A Brief Tutorial on Wavelets

Robot Vision Lab, Purdue University
January 1998

Let's say we want to develop an orthonormal representation for a set of signals (or images), but we wish to do so in such a manner that ephemeral phenomena (infrequently occurring but important transitions) do not get lost in the process.

To explain this requirement, consider the following waveform:

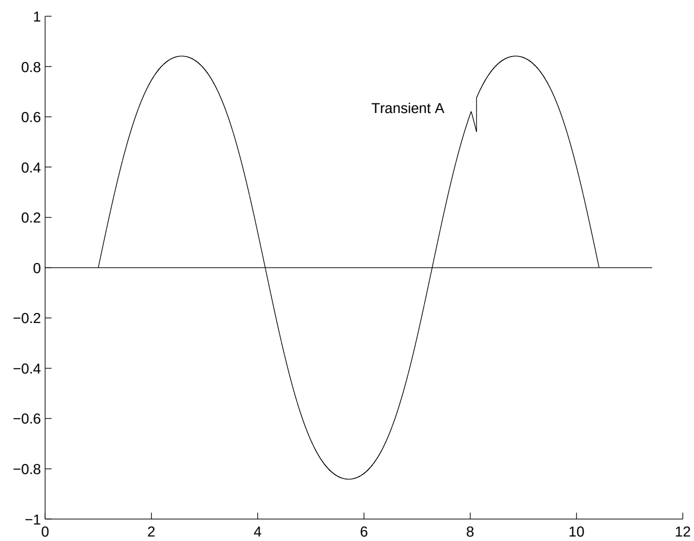

Figure 1: *A sin(x) signal with a small transient at point (A).*

In a Fourier-based orthonormal expansion of this waveform, the energy in the transient at (A) would get averaged over the entire duration of the waveform, the result being that the magnitude of the high-frequency components that represent the transient at (A) would be so miniscule in relation to the magnitude of the low-frequency components that the former would probably disappear in any thresholding operation. The magnitude of each high-frequency component would be proportional to the inner product of the high-frequency sinusoid with the transient integrated over the duration of the entire waveform.

Ephemeral phenomena are much better captured by a hierarchical representation in which short-duration changes can be captured by inner products with short-duration basis functions. Consider the following function

$$\Psi(t) = (1 - t^2)e^{-\frac{1}{2}t^2} \tag{1}$$

As many of you might remember, this function is the LoG edge detection operator in computer vision. The LoG operator derives its power from first smoothing the function with a Gaussian and then taking a Laplacian of the smoothed function.

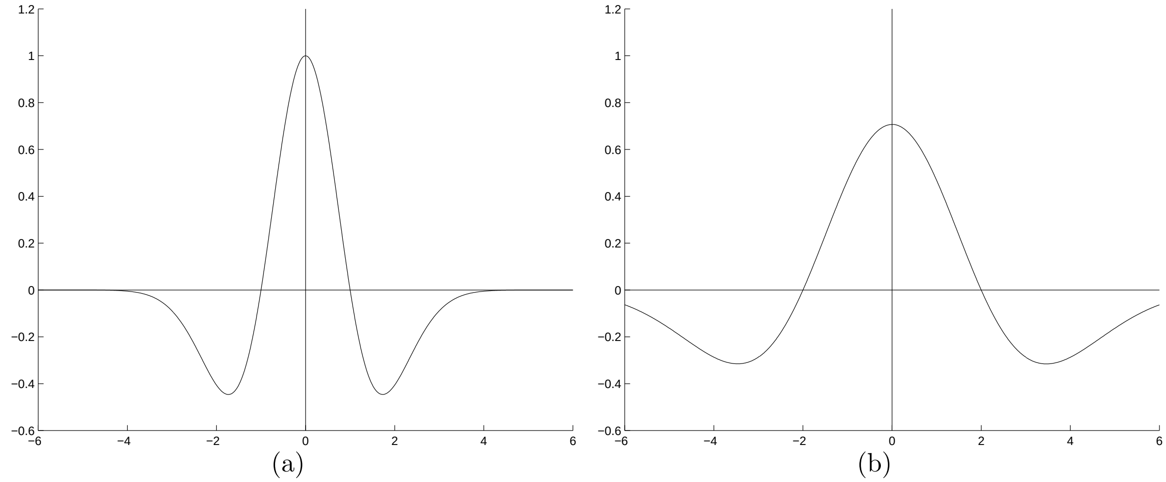

Figure 2: *(a) Mexican-hat function. (b) Dilated version of (a).*

A set of translated and non-overlapping versions of the above function could constitute an orthonormal basis (but, by no means, a complete orthonormal basis). So we could try to represent a function f(t) by the following set of coefficients:

$$T_n(f) = \int_{-\infty}^{+\infty} f(t)\Psi(t - nb)dt, \tag{2}$$

where b is the distance between the successive positions of the basis $\Psi(t)$. If the duration of the $\Psi(t)$ was comparable to the duration of the localized transient at (A) in Figure 1, this calculation here will do a much better job of capturing that transient.

The coefficients computed by the above integration will capture changes in the signal, but only at a resolution level that will be determined by the amount of smoothing incorporated in the basis function $\Psi(t)$ and the extent of that function. So we could try to add additional coefficients to those extracted previously by taking inner products of $f(t)$ with dilated and contracted versions of $\Psi(t)$. Figure 2 shows a Mexican hat function and its dilated version where the dilation is by a factor of 2.

For example, to capture changes in $f(t)$ at a factor of 2 coarser resolution, we could use the following basis function

$$\begin{aligned} \Psi_2(t) &= \frac{1}{\sqrt{2}}\Psi(\frac{t}{2}) \\ &= \frac{1}{\sqrt{2}}\left[1 - \left(\frac{t}{2}\right)^2\right] e^{-\frac{1}{2}(\frac{t}{2})^2} \end{aligned} \tag{3}$$

Using these scaled versions of $\Psi(t)$, we can now define a more complete representation of the function $f(t)$ by

$$T_{m,n}(f) = \int f(t)\Psi_{2^m}(t - nb)dt, \quad m = 0, 1, 2, ... \tag{4}$$

The only problem with these inner products is that since at each scale we want the translation to be non-overlapping, we should not be translating by the same extent, that is by nb, at every scale. A more appropriate translation would be by $n2^m b_0$ for scale m. This means that the dilated $\Psi(t)$ would be translated by larger extents than the contracted versions. This is reflected in the following formula:

$$\begin{aligned} T_{m,n}(f) &= \int f(t)\Psi_{2^m}(t - n2^m b_0)dt \\ &= \frac{1}{2^{m/2}} \int f(t)\Psi\left(\frac{t - n2^m b_0}{2^m}\right) dt \\ &= \frac{1}{2^{m/2}} \int f(t)\Psi\left(2^{-m}t - nb_0\right) dt, \quad m = 0, 1, 2, ... \end{aligned} \tag{5}$$

This was the formalism that was used for many years for calculating hierarchical representations for signals and images. But note that in the form it has been presented, it only captures changes (meaning, the detail) at different resolutions. It indirectly discards the smooth part of the underlying signal; the smooth part gets ignored because of the second-derivative inherent in the LoG function. The coefficients that are calculated for different values of m could be represented pictorially by

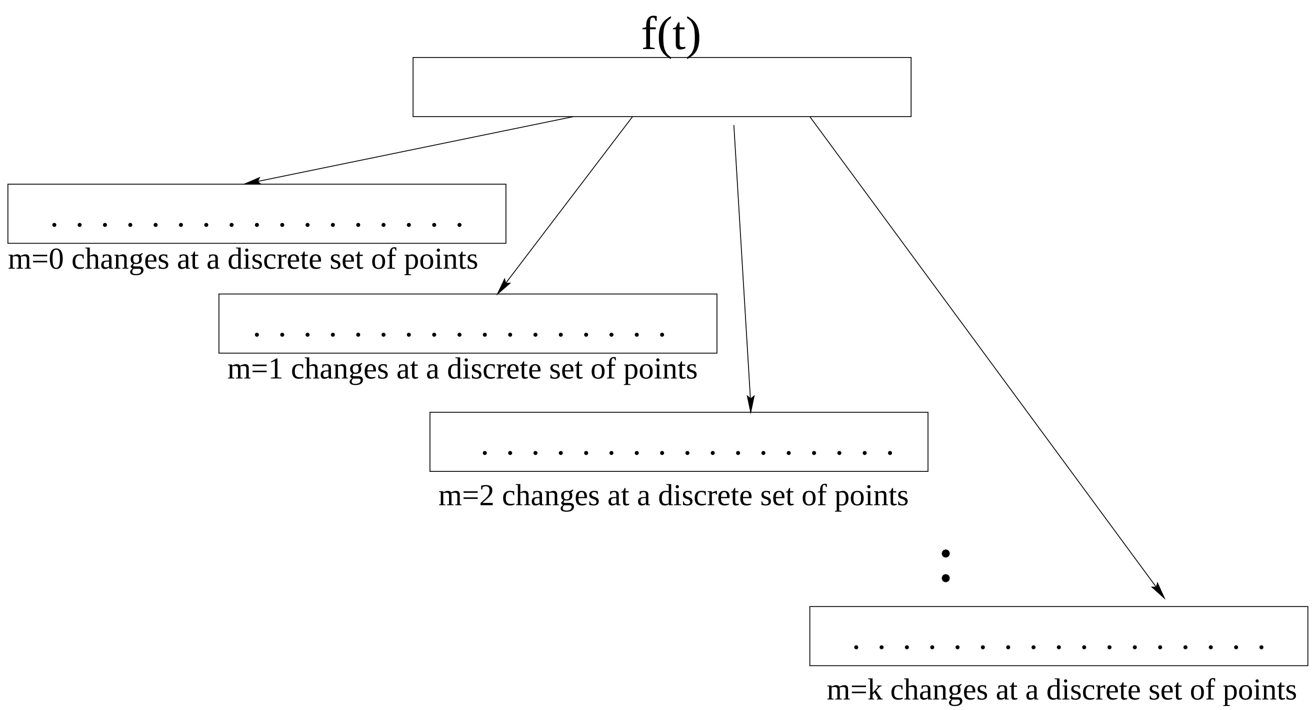

Figure 3: *Applying the Mexican-hat function at different scales yields coefficients that represent changes at those scales.* $m = 0$ *corresponds to the basic wavelet, that is without any dilation.* $m = 1$ *corresponds to dilation of the wavelet by a factor of 2.*

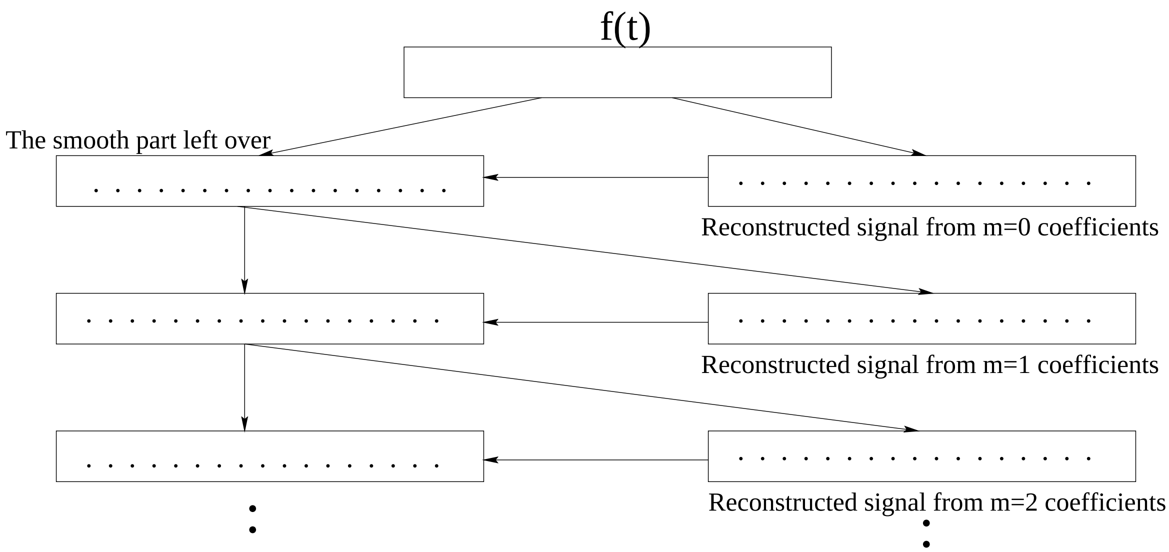

Figure 4: *This figure shows how a wavelet analysis using the Mexican-hat function can be cast in such a way as to show the smooth and the detail parts separately at different scales.*

However, in order to retain the smooth part, we could recast the processing involved in the following form that we will now discuss. Consider first the changes detected at the scale $m = 0$. From the coefficients $T_{0,n}$, we can try to reconstruct the detail part of the function at scale 0 as follows

$$D_0(t) \quad = \quad \sum_n T_{0,n}(f)\Psi_{2^0}(t - nb_0) \tag{6}$$

where $\Psi_{2^0}(t - nb_0)$ is the same thing as $\Psi(t)$. Clearly, if we subtract this from the function $f(t)$, we'll get the left-over smooth part, represented by $S_0(t)$, corresponding to that scale

$$S_0(t) \quad = \quad f(t) \; - \; D_0(t) \tag{7}$$

Figure 5 shows a step function and the detail and the smooth parts of the function at level $m = 1$ when a Mexican-hat function is used as a wavelet. To extract the detail and the smooth parts at level $m = 2$, we first write down the form of the wavelet at this level of resolution

$$\begin{aligned} \Psi_{2^1}(t) \quad &= \quad \frac{1}{\sqrt{2}}\Psi_{2^0}(\frac{t}{2}) \\ &= \quad \frac{1}{\sqrt{2}}\left[1 - \left(\frac{t}{2}\right)^2\right] e^{-\frac{1}{2}(\frac{t}{2})^2} \end{aligned} \tag{8}$$

By taking inner products of this wavelet with the smooth part obtained at $m = 1$ level, we obtain

$$T_{1,n}(f) = \int S_0(t)\Psi_{2^1}(t - n2b)dt \tag{9}$$

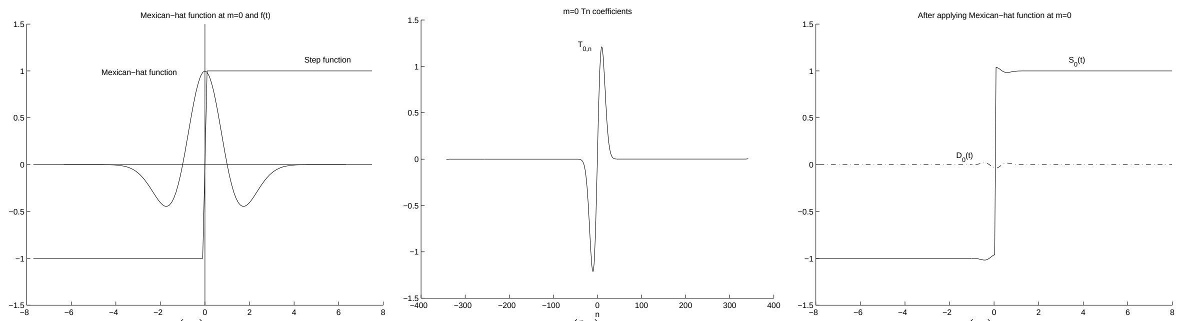

Figure 5: *(a) Mexican-hat function superimposed on a step function. (b) Coefficients $T_{0,n}$. (c) The dashed line represents the detail part, $D_0(t)$, and the solid line the smooth part, $S_0(t)$.*

As before, these coefficients can be used to reconstruct the detail part corresponding to $m = 2$ level, and so on. This process can be continued to yield the detail and the smooth parts at different levels of resolution in the manner shown in Fig. 4.

This is basically how modern wavelet processing is done, but with a more rigorous attention to completeness issues. In what follows, we will present a more thorough analysis of the wavelets.

1 A Vector Space Formalization of Wavelets

Let V_0 be the vector space in which a given function $f(t)$ can be represented without error. There must exist in V_0 an orthonormal basis set[1]

$$\Phi_{0,n}(t)$$

such that

$$lim_{K\rightarrow\infty}||f(t) - \sum_{k=0}^{K} \alpha_k \Phi_{0,k}(t)|| \quad = \quad 0 \tag{10}$$

Now let's consider another vector space V_1 such that $f(\frac{t}{2}) \in V_1 \Longleftrightarrow f(t) \in V_0$, $\forall\ f(t)$. Note that $f(\frac{t}{2})$ is a dilated version of $f(t)$. Then there must exist in the vector space V_1 an orthonormal basis set

$$\Phi_{1,n}(t)$$

[1]In this report – and also in many other wavelet related contributions – the functions denoted by Φ are used to represent signals and images. On the other hand, the functions represented by Ψ are used to detect changes in signals and images.

such that

$$lim_{K\to\infty}||f(\frac{t}{2}) - \sum_{k=0}^{K}\beta_k\Phi_{1,k}(t)|| \quad = \quad 0 \tag{11}$$

Note that since $f(\frac{t}{2})$ is a smoother version of $f(t)$, the space V_1 must be a subspace of V_0. It also stands to reason that from the orthonormal basis functions $\Phi_{0,k}(t)$ we should be able to obtain $\Phi_{1,k}(t)$ by

$$\Phi_{1,k}(t) = \frac{1}{\sqrt{2}}\Phi_{0,k}(\frac{t}{2}) \tag{12}$$

Similarly, we could consider a vector space V_2 in which $f(\frac{t}{2^2})$ can be represented without error, and so on. We can thus create a hierarchy of spaces

$$... \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset ... \tag{13}$$

where we have extended the hierarchy in both directions. An orthonormal basis functions for the space V_j will be given by

$$V_j : \qquad \Phi_{j,k}(t) = 2^{\frac{-j}{2}}\Phi_{0,k}(\frac{t}{2^j}) \tag{14}$$

Let's focus on V_1 for a moment. V_1 is a subspace of V_0. Let W_1 represent that part of V_0 that is orthogonal to V_1 in V_0. What that means is that V_1 and W_1 taken together constitute the space V_0:

$$V_0 \quad = \quad V_1 \oplus W_1$$

If for each V_j we define a W_j in a similar manner, we have

$$V_{j-1} \quad = \quad V_j \oplus W_j$$

We can show these relationships pictorially as

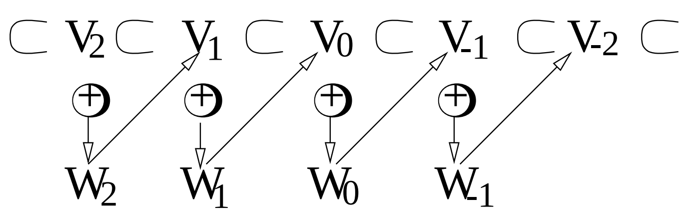

Figure 6: *The relationship between the V_j spaces and the W_j spaces is shown here.*

Given a function $f(t)$, we can project it onto V_j by applying the projection operator P_j

$$f_j(t) \;=\; P_j f(t) \tag{15}$$

At resolution level 1, the projection can be obtained by:

$$\begin{aligned} f_1(t) &= P_1 f(t) \\ &= \sum_k S_{1,k}\Phi_{1,k}(t) \end{aligned} \tag{16}$$

where $S_{1,k}$ are the inner products of $f(t)$ and $\Phi_{1,k}$.

$$S_{1,k} = < f(t), \Phi_{1,k}(t) > \tag{17}$$

We will now assume that at a given resolution the orthonormal functions $\Phi_{1,k}(t)$ are simply translated versions of one another. For example, at resolution level 1

$$\Phi_{1,k}(t) = \Phi_{1,0}(t - 2k) \tag{18}$$

where the multiplier 2 for the translation accounts for making the translation proportional to the dilation at this level of resolution. But since

$$\Phi_{1,0}(t) = \frac{1}{\sqrt{2}}\Phi_{0,0}(\frac{t}{2}) \tag{19}$$

we have

$$\begin{aligned} \Phi_{1,k}(t) &= \frac{1}{\sqrt{2}}\Phi_{0,0}(\frac{t-2k}{2}) \\ &= \frac{1}{\sqrt{2}}\Phi_{0,0}(2^{-1}t - k) \end{aligned} \tag{20}$$

In a similar manner, the basis functions of the V_j space will be given by

$$\Phi_{j,k}(t) = 2^{-\frac{j}{2}}\Phi_{0,0}(2^{-j}t - k) \tag{21}$$

The function $f(t)$ may now be projected into the V_j space by

$$\begin{aligned} f_j(t) = P_j f(t) &= \sum_k S_{j,k}\Phi_{j,k}(t) \\ &= \sum_{k'} S_{j+1,k'}\Phi_{j+1,k'}(t) + \sum_{k'} D_{j+1,k'}\Psi_{j+1,k'}(t) \end{aligned} \tag{22}$$

where $S_{j+1,k'}$ and $D_{j+1,k'}$ are the coefficients at level $j+1$ from the smooth part and the detail part, respectively. We now can keep projecting function $f(t)$ into a further resolution level, $j+1$:

$$S_{j+1,k} = < f_j(t), \Phi_{j+1,k}(t) > \tag{23}$$

This gives us a mechanism for creating smoother and smoother versions of a function $f(t)$.

But in many applications, we are specifically interested in extracting the detail at different resolutions, as opposed to the smoothed part at different resolutions. Of course, at any resolution the detail and the smooth part are intimately related by a mere subtraction. For example, the detail corresponding to the subspace W_1 is given by

$$\begin{aligned} e_1(t) &= f(t) - [projection\ of\ f(t)\ onto\ V_1] \\ &= f(t) - P_1 f(t) \end{aligned} \tag{24}$$

and, in general, we can write

$$\begin{aligned} e_{j+1}(t) &= [projection\ of\ f(t)\ onto\ V_j] - [projection\ of\ f(t)\ onto\ V_{j+1}] \\ &= P_j f(t) - P_{j+1} f(t) \end{aligned} \tag{25}$$

The smooth and the detail part at different resolutions can be shown simultaneously in the following form:

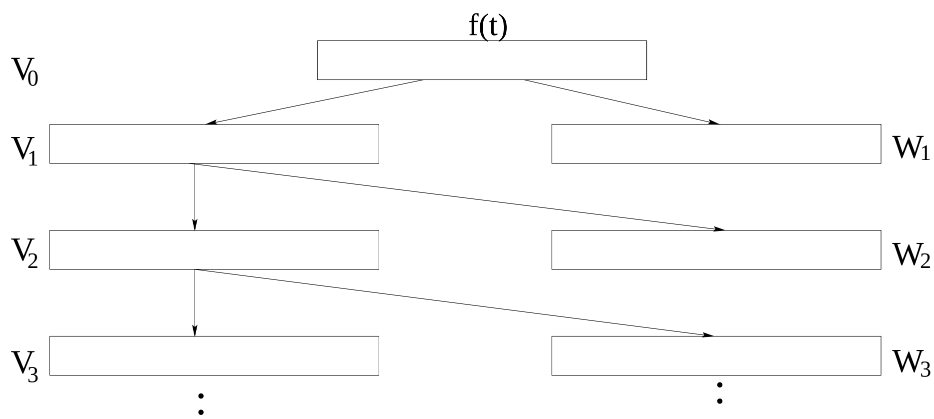

Figure 7: *Shows schematically the extraction of the smooth versions of f(t) at different scales and the extraction of the details at the same scales.*

We can also write down directly closed-form expressions for projecting f(t) directly onto the subspaces W_1, W_2, etc. Recall that W_0 is spanned by the orthonormal basis set

$$\Psi_{0,k}(t)$$

From the relationship between the space V_j and the space V_0, we can write

$$\Psi_{j,0}(t) = \frac{1}{2^{\frac{j}{2}}} \Psi_{0,0}(\frac{t}{2^j}) \tag{26}$$

If we make the usual wavelet assumption that the space W_0 is spanned by a set of functions that are orthonormal by virtue of their displacement property, we can write

$$\Psi_{0,k}(t) = \Psi_{0,0}(t-k) \tag{27}$$

for resolution level 0. The same relation for resolution level j becomes

$$\Psi_{j,k}(t) \;=\; \Psi_{j,0}(t - k2^j) \tag{28}$$

By substituting one of the previous equations, we can now write

$$\begin{aligned}\Psi_{j,k}(t) \quad &= \quad 2^{-\frac{j}{2}}\Psi_{0,0}(\frac{t-k2^j}{2^j} \\ &= \quad 2^{-\frac{j}{2}}\Psi_{0,0}(2^{-j}t - k)\end{aligned} \tag{29}$$

2 Self-replication Constraints on Φ and Ψ

2.1 Constraint equation for Φ

In this section, we will show that our base wavelets $\Phi_{0,0}$ and $\Psi_{0,0}$ must obey certain mathematical constraints that can be referred to as the self-replicating constraints. For the discussion in this section, we will use the simpler notation Φ and Ψ for $\Phi_{0,0}$ and $\Psi_{0,0}$, respectively.

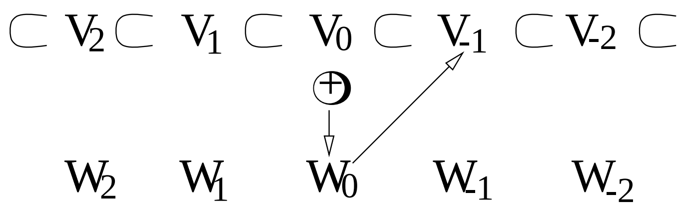

Figure 8: *Same as Fig. 5 except that now we are focusing specifically on the relationship between V_0 , W_0 and V_{-1}*

Recall, $\Phi(t)$ are the basis functions in the space V_0. But the space of functions V_0 is a subspace of the space of functions V_{-1}, in just the same manner that the space of functions V_1 is a subspace of the space of functions V_0 (see Eq. (13)). Therefore, it must be the case that any function in V_0 can be represented by a linear combination of the basis functions of V_{-1}. Extrapolating from the basis functions for V_j we showed in Eq. (14), the zero-displacement basis function for the space V_{-1} is given by

$$\Phi_{-1,0}(t) = \sqrt{2}\Phi(2t) \tag{30}$$

and therefore the displacement by k function for this level of resolution would be given by

$$\begin{aligned}\Phi_{-1,k}(t) \quad &= \quad \Phi_{-1,0}(t - \frac{k}{2}) \\ &= \quad \sqrt{2}\Phi(2(t - \frac{k}{2})) \\ &= \quad \sqrt{2}\Phi(2t - k)\end{aligned} \tag{31}$$

where we have used the abbreviated notation for $\Phi_{0,0}$. Since any function in V_0 can be expressed as a linear sum of these basis functions, the property must also hold for $\Phi(t)$. So we can write

$$\begin{aligned}\Phi(t) &= \sum_n h_n \Phi_{-1,n}(t) \\ &= \sqrt{2} \sum_n h_n \Phi(2t - n)\end{aligned} \tag{32}$$

Since the function Φ appears on both sides of this equation, we will refer to it as a self-replicating constraint. Evidently, Φ must be a solution of this equation.

It's easier to utilize the frequency domain version of the above equation. Taking Fourier transform of both sides of the equation, we can write

$$\begin{aligned}\hat{\Phi}(w) &= \frac{1}{\sqrt{2}} \sum_k h_k e^{-\frac{iwk}{2}} \hat{\Phi}(\frac{w}{2}) \\ &= m_0(\frac{w}{2}) \hat{\Phi}(\frac{w}{2})\end{aligned} \tag{33}$$

where m_0, a function of periodicity 2π, is given by

$$m_0(\frac{w}{2}) = \frac{1}{\sqrt{2}} \sum_k h_k e^{-iwk} \tag{34}$$

2.2 Constraint Equation for Ψ

Again invoking the property that any function in W_0, which is a subspace of V_{-1}, must be expressible as a linear combination of a basis set of functions in V_{-1}, we can express Ψ as

$$\Psi(t) = \sqrt{2} \sum_n g_n \Phi(2t - n) \tag{35}$$

where g_n is a set of coefficients for high-pass filter.

As we will see later, the coefficients g_n and h_n will play a direct role when we try to relate the coefficients at a given level of resolution with the coefficients at the previous level of resolution. Therefore, they will play a central role in the implementation of the calculation of the wavelet coefficients at multiple levels.

2.3 Inter-level Constraint Equations for Φ and Ψ

In this subsection, we would like to extend the constraints in Eqs. (32) and (35) to inter-level constraints. At resolution level $j+1$, from Eqs. (21) and (32), $\Phi_{j+1,k}$ can be presented as:

$$\begin{aligned}\Phi_{j+1,k}(x) &= 2^{\frac{-(j+1)}{2}} \Phi(2^{-(j+1)}x - k) \\ &= 2^{\frac{-(j+1)}{2}} \sqrt{2} \sum_n h_n \Phi(2(2^{-(j+1)}x - k) - n)\end{aligned}$$

$$
\begin{aligned}
&= 2^{\frac{-j}{2}} \sum_n h_n \Phi(2^{-j}x - (2k+n)) \\
&= \sum_m h_{m-2k} \Phi_{j,m}(x) \qquad (36)
\end{aligned}
$$

In the same manner, $\Psi_{j+1,k}$ can be presented as:

$$
\Psi_{j+1,k}(x) = \sum_m g_{m-2k} \Phi_{j,m}(x) \qquad (37)
$$

Equations (36) and (37) here for resolution level j+1 correspond to the equations (32) and (35) shown earlier for resolution level 0.

In next section, we will derive a relationship between the coefficients g_n's and the coefficients h_n's.

3 Relationship between the h_k and g_k coefficients

We will now show that the h_k coefficients in Eq. (32) and the g_k coefficients in Eq. (35) are related. In particular, we will show that the latter coefficients can be derived from the former. We will start by deriving a constraint equation for the Fourier transform of the h_k sequence. Recall from Eq. (34) that this Fourier transform is denoted $m_0(w)$.

3.1 An equation that must be satisfied by m_0

Since $\{\Phi(t-k)\}$ for different displacement values of the displacement index k form an orthonormal basis, it must be true

$$
\begin{aligned}
\delta_{k,0} &= <\Phi(t), \Phi(t-k)> \\
&= \int_{-\infty}^{+\infty} \hat{\Phi}(w) e^{-iwk} \overline{\hat{\Phi}(w)} dw \\
&= \int_{-\infty}^{+\infty} |\hat{\Phi}(w)|^2 e^{-iwk} dw \\
&= \int_0^{2\pi} \sum_n |\hat{\Phi}(w + 2n\pi)|^2 e^{-iwk} dw \qquad (38)
\end{aligned}
$$

The last expression on the right hand side follows from the simple fact that instead of integrating over the interval $(-\infty, +\infty)$, we can first integrate over the window $[0, 2\pi]$, slide the integrand to, say, the left by 2π, integrate again over the window $[0, 2\pi]$, etc. Since the left hand side equals the Kronecker's delta ($\delta_{i,j} = 1$ if $i = j$, 0 if $i \neq j$), it must be the case that the integrand under the $(0, 2\pi)$ integral is a constant as follows:

$$
\sum_n |\hat{\Phi}(w + 2n\pi)|^2 = \frac{1}{2\pi} \qquad (39)
$$

Now from Eq. (33) we can write

$$
\begin{aligned}
\sum_n |\hat{\Phi}(w+2n\pi)|^2 &= \sum_n |m_0(\frac{w}{2}+n\pi)\hat{\Phi}(\frac{w}{2}+n\pi)|^2 \\
&= \sum_{odd\ terms} |m_0(\frac{w}{2}+\pi+2n\pi)|^2|\hat{\Phi}(\frac{w}{2}+\pi+2n\pi)|^2 \\
&+ \sum_{even\ terms} |m_0(\frac{w}{2}+2n\pi)|^2|\hat{\Phi}(\frac{w}{2}+2n\pi)|^2 \qquad (40)
\end{aligned}
$$

But since m_o is 2π-periodic and the result of Eq. (39), we have

$$
\begin{aligned}
\sum_n |\hat{\Phi}(w+2n\pi)|^2 &= \sum_{odd\ terms} |m_0(\frac{w}{2}+\pi)|^2|\hat{\Phi}(\frac{w}{2}+\pi+2n\pi)|^2 \\
&+ \sum_{even\ terms} |m_0(\frac{w}{2})|^2|\hat{\Phi}(\frac{w}{2}+2n\pi)|^2 \\
&= \frac{1}{2\pi}\left(|m_0(\frac{w}{2})|^2 \ + \ |m_0(\frac{w}{2}+\pi)|^2\right) \qquad (41)
\end{aligned}
$$

Now using Eq. (39), we conclude

$$
|m_0(\frac{w}{2})|^2 + |m_0(\frac{w}{2}+\pi)|^2 = 1 \qquad (42)
$$

3.2 Fourier Transform Properties of Functions in W_0

In this subsection we will derive a useful property obeyed by the Fourier transforms of the functions in W_0. Eventually, we will use this property to delineate the relationship between the h_k and g_k coefficients, the main focus of Section 3.

Recall, W_0 is a subspace of V_{-1}. Therefore, any function in W_0 should be expressible in terms of the basis functions, $\Phi_{-1,k}$, of V_{-1}. Let's consider a function $q(t)$ in W_0. We can write

$$
\begin{aligned}
q(t) &= \sum_k l_k \Phi_{-1,k}(t) \\
&= \sqrt{2}\sum_k l_k \Phi(2t-k) \qquad (43)
\end{aligned}
$$

The Fourier transform of the above equation is

$$
\begin{aligned}
\hat{q}(w) &= \frac{1}{\sqrt{2}}\sum_k l_k e^{\frac{-ikw}{2}}\hat{\Phi}(\frac{w}{2}) \\
&= m_q(\frac{w}{2})\hat{\Phi}(\frac{w}{2}) \qquad (44)
\end{aligned}
$$

where the 2π-periodic function $m_q(\frac{w}{2})$ is the discrete Fourier transform of the sequence l_k:

$$
m_q(\frac{w}{2}) = \frac{1}{\sqrt{2}}\sum_k l_k e^{\frac{-ikw}{2}} \qquad (45)
$$

Since $q \in W_0$, $q(t)$ is orthogonal to the basis functions of V_0. That is $q \perp \Phi(t-k)$. In this case, the inner product of q and Φ should be 0:

$$\int_{-\infty}^{+\infty} q(t)\Phi(t-k)dt = 0 \tag{46}$$

Substituting Fourier integrals for each of the terms in the integrand, we get

$$\int_{-\infty}^{+\infty} \hat{q}(w)\overline{\hat{\Phi}(w)}e^{-ikw}dw \quad = \quad 0 \tag{47}$$

which is the same as

$$\int_{0}^{2\pi} \sum_n \hat{q}(w+2n\pi)\overline{\hat{\Phi}(w+2n\pi)}e^{-ikw}dw \quad = \quad 0 \tag{48}$$

This implies that

$$\sum_n \hat{q}(w+2n\pi)\overline{\hat{\Phi}(w+2n\pi)} \quad = \quad 0 \tag{49}$$

The $\hat{q}(w)$ obtained in Eq. (44) is plugged into the above equation. We also apply the constraint equation (33) to substitute the $\hat{\Phi}(w+2n\pi)$ term in the above equation. Hence we have

$$\sum_n m_q(\frac{w}{2}+n\pi)\overline{m_0(\frac{w}{2}+n\pi)}|\hat{\Phi}(\frac{w}{2}+n\pi)|^2 = 0 \tag{50}$$

We split the summation into two halves, one involving only the even values of the index n and the other only the odd values of the same index. This operation is similar to what was done in Eq. (41). Following steps identical to those shown earlier, we obtain

$$m_q(\frac{w}{2})\overline{m_0(\frac{w}{2})} \; + \; m_q(\frac{w}{2}+\pi)\overline{m_0(\frac{w}{2}+\pi)} = 0 \tag{51}$$

We introduce another 2π-periodic function $\lambda(w)$ as follow:

$$\lambda(\frac{w}{2}) \quad = \quad \frac{m_q(\frac{w}{2})}{\overline{m_0(\frac{w}{2}+\pi)}} \tag{52}$$

$$\Longrightarrow m_q(\frac{w}{2}) \quad = \quad \lambda(\frac{w}{2})\overline{m_0(\frac{w}{2}+\pi)} \tag{53}$$

The m_q is then plugged into Eq. (51).

$$\lambda(\frac{w}{2})\overline{m_0(\frac{w}{2}+\pi)m_0(\frac{w}{2})} \; + \; \lambda(\frac{w}{2}+\pi)\overline{m_0(\frac{w}{2}+2\pi)m_0(\frac{w}{2}+\pi)} \;=\; 0 \tag{54}$$

where m_0 is also a 2π-periodic function. $m_0(\frac{w}{2}+2\pi) = m_0(\frac{w}{2})$. From Eq. (42), we know that $m_0(\frac{w}{2}+\pi)$ and $m_0(\frac{w}{2})$ can not be 0 at the same time. The above equation can be simplified as follow:

$$\begin{aligned} \lambda(\frac{w}{2}) \quad &+ \quad \lambda(\frac{w}{2}+\pi) = 0 && (55) \\ \Longrightarrow \lambda(\frac{w}{2}) \quad &= \quad e^{j\frac{w}{2}} v(w) && (56) \end{aligned}$$

where $v(w)$ is another 2π-periodic function. From Eq. (44) and 53, we have

$$\begin{aligned} \hat{q}(w) \quad &= \quad [e^{\frac{iw}{2}} \overline{m_0(\frac{w}{2}+\pi)} \hat{\Phi}(\frac{w}{2})] v(w) && (57) \\ Also\ when \quad q \ &\in \quad W_0,\ we\ have \\ q(t) \quad &= \quad \sum_k c_k \Psi(t-k) \\ \Longrightarrow \hat{q}(w) \quad &= \quad [\sum_k c_k e^{-ikw}] \hat{\Psi}(w) && (58) \end{aligned}$$

The first term in Eq. (57) is invariant with q, so is the second term in Eq. (58). It suggests that the analyzer wavelet Ψ can be expressed as

$$\begin{aligned} \hat{\Psi}(w) \quad &= \quad e^{\frac{iw}{2}} \overline{m_0(\frac{w}{2}+\pi)} \hat{\Phi}(\frac{w}{2}) \\ \overset{Eq.34}{\Longrightarrow} \quad &= \quad e^{\frac{iw}{2}} \frac{1}{\sqrt{2}} \sum_k \overline{h_k} e^{j(\frac{w}{2}+\pi)k} \hat{\Phi}(\frac{w}{2}) \\ \Longrightarrow \quad &= \quad \frac{1}{\sqrt{2}} \sum_k e^{j\pi k} \overline{h_k} e^{\frac{-j(-k-1)w}{2}} \hat{\Phi}(\frac{w}{2}) \\ \Longrightarrow \quad &= \quad \frac{1}{\sqrt{2}} \sum_k (-1)^k \overline{h_k} e^{\frac{-j(-k-1)w}{2}} \hat{\Phi}(\frac{w}{2}) \\ \overset{Let\ n=-k-1}{\Longrightarrow} \quad &= \quad \frac{1}{\sqrt{2}} \sum_n (-1)^{-n-1} \overline{h_{-n-1}} e^{\frac{-jnw}{2}} \hat{\Phi}(\frac{w}{2}) \\ \Longrightarrow \quad &= \quad \frac{1}{\sqrt{2}} \sum_n g_n e^{\frac{-jnw}{2}} \hat{\Phi}(\frac{w}{2}) && (59) \end{aligned}$$

Taking the inverse Fourier transform for the above equation, we can have $\Psi(t)$ as follow:

$$\Psi(t) = \sum_n g_n \Phi(2t-n) \qquad (60)$$

Now we have the relationship between g_n and h_n as follow:

$$g_n = (-1)^{-n-1} \overline{h_{-n-1}} \qquad (61)$$

4 Creating a Wavelet Basis

The main goal in this section is to explain how to create a wavelet basis functions. We will present simple recipes to create orthonormal bases, above all, of practical use. For that, we divided the section in four parts. In the first part, we summarize the conditions presented in previous sections which are necessary to define a Multi-Resolution Analysis (MRA) of a wavelet basis function. Following that, we discuss the necessity of such Multi-Resolution Analysis and present two very useful bases - Haar and Meyer. These bases were originally proposed without using Multi-Resolution Analysis, but still, they satisfy the MRA conditions explained here. Our goal in this subsection is to demonstrate that, though not mandatory, MRA offers an easy way to design families of wavelet bases, that among other things, present compact support.

In the third section, we will analyze some relaxations that can be applied to the original constraints. These relaxations will lead us to an algorithm to find Ψ - The Cascade Algorithm - which will be finally presented in the last part of this section.

4.1 Multi-Resolution Analysis

From the discussion in the previous three sections, we can now summarize the conditions necessary to derive a Multi-Resolution Analysis (MRA) of a wavelet basis. Those conditions are:

The existence of a hierarchy of subspaces:

$$... \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset ...L^2(\mathcal{R}) \tag{62}$$

A "ladder" relationship between the subspaces:

$$f(.) \in V_j \Longleftrightarrow f(2^j.) \in V_0 \tag{63}$$

Invariance to integer displacement:

$$f(.) \in V_0 \Longrightarrow f(. - n) \in V_0 \tag{64}$$

Existence of an orthonormal basis function that spans V_0

$$\{\Phi_{0,n}; n \in \mathcal{Z}\} \tag{65}$$

Some of these conditions may be expressed in another form, more useful in future derivations. For example, Eqs. (39) and (42), reproduced below, are equivalent to Eq. (65) above, and will be used instead when we refer to the orthonormality property.

$$\sum_n |\hat{\Phi}(w + 2n\pi)|^2 = \frac{1}{2\pi}$$

$$|m_0(\frac{w}{2})|^2 + |m_0(\frac{w}{2}+\pi)|^2 = 1$$

What is important to mention here is that if all these conditions are satisfied, we have a ladder of closed subspaces $(V_j)_{j\in\mathcal{Z}}$ and then there exists an associated orthornormal wavelet basis $\{\Psi_{j,k}\}$ such that

$$P_{j-1} \quad = \quad P_j + \sum_k <., \Psi_{j,k}(t) > \Psi_{j,k}(t) \tag{66}$$

with

$$\Psi(t) \quad = \quad \sqrt{2}\sum_n g_n \Phi(2t-n) \tag{67}$$

and

$$g_n = (-1)^n \overline{h_{-n+1}} \tag{68}$$

$$h_n \quad = \quad <\Phi, \Phi_{-1,n}>, \quad \sum_{n\in\mathcal{Z}} |h_n|^2 = 1 \tag{69}$$

With the equations above, we can state our first recipe to create a wavelet basis:

1. Choose an appropriate Φ,
2. Find h_n from (69) above, and then g_n,
3. Finally, use g_n and Φ to find Ψ

In order to apply this recipe, first, we need to define "appropriate". An appropriate Φ is an orthonormal fuction satisfying Eq. (65), which can be very difficult to ascertain. Besides, it is usually desirable that Φ be compactly supported (non-zero only in a small interval), otherwise, we would have the similar problems of non-locality found in the Fourier Analysis. In case Φ is compactly supported, we can find a finite number of h_n and hence a finite number of Ψ's. Otherwise, we need to find analytic expressions for both, which may not be easy to compute, since it would envolve the inner product of infinitely many Φ – to obtain h_n – and the summation of infinitely many Φ – that leads to Ψ.

Remarks:

1. From the derivation that led to Eq. (67), we can conclude that Ψ is not unique for a given Φ. In fact, any $\hat{\Psi}' = \rho(w)\ \hat{\Psi}$, with ρ being 2π-periodic and $|\rho| = 1$, would be valid too.

4.2 Why MRA?

The question that arises from the difficulties explained in the previous subsection is: Do we need a MRA to find an orthonormal wavelet basis or there is another way to find it?

The answer to this is: No, we don't need MRA. Despite the fact that every orthonormal wavelet basis of practical interest known to this date is associated with a MRA, there are many "pathological" Ψ's that are not even derivable from a MRA (they cannot be explained by MRA).

One such case is due to J. L. Journé and can be found in Daubechies' book (include the example here in the future). In this example, $\hat{\Psi}$ is chosen not continuous, which is in general the easiest way to obtain "non-MRA" basis. Another example is the Mexican Hat, in the introduction of this tutorial. Such $\Psi_{j,k}$ constitutes an orthonormal basis for $L^2(\mathcal{R})$, but there is no known Φ associated with it.

Other than these "pathological" cases, there are two very important wavelet bases that were created before MRA was proposed, but they both are derivable from a MRA. The first is the Haar wavelet and is the most simple case of study. The second one is the Meyer Wavelet and it is a wavelet basis with very good properties: smoothness and fast decay. However, the derivation of the Meyer wavelet without resorting to MRA is definitely even more complicated and some steps were regarded as "magical". It was only after the MRA was applied to the Meyer's Φ that they could be fully (and simply) understood.

Next, we will present both the Haar and Meyer Wavelets.

4.2.1 Haar Wavelet

As we mentioned before, the Haar Wavelet is the most simple one. Yet, since it is compactly supported, the application of the recipe we gave by the end of 4.1 becomes trivial.

In the Haar Wavelet,

$$\Phi(x) = \begin{cases} 1 & : \quad 0 \leq x \leq 1 \\ 0 & : \quad otherwise \end{cases}$$

From the recipe, Eq. (69), and the fact that Φ is compactly supported,

$$h_n = \sqrt{2} \int dx \Phi(x) \overline{\Phi(2x-n)} = \begin{cases} \frac{1}{\sqrt{2}} & : \quad n = 0, 1 \\ 0 & : \quad otherwise \end{cases}$$

Again, from our recipe, Eqs. (68) and (67),

$$\Psi(x) = \frac{1}{\sqrt{2}} \Phi_{-1,0} - \frac{1}{\sqrt{2}} \Phi_{-1,0} = \begin{cases} 1 & : \quad 0 \leq x \leq \frac{1}{2} \\ -1 & : \quad \frac{1}{2} \leq x \leq 1 \\ 0 & : \quad otherwise \end{cases}$$

Though the Haar Wavelet is very simple to compute, for being not continuous, its Fourier Transform presents a very poor decay: $|w|^{-1}$, which means that it has a bad frequency localization (no better than the WFT, for what that matters).

4.2.2 Meyer Wavelet

For the Meyer Wavelet, the application of the recipe is done in the frequency space. Therefore, given:

$$\hat{\Phi}(w) = \begin{cases} (2\pi)^{-\frac{1}{2}} & : \quad w \leq \frac{2\pi}{3} \\ (2\pi)^{-\frac{1}{2}} \cos[\frac{\pi}{2}\nu(\frac{3}{2\pi}|w| - 1)] & : \quad \frac{2\pi}{3} \leq w \leq \frac{4\pi}{3} \\ 0 & : \quad otherwise \end{cases}$$

And since Φ is compactly supported, we obtain,

$$\begin{aligned} \hat{\Psi}(w) &= e^{\frac{iw}{2}} \overline{m_0(\frac{w}{2} + \pi)} \hat{\Phi}(\frac{w}{2}) \\ &= \sqrt{2\pi} e^{\frac{iw}{2}} \sum_{n \in \mathcal{Z}} \hat{\Phi}(w + 2\pi(2n+1)) \hat{\Phi}(\frac{w}{2}) \\ &= \sqrt{2\pi} e^{\frac{iw}{2}} [\hat{\Phi}(w + 2\pi) + \hat{\Phi}(w - 2\pi)] \hat{\Phi}(\frac{w}{2}) \end{aligned}$$

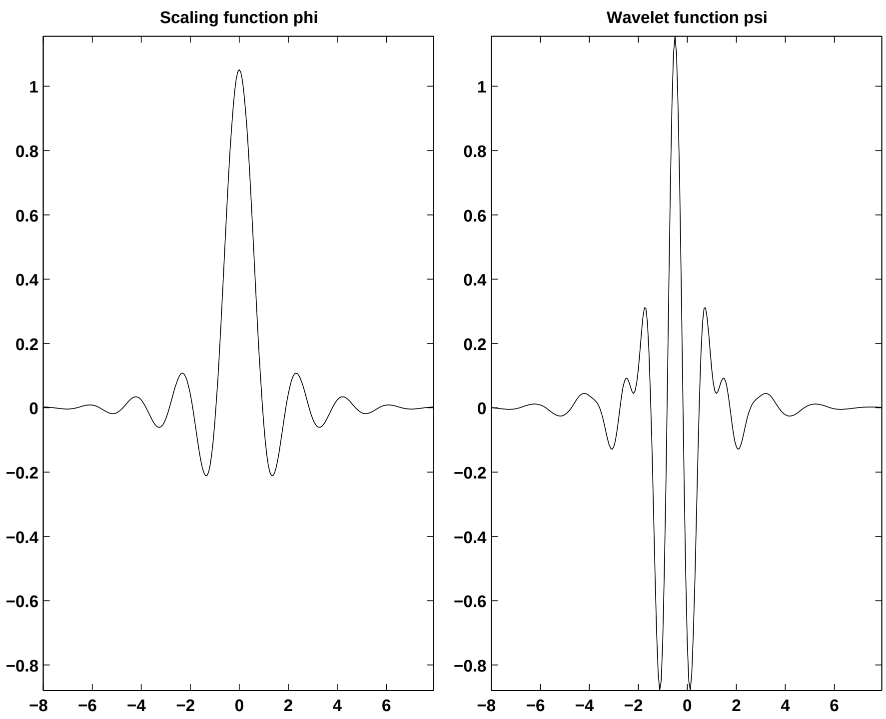

Figure 9: *Meyer Wavelet*

4.3 Relaxing the Constraints

4.3.1 Riesz vs. Orthonormal Bases

As we showed in subsection 4.1, to apply the recipe, we need first to choose a Φ and prove it satisfies all the constraints in Eqs. (62)-(65). This can be a very tedious process, unless we can relax some of these constraints.

The constraints (62)-(64) define a ladder of subspaces that is central to the MRA, so there is not much that can be done here. However, for Eq. (65), it can be proven that Φ does not need to constitute an orthonormal basis, but rather, it only needs to constitute an Riesz basis. Once we choose such Φ, we can construct an orthornormal basis, Φ', derived from Φ. This is called orthonormalization, and can be thought as analogous to the orthonormalization of vectors in vector spaces.

A Riesz basis for V_0 is a set $\{\Phi(t-k)\}$ for which the following equation can be satisfied.

$$A \sum_k |c_k|^2 \leq || \sum_k c_k \Phi(. - k)||^2 \leq B \sum_k |c_k|^2 \tag{70}$$

where A > 0, B $< \infty$ and $(c_k)_{k \in \mathcal{Z}} \in (\mathcal{Z})$

But, from Eq. (38),

$$\begin{aligned} || \sum_k c_k \Phi(. - k)||^2 &= \int dw | \sum_k c_k e^{-ikw} \hat{\Phi}(w)|^2 && (71) \\ &= \int_0^{2\pi} dw | \sum_k c_k e^{-ikw}|^2 \sum_l |\hat{\Phi}(w + 2\pi l)|^2 && (72) \end{aligned}$$

and

$$\sum_k |c_k|^2 = (2\pi)^{-1} \int_0^{2\pi} dw | \sum_k c_k e^{-ikw}|^2 \tag{73}$$

so that Eq. (70) is equivalent to

$$(2\pi)^{-1} A \leq \sum_l |\hat{\Phi}(w + 2\pi l)|^2 \leq (2\pi)^{-1} B \tag{74}$$

and we can derive an orthonormal basis Φ' by applying the orthonormalization trick:

$$\hat{\Phi}'(w) = (2\pi)^{-\frac{1}{2}} \left[\sum_l |\hat{\Phi}(w + 2\pi l)|^2 \right]^{-\frac{1}{2}} \hat{\Phi}(w) \tag{75}$$

This means that if we start from a Riesz basis, Φ, we can follow another recipe to create a wavelet basis, Ψ:

1. Choose a Φ with reasonable decay,
2. Apply the "orthonormalization trick", as in Eq. (75), to obtain Φ',
3. Find h_n from (69), and then g_n, but using Φ' instead of Φ
4. Finally, use g_n and Φ' to find Ψ'

An example of a basis created using this recipe is the Battle-Lemarié family. (include the example here)

But this recipe doesn't sound much simpler than the previous one, since we still need to prove (though it's easier than before) that Φ is a Riesz basis. Besides, from Eq. (75), we can see that no matter how we choose Φ, the support Φ' will always be the entire $\mathcal{R}$. As well as the support of Ψ'.

Remarks:

1. For some wavelet bases, such as Battle-Lemarié, both Φ' and Ψ' have a very good (exponential) decay.
2. The Meyer wavelet, is very "smooth" ($\in C^\infty$), and decays faster than any inverse polynomial, but it's not exponential.
3. In general, there is a trade-off between "smoothness" and fast decay, and Ψ can not be both smooth and have a fast decay. In other words, we have to choose for exponential decay in either time or frequency. In practice, time is preferred over frequency.

4.3.2 Compactly Supported Orthonormal Bases

The process of choosing Φ, applying the orthonormalization trick and proving that it defines a MRA (satisties Eqs. (62)-(65)) doesn't seem simple enough, specially when it comes to writing a program to compute the wavelet transform of a signal. Besides that, as we showed in the previous subsection, it creates an infinitely supported basis function.

In this subsection, we will discuss another method for creating wavelet bases that start from m_0, Eq. (42), rather than from Φ or V_j as before. With this method, we can relax the orthonormality constraint, Eq. (65), so that we concentrate only on compactly supported orthonormal bases. Although those wavelet bases cannot, in general, be written in a closed analytic form, the method itself suggests an algorithm - Cascade Algorithm - that can be applied to numerically compute Ψ with arbitrarily high precision.

As we showed in other examples of wavelet bases, if we make Φ compactly supported,

$$h_n = < \Phi, \Phi_{-1,n} > \tag{76}$$

$$= \sqrt{2} \int dx \Phi(x) \overline{\Phi(2x-n)} \tag{77}$$

h_n will have only finitely many nonzeros, and Ψ reduces to a finite linear combination of compactly supported functions Φ.

By the end of last subsection, we claimed in Remark 3, that there exists a trade-off between exponential decay in either time or frequency. This claim is due to a theorem that we state without proof herein:

If the $\Psi_{j,k}(x)$ constitute an orthonormal set in $L^2(\mathcal{R})$, with $|\Psi(x)| \leq C(1+|x|)^{-m-1-\varepsilon}$, $\Psi \in C^m(\mathcal{R})$ and $\Psi^{(l)}$ bounded for $l \leq m$, then $\int x^l\ \Psi(x)dx = 0$ for $l = 0, 1, ..., m$

In plain english, this means that the number of central moments, m, that $\Psi_{j,k}$ possesses - therefore the steepness with which $\Psi_{j,k}$ goes to zero - also determines the smoothness of Ψ ($\Psi \in C^m(\mathcal{R})$).

The claim in Remark 3 itself is a Corollary of the above theorem and basically states that if $m = \infty$ (exponential decay), then $\Psi \equiv 0$.

Another corollary of the same theorem, and more important here, states that

Assume that the $\Psi_{j,k}(x)$ constitute an orthonormal basis of wavelets, assiciated with a multi-resolution analysis. If $|\Phi|$, $|\Psi| \leq C(1+|x|)^{-m-1-\varepsilon}$ and $\Psi \in C^m(\mathcal{R})$ with $\Psi^{(l)}$ bounded for $l \leq m$, then m_0, as defined in Eq. (34), factorizes as

$$m_0(w) = \left(\frac{1+e^{-iw}}{2}\right)^{m+1} L(w) \tag{78}$$

where L is 2π-periodic.

In other words, m_0 becomes a trigonometric polynomial and, from Eq. (42) and the orthonormality requirement of $\Phi_{0,n}$, we have:

$$|m_0(w)|^2 + |m_0(w+\pi)|^2 = 1 \tag{79}$$

Let's rewrite $|m_0(w)|^2$ with,

$$M_0(w) = |m_0(w)|^2 \tag{80}$$

and since L is also a trigonometric polynomial (for convenience in $\sin(w)$), we can rewrite M_0 as

$$M_0(w) = \left(\cos^2\frac{w}{2}\right)^N P\left(\sin^2\frac{w}{2}\right) \tag{81}$$

Now, making $y = \cos^2 \frac{w}{2}$ and $\sin^2 \frac{w}{2} = (1 - y)$, and performing some algebraic simplifications, we can write equation (79) above as

$$(1-y)^N P(y) + y^N P(1-y) = 1 \tag{82}$$

Finally, due to Bezout, we state as a Lemma:

If p_1, p_2 are two polynomials, of degree n_1, n_2 respectively, with no common zeros, then there exist unique polynomials q_1, q_2, of degree $n_2 - 1$, $n_1 - 1$, respectively, so that

$$p_1(x)q_1(x) + p_2(x)q_2(x) = 1 \tag{83}$$

Applied to our case above, it implies that there exist two polynomials q_1, q_2, of degree $\leq N - 1$, so that

$$(1-y)^N q_1(y) + y^N q_2(y) = 1 \tag{84}$$

or, replacing $y \Longleftrightarrow 1 - y$

$$(1-y)^N q_2(1-y) + y^N q_1(1-y) = 1 \tag{85}$$

But the uniqueness of q_1, q_2 thus implies that $q_2(y) = q_1(1-y)$ and therefore, we only need to find q_1 such that

$$q_1(y) = (1-y)^{-N} \left[1 - y^N q_1(1-y)\right] \tag{86}$$

The Taylor expansion for $(1-y)^{-N}$ is given by

$$q_1(y) = \sum_{k=0}^{N-1} \binom{N+k-1}{k} y^k + O(y^N) \tag{87}$$

But from the Lemma above, we know that the degree of $q_1 \leq N - 1$. Therefore, q_1 has to be equal to its Taylor exapansion truncated after N terms, or

$$q_1(y) = \sum_{k=0}^{N-1} \binom{N+k-1}{k} y^k \tag{88}$$

Let's denote this solution for q_1 by P_N. We know that P_N is the unique lowest degree solution, but there exist many solutions of higher degree. For any such higher degree solution, we have

$$(1-y)^N[P(y)-P_N(y)]+y^N[P(y-1)-P_N(y-1)]=0 \tag{89}$$

which already implies that $P-P_N$ is divisible by y^N,

$$P(y)-P_N(y)=y^N\tilde{P}(y) \tag{90}$$

and

$$\tilde{P}(y)+\tilde{P}(1-y)=0 \tag{91}$$

i.e. $\tilde{P}$ is antisymmetric with respect to $\frac{1}{2}$

To summarize all our findings so far, we have:

A trigonometric polynomial m_0 of the form

$$m_0(w)=\left(\frac{1+e^{-iw}}{2}\right)^{m+1}L(w) \tag{92}$$

satisfies Eq. (79) if and only if $|L(w)|^2$ can be written as

$$|L(w)|^2=P\left(\sin^2\frac{w}{2}\right) \tag{93}$$

with

$$P(y)=P_N(y)-y^NR(\frac{1}{2}-y) \tag{94}$$

where

$$P_N(y)=\sum_{k=0}^{N-1}\binom{N+k-1}{k}y^k \tag{95}$$

and R is an odd polynomial, chosen such that $P(y)\leq 0$ for $y\in[0,1]$

Now we are almost ready for our final recipe. However, we have characterized $|m_0(w)|^2$ and we still need to "extract the square root" of it. This procedure is called spectral factorization in the engineering literature, and all we need is to state another Lemma.

Let A be a positive trigonometric polynomial where $A(w)=A(-w)$; A is necessarily of the form

$$A(w)=\sum_{m=0}^{M}a_m\cos mw \tag{96}$$

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

$$B(w) = \sum_{m=0}^{M} b_m e^{imw} \tag{97}$$

such that $|B(w)|^2 = A(w)$ *and* a_m, $b_m \in \mathcal{R}$

The consequence of this is that given $|L(w)|^2$ factorized as:

$$\begin{aligned} |L(w)|^2 &= \left[\frac{1}{2}|a_M| \prod_{j=1}^{J} |z_j|^{-2} \prod_{l=1}^{L} |r_l|^{-1}\right] \left|\prod_{j=1}^{J} (e^{-iw} - z_j)(e^{-iw} - \overline{z_j})\right|^2 \\ & \left|\prod_{k=1}^{K} (e^{-iw} - e^{i\alpha_k})(e^{-iw} - e^{-i\alpha_k})\right|^2 \left|\prod_{l=1}^{L} (e^{-iw} - r_l)\right|^2 \end{aligned} \tag{98}$$

We can obtain $L(w)$ as

$$\begin{aligned} L(w) &= \left[\frac{1}{2}|a_M| \prod_{j=1}^{J} |z_j|^{-2} \prod_{l=1}^{L} |r_l|^{-1}\right]^{\frac{1}{2}} \left|\prod_{j=1}^{J} (e^{-iw} - z_j)(e^{-iw} - \overline{z_j})\right| \\ & \left|\prod_{k=1}^{K} (e^{-iw} - e^{i\alpha_k})(e^{-iw} - e^{-i\alpha_k})\right| \left|\prod_{l=1}^{L} (e^{-iw} - r_l)\right| \end{aligned} \tag{99}$$

Now we can write our recipe as follows:

1. Choose the number of vanishing moments, N, for your wavelet basis Ψ.
2. Find P_N given by (95)
3. Choose R, such that $P(y) \leq 0$ for $y \in [0, 1]$
4. Replace P_N and R in (94) to obtain P
5. Factorize $|L(w)|^2$ as in (98) and then in (99)
6. Replace $L(w)$ in (92) and compare it to (34). That give us h_n.

Let's now apply the recipe above to a simple example.

Step 1

$$N = 2 \tag{100}$$

Step 2

$$
\begin{aligned}
P_N(y) &= \sum_{k=0}^{1} \binom{k+1}{k} y^k \\
&= 1 + 2y
\end{aligned}
$$

Step 3

$$R \equiv 0$$

Step 4

$$
\begin{aligned}
P(y) &= P_N(y) \\
&= 1 + 2y
\end{aligned}
$$

Step 5

$$
\begin{aligned}
|L(w)|^2 &= P\left(\sin^2 \frac{w}{2}\right) \\
|L(w)|^2 &= 1 + 2\sin^2 \frac{w}{2} \\
&= 2 - \cos w \\
&= 2 - \frac{e^{iw} + e^{-iw}}{2} \\
&= -\frac{e^{iw}}{2}[1 - 4e^{-iw} + (e^{-iw})^2]
\end{aligned}
$$

$$
\begin{aligned}
|P_L(z)|^2 &= 1 - 4z + z^2 \\
&\Longrightarrow z_j = 2 \pm \sqrt{3}
\end{aligned}
$$

$$
L(w) = \left(\frac{1}{2}\,\frac{1}{2+\sqrt{3}}\right)^{\frac{1}{2}} \left(e^{-iw} - (2+\sqrt{3})\right)
$$

Step 6

$$
\begin{aligned}
m_0(w) &= \left(\frac{1+e^{-iw}}{2}\right)^2 L(w) \\
&= \frac{1}{4(4+\sqrt{3})^{\frac{1}{2}}}[(2+\sqrt{3}) + (3+2\sqrt{3})e^{-iw} + \sqrt{3}e^{-2iw} - e^{-3iw}]
\end{aligned}
$$

Comparing to

$$m_0(w) = \frac{1}{\sqrt{2}} \sum_k h_k e^{-iwk}$$

We obtain $h_0 = 0.482963$, $h_1 = 0.836516$, $h_2 = 0.224144$, $h_3 = -.129409$

5 Obtaining the wavelet coefficients S_j and d_j

In this section we will show how the wavelet coefficients of a signal can be calculated in accordance with Fig. 7. We will present a recursive approach here that will allow us to calculate the wavelet coefficients at each scale from the scaling wavelet coefficients at the previous scale.

The analyzer and the scaling wavelet coefficients at scale j are given by

$$\begin{aligned} D_{j,k} &= < f_{j-1}(t), \Psi_{j,k}(t) > && (101)\\ S_{j,k} &= < f_{j-1}(t), \Phi_{j,k}(t) > \quad , \ k = 0, 1, ..., N && (102) \end{aligned}$$

We will now show that both of these sets of coefficients can be obtained from $s_{j-1,k}$, $k = 0, 1, ..., N-1$. The Ψ wavelets at scale j are given by

$$\begin{aligned} \Psi_{j,k} &= 2^{\frac{-j}{2}} \Psi(2^{-j}t - k)\\ &= 2^{\frac{-j}{2}} \sum_n g_n 2^{\frac{1}{2}} \Phi(2^{-j+1}t - 2k - n)\\ &= 2^{\frac{-j+1}{2}} \sum_n g_n \Phi(2^{-j+1}t - (2k+n))\\ &= \sum_n g_n \Phi_{j-1,2k+n}(t)\\ &= \sum_n g_{n-2k} \Phi_{j-1,n}(t) && (103) \end{aligned}$$

Therefore,

$$\begin{aligned} D_{j,k} &= < f_{j-1}(t), \Psi_{j,k}(t) > && (104)\\ &= \sum_n \overline{g_{n-2k}} < f_{j-1}(t), \Phi_{j-1,n}(t) >\\ &= \sum_n \overline{g_{n-2k}} D_{j-1,n} \end{aligned}$$

Again, we apply the similar step to find the wavelet coefficients for the smooth part of the signal. from Eq. (32), we know the scaling wavelet can be expressed as a constraint equation which can be used to express the $\Phi_{j,k}$ as follow:

$$\Phi_{j,k} = 2^{\frac{-j}{2}} \Phi(2^{-j}t - k)$$

$$
\begin{aligned}
&= 2^{\frac{-j}{2}} \sum_n h_n 2^{\frac{1}{2}} \Phi(2^{-j+1}t - 2k - n) \\
&= 2^{\frac{-j+1}{2}} \sum_n h_n \Phi(2^{-j+1}t - (2k+n)) \\
&= \sum_n h_n \Phi_{j-1,2k+n}(t) \\
&= \sum_n h_{n-2k} \Phi_{j-1,n}(t) \qquad (105)
\end{aligned}
$$

The smooth part wavelet coefficients can be extracted by the following:

$$
\begin{aligned}
S_{j,k} &= < f_{j-1}, \Phi_{j,k} > \qquad (106) \\
&= \sum_n \overline{h_{n-2k}} < f_{j-1}, \Phi_{j-1,n} > \\
&= \sum_n \overline{h_{n-2k}} s_{j-1,n}
\end{aligned}
$$