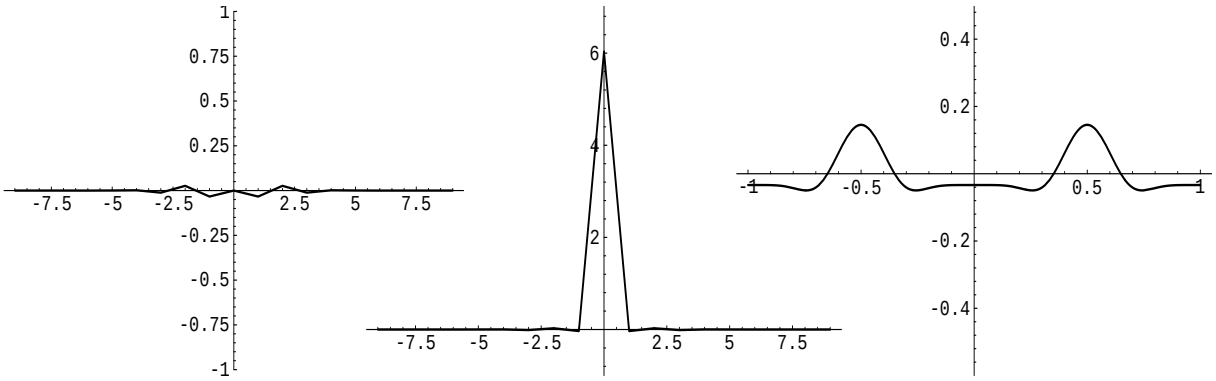

Figure 10: γ^0, γ, and $\hat{\gamma}^0$ for "Coiflet 18" low-pass QMF.

Thus $c[F^*u] = 2c[u] - \langle u|C_f|u\rangle$. C_f is evidently a convolution matrix: $C_f(i,j) = \gamma(i-j)$ so that $C_f u = \gamma * u$. The function γ is defined by the following formula:

$$\gamma(n) \stackrel{\text{def}}{=} \sum_k k f(k-n)\bar{f}(k+n). \tag{20}$$

From this formula it is easy to see that $\gamma(n) = \bar{\gamma}(-n)$, thus $\hat{\gamma}(\xi) = \hat{\bar{\gamma}}(-\xi) = \overline{\hat{\gamma}}(\xi) \Rightarrow \hat{\gamma} \in \mathbf{R}$. This symmetry of γ makes the matrix C_f selfadjoint. Along its main diagonal, $C_f(i,i) = \gamma(0) = c[f]$. Other diagonals of C_f are constant, and if f is supported in the finite interval $[a,b]$, then $C_f(i,j) = \gamma(i-j) = 0$ for $|i-j| > |b-a|$.

We can subtract the diagonal from C_f by writing $C_f = C_f^0 + c[f]I$, which is the same as the decomposition $\gamma(n) = \gamma^0(n) + c[f]\delta(n)$. This gives a decomposition of the phase response matrix:

$$FXF^* = 2X - c[f]I - C_f^0.$$

Thus FXF^* is multiplication by the linear function $2x - c[f]$ minus convolution with γ^0. We will say that f has a *linear phase response* if $\gamma^0 \equiv 0$.

Proposition 1 *Suppose that* $f = \{f(n) : n \in \mathbf{Z}\}$ *satisfies* $\sum_k \bar{f}(k-n)f(k+n) = \delta(n)$ *for*

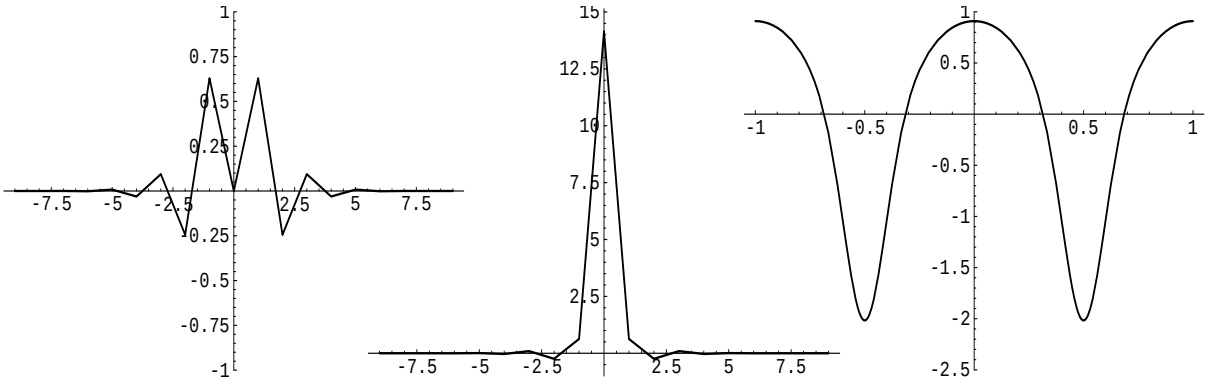

Figure 11: γ^0, γ, and $\hat{\gamma}^0$ for "Daubechies 18" high-pass QMF.

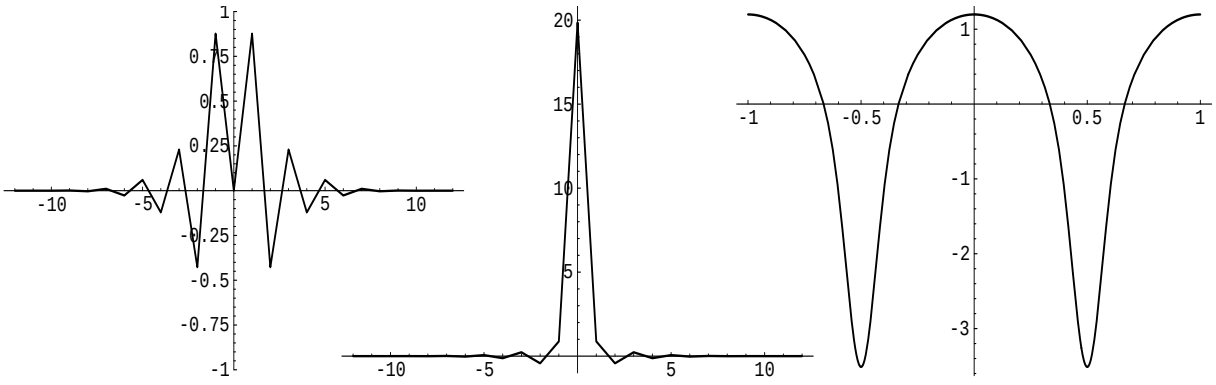

Figure 12: γ^0, γ, and $\hat{\gamma}^0$ for "Vaidyanathan 24" low-pass QMF.

$n \in \mathbf{Z}$. *If f is Hermitean symmetric or antisymmetric about some integer or half integer T, then the phase response of f is linear.*

Proof: We have $f(n) = \pm\bar{f}(2T - n)$ for all $n \in \mathbf{Z}$, taking + in the symmetric case and − in the antisymmetric case. Now $\gamma^0(0) = 0$ for all filters. For $n \neq 0$ we have

$$\begin{aligned}\gamma^0(n) &= \sum_k kf(k-n)\bar{f}(k+n) = \sum_k k\bar{f}(2T-k+n)f(2T-k-n) \\ &= 2T\sum_k \bar{f}(k+n)f(k-n) - \sum_k k\bar{f}(k+n)f(k-n) = \quad 0 - \gamma^0(n).\end{aligned}$$

Thus we have $\gamma^0(n) = 0$ for all $n \in \mathbf{Z}$. □

The linear function shifts the center of energy x to $2x - c[f]$, and the convolution operator γ^0 perturbs this by a "deviation" $\langle u, \gamma^0 * u\rangle / \|u\|^2$. We can denote the maximum value of this perturbation by $d[f]$. By Plancherel's theorem and the convolution theorem, the deviation is $\langle \hat{u}, \hat{\gamma}^0 \hat{u}\rangle / \|u\|^2$ and its maximum value is given by the maximum absolute value of $\hat{\gamma}^0(\xi)$:

$$d[f] = \sup\{|\hat{\gamma}^0(\xi)| : \xi \in [0,1]\}. \tag{21}$$

Now $\gamma^0(n) = \bar{\gamma}^0(-n)$ is symmetric just like γ, so its Fourier transform $\hat{\gamma}^0$ is purely real and can be computed using only cosines as follows:

$$\hat{\gamma}^0(\xi) = 2\sum_{n=1}^{\infty} \gamma(n)\, \cos 2\pi n\xi. \tag{22}$$

The critical points of $\hat{\gamma}^0$ are found by differentiating Equation 22:

$$\hat{\gamma}_0'(\xi) = -4\pi \sum_{n=1}^{\infty} n\gamma(n)\, \sin 2\pi n\xi. \tag{23}$$

It is evident that $\xi = 0$ and $\xi = \frac{1}{2}$ are critical points. We will look at a list of 17 example QMFs of lengths 2 through 30, for which $|\hat{\gamma}^0(\xi)|$ achieves its maximum at $\xi = \frac{1}{2}$, where

$$\hat{\gamma}^0\left(\frac{1}{2}\right) = 2\sum_{n=1}^{\infty}(-1)^n\gamma(n) = 2\sum_{k=-\infty}^{\infty}\sum_{n=1}^{\infty}(-1)^n kf(k-n)\bar{f}(k+n). \tag{24}$$

Graphs of $\hat{\gamma}^0$ for some of the example QMFs can be seen in Figures 9 through 12.

Values of the quantities $c[f]$ and $d[f]$ for example QMFs are listed in Table 2. Notice that if $g(n) = (-1)^n\bar{h}(2M+1-n)$, so that h and g are a conjugate pair of filters, and $|\operatorname{supp} g| = |\operatorname{supp} h| = 2M$ is the length of the filters, then $d[g] = d[h]$ and $c[g] + c[h] = 2M - 1$. This also implies that $C_h(i,j) = -C_g(i,j)$, so that the function $\hat{\gamma}^0$ corresponding to the filter h is just the negative of the one corresponding to g.

We can put the preceding formulas together into a single theorem:

Theorem 15 (QMF Phase Shifts) *Suppose that $u \in \ell^2$ and that $F : \ell^2 \to \ell^2$ is convolution and decimation by two with an orthogonal QMF $f \in \ell^1$. Suppose that $c[u]$ and $c[f]$ both exist. Then*

$$c[F^*u] = 2c[u] - c[f] - \langle u, \gamma^0 * u\rangle / \|u\|^2,$$

f	$\lvert\text{supp } f\rvert$	H or G	$c[f]$	$d[f]$
B	18	H	2.4439712920	2.6048841893
		G	14.5560287079	2.6048841893
C	6	H	3.6160691415	0.4990076823
		G	1.3839308584	0.4990076823
	12	H	4.0342243997	0.0868935216
		G	6.9657756002	0.0868935217
	18	H	6.0336041704	0.1453284669
		G	10.9663958295	0.1453284670
	24	H	8.0333521640	0.1953517707
		G	14.9666478359	0.1953517692
	30	H	10.0333426139	0.2400335062
		G	18.9666573864	0.2400330874
D	2	H	0.5000000000	0.0000000000
		G	0.5000000000	0.0000000000
	4	H	0.8504809471	0.2165063509
		G	2.1495190528	0.2165063509
	6	H	1.1641377716	0.4604317871
		G	3.8358622283	0.4604317871
	8	H	1.4613339067	0.7136488576
		G	5.5386660932	0.7136488576
	10	H	1.7491114972	0.9711171403
		G	7.2508885027	0.9711171403
	12	H	2.0307505738	1.2308332718
		G	8.9692494261	1.2308332718
	14	H	2.3080529576	1.4918354676
		G	10.6919470423	1.4918354676
	16	H	2.5821186257	1.7536045071
		G	12.4178813742	1.7536045071
	18	H	2.8536703515	2.0158368941
		G	14.1463296483	2.0158368941
	20	H	3.1232095535	2.2783448731
		G	15.8767904464	2.2783448731
V	24	H	19.8624838621	3.5116226595
		G	3.1375161379	3.5116226595

Table 2: Center-of-energy shifts and errors for some example QMFs.

where $\gamma^0 \in \ell^2$ is the sequence

$$\gamma^0(n) = \begin{cases} 0, & \text{if } n = 0, \\ \sum_k k f(k-n)\bar{f}(k+n), & \text{if } n \neq 0. \end{cases}$$

The last term satisfies the sharp inequality

$$|\langle u, \gamma^0 * u\rangle| \leq d[f]\,\|u\|^2,$$

where

$$d[f] = 2\left|\sum_{k=-\infty}^{\infty}\sum_{n=1}^{\infty}(-1)^n k f(k-n)\bar{f}(k+n)\right|.$$

□

If $d[f]$ is small, then we can safely ignore the deviation of F^*u from a pure shift of u by $c[f]$. In that case, we will say that $c[F^*u] \approx 2c[u] - c[f]$ and $c[Fu] \approx \frac{1}{2}c[u] + \frac{1}{2}c[f]$. We note that the "C" filters have the smallest errors $d[f]$; these are the filters to use if we wish to extract reasonably accurate position information.

If we apply a succession of filters $F_1^*F_2^*\cdots F_L^*$, then by induction on L we can compute the shifts as follows:

$$c[F_1^*F_2^*\cdots F_L^*u] = 2^L c[u] - 2^{L-1}c[f_L] - \cdots - 2^1 c[f_2] - c[f_1] - \epsilon^*, \tag{25}$$

where

$$|\epsilon^*| \leq 2^{L-1}d[f_L] + \cdots + 2^1 d[f_2] + d[f_1]. \tag{26}$$

Similarly, if $v = F_1^*F_2^*\cdots F_L^*u$, so that $F_L\cdots F_2F_1v = u$, then the following holds:

$$c[F_L\cdots F_2F_1v] = 2^{-L}c[v] + 2^{-L}c[f_1] + 2^{-L+1}c[f_2] + \cdots + 2^{-1}c[f_L] + \epsilon, \tag{27}$$

where

$$|\epsilon| \leq 2^{-1}d[f_L] + \cdots + 2^{-L+1}d[f_2] + 2^{-L}d[f_1]. \tag{28}$$

Now suppose that (h, g) is a conjugate pair of QMFs, so that $f_i \in \{h, g\}$ for each $i = 1, 2, \ldots, L$. Then $d[f_i]$ is constantly $d[h]$ and we have the simpler estimates for the deviation from a pure shift:

$$|\epsilon^*| \leq (2^L - 1)\,d[h] \approx 2^L d[h] \qquad \text{and} \qquad |\epsilon| \leq (1 - 2^{-L})\,d[h] \approx d[h]. \tag{29}$$

Suppose that we encode the sequence of filters $F_1^*F_2^*\cdots F_L^*$ as the integer $b = b_1 2^{L-1} + b_2 2^{L-2} + \cdots + b_L 2^0$, where

$$b_k = \begin{cases} 0, & \text{if } F_k = H; \\ 1, & \text{if } F_k = G. \end{cases} \tag{30}$$

Then we can write $c[f_k] = b_k c[g] + (1 - b_k)c[h] = c[h] + b_k(c[g] - c[h])$. Notice that the bit-reversal of b, considered as an s-bit binary integer, is the integer $b' = b_1 2^0 + b_2 2^1 + \cdots + b_L 2^{L-1}$. This simplifies the formula for the phase shift as follows:

Corollary 2 *If h and g are a conjugate pair of QMFs with centers of energy $c[h]$ and $c[g]$, respectively, then*

$$c[F_1^* F_2^* \cdots F_L^* u] = 2^L c[u] - \left(2^L - 1\right) c[h] - (c[g] - c[h])\, b' - \epsilon^*, \tag{31}$$

where $|\epsilon^*| \leq (2^L - 1)\, d[h]$ and $b = b_1 2^{L-1} + b_2 2^{L-2} + \cdots + b_L$ encodes the sequence of filters as in Equation 30, and b' is the bit-reversal of b considered as an L-bit binary integer.

Proof: We observe that

$$\begin{aligned} c[F_1^* F_2^* \cdots F_L^* u] &= 2^L c[u] - \sum_{k=1}^{L} 2^{L-k} \Big[c[h] + b_{L-k+1}(c[g] - c[h]) \Big] - \epsilon^* \\ &= 2^L c[u] - c[h] \sum_{s=0}^{L-1} 2^s - (c[g] - c[h]) \sum_{s=0}^{L-1} b_{s+1} 2^s - \epsilon^* \\ &= 2^L c[u] - \left(2^L - 1\right) c[h] - (c[g] - c[h]) b' - \epsilon^*. \end{aligned}$$

The estimate on ϵ^* follows from Equation 29. □

4.2 Shifts in the Periodic Case

Defining a center of energy for a periodic signal is problematic. However, if a periodic signal contains a component with a distinguishable scale much shorter than the period, then it may be desirable to locate this component within the period. If the component is characterized by a large amplitude found by filtering, then we can locate it by interpreting the "position" information of the filter output. We must adjust this position information by the center-of-energy shift caused by filtering, and allow for the deviation due to phase nonlinearity. In the periodic case, the shift can be approximated by a cyclic permutation of the output coefficients.

We can compute the center of energy of a nonzero q-periodic sequence u_q as follows:

$$c[u_q] = \frac{1}{\|u_q\|^2} \sum_{k=0}^{q-1} k |u_q(k)|^2.$$

Since $c[u_q]$ is a convex combination of $0, 1, \ldots, q-1$, we have $0 \leq c[u_q] \leq q-1$. Now suppose that u_q is the q-periodization of u and that all but ϵ of the energy in the sequence u comes from coefficients in one period interval $J_0 \stackrel{\text{def}}{=} [j_0 q, j_0 q + q - 1]$, for some integer j_0 and some positive $\epsilon \ll 1$. We must also suppose that u has a finite position uncertainty which is less than q. These conditions may be succinctly combined into the following:

$$\left(\sum_{j \notin J_0} \left[j - (j_0 + \frac{1}{2}) q \right]^2 |u(j)|^2 \right)^{\frac{1}{2}} < q\epsilon \|u\|. \tag{32}$$

For sufficiently small ϵ, we can compute the center of energy of u_q in terms of the center of energy of u. We summarize the calculation as follows: if almost all of the energy of u is

concentrated on an interval of length q, then transient features of u have a scale smaller than q and will become transient features of u_q upon q-periodization. Assuming that u satisfies Equation 32 with $\epsilon < \frac{1}{8q}$, we can use the following approximation to locate the center of energy of a periodized sequence to within one index:

$$c[u_q] \stackrel{\text{def}}{=} c[u] \bmod q. \tag{33}$$

We interpret the expression "$x \bmod q$" to mean the unique real number x' in the interval $[0, q[$ such that $x = x' + nq$ for some integer n.

We next use the following approximation:

$$\begin{aligned} c[F_{2q}^* u_q] &= c[(F^* u)_{2q}] = c[F^* u] \bmod 2q \\ &= 2c[u] - c[f] - \langle u, \gamma^0 * u \rangle / \|u\|^2 \bmod 2q. \end{aligned}$$

This result is proved in [13]. Now $\langle u, \gamma^0 * u \rangle / \|u\|^2$ is bounded by $d[f]$ so we plan to ignore it as before, though we must still verify that the QMFs satisfy Equation 32 with sufficiently small ϵ. Table 3 shows the value of ϵ for a few example QMFs and a few example periodizations. In all cases $\epsilon < 1$, so the table lists only the digits after the decimal point.

References

[1] Albert Cohen. *Ondelettes, Analyses Multirésolutions et Traitement Numérique du Signal*. PhD thesis, Université Paris IX Dauphine, 1990.

[2] Ronald R. Coifman, Yves Meyer, and Mladen Victor Wickerhauser. Size properties of wavelet packets. In Mary Beth Ruskai, Gregory Beylkin, Ronald Coifman, Ingrid Daubechies, Stéphane Mallat, Yves Meyer, and Louise Raphael, editors, *Wavelets and Their Applications*, pages 453–470. Jones and Bartlett, Boston, 1992.

[3] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, XLI:909–996, 1988.

[4] Gerald B. Folland. *Harmonic Analysis in Phase Space*. Number 122 in Annals of Mathematics Studies. Princeton University Press, Princeton, New Jersey, 1989.

[5] Michael Frazier, Bjørn Jawerth, and Guido Weiss. *Littlewood–Paley Theory and the Study of Function Spaces*. Number 79 in CBMS Regional Conference Lecture Notes. American Mathematical Society, Providence, Rhode Island, 1990.

[6] Nikolaj Hess-Nielsen. *Time-Frequency Analysis of Signals Using Generalized Wavelet Packets*. PhD thesis, University of Aalborg, 1992.

[7] Nikolaj Hess-Nielsen. Control of frequency spreading of wavelet packets. *Applied and Computational Harmonic Analysis*, 1(2):157–168, March 1994.

[8] Yves Meyer. De la recherche pétrolière à la géometrie des espaces de Banach en passant par les paraproduits. Technical report, École Polytechnique, Palaiseau, 1985–1986.

f	$\lvert\text{supp } f\rvert$	H or G	$q=2$ $q=16$	$q=4$ $q=18$	$q=6$ $q=20$	$q=8$ $q=22$	$q=10$ $q=24$	$q=12$ $q=26$	$q=14$ $q=28$
B	18	H	.703612 .001415	.279300	.142238	.074249	.033688	.014072	.005406
		G	.734120 .001590	.324821	.163452	.087139	.038976	.016137	.006156
C	6	H	.247013	.102745					
		G	.268885	.069768					
	12	H	.263115	.072831	.033281	.010694	.001009		
		G	.251051	.070544	.028711	.009039	.001205		
	18	H	.299435 .000040	.100032	.052849	.018963	.007231	.002661	.000621
		G	.291211 .000045	.098243	.046702	.017889	.007332	.002556	.000708
	24	H	.329096 .000890	.120402 .000331	.065564 .000036	.027330 .000002	.014121	.005809	.002328
		G	.322880 .000936	.119051 .000367	.060292 .000039	.027004 .000002	.013983	.005754	.002531
	30	H	.354113 .002035	.136558 .000958	.075916 .000291	.035107 .000138	.020482 .000026	.009303 .000002	.004743 .000000
		G	.349093 .002111	.135636 .001051	.071338 .000285	.035330 .000134	.020121 .000024	.009401 .000002	.005009 .000000
D	4	H	.171193						
		G	.273971						
	6	H	.304120	.050230					
		G	.259392	.073125					
	8	H	.308900	.102651	.017895				
		G	.323009	.122720	.023634				
	10	H	.342554	.135552	.040530	.006627			
		G	.449328	.116023	.053618	.008251			
	12	H	.422494	.137647	.058646	.016224	.002475		
		G	.463486	.160047	.064599	.020210	.002964		
	14	H	.524235	.169394	.072909	.023686	.006412	.000924	
		G	.508880	.223013	.076843	.029062	.007680	.001077	
	16	H	.524480	.210433	.085366	.032061	.009408	.002489	.000344
		G	.587024	.220427	.103528	.038321	.011119	.002899	.000393
	18	H	.564454 .000128	.243878	.102607	.045068	.014338	.003662	.000948
		G	.636888 .000144	.238832	.128066	.050826	.016666	.004213	.001082
	20	H	.634131 .000354	.248979 .000047	.120135	.051443	.024453	.006775	.001411
		G	.672192 .000398	.282813 .000053	.138670	.060597	.025714	.007739	.001591
V	24	H	.872011 .006270	.390176 .001937	.217686 .000629	.116186 .000191	.062451	.036782	.017151
		G	.829783 .005653	.355441 .001764	.190529 .000574	.101064 .000175	.057180	.034695	.015266

Table 3: Concentration of energy for some example orthogonal QMFs.

[9] Yves Meyer. *Ondelettes et Opérateurs*, volume I: Ondelettes. Hermann, Paris, 1990.

[10] Yves Meyer. *Ondelettes et Opérateurs*, volume II: Opérateurs. Hermann, Paris, 1990.

[11] G. Polya and G. Szegö. *Aufgaben und Lehrsätze aus der Analysis*, volume II. Springer, Berlin, 1971.

[12] Mladen Victor Wickerhauser. INRIA lectures on wavelet packet algorithms. In Pierre-Louis Lions, editor, *Problémes Non-Linéaires Appliqués, Ondelettes et Paquets D'Ondes*, pages 31–99. INRIA, Roquencourt, France, 17–21 June 1991. Minicourse lecture notes.

[13] Mladen Victor Wickerhauser. *Adapted Wavelet Analysis from Theory to Software*. AK Peters, Ltd., Wellesley, Massachusetts, 9 May 1994. With optional diskette.