THE FRAMEWORK OF A DYADIC WAVELETS ANALYSIS

By D.S.G. POLLOCK

University of Leicester

Email: stephen_pollock@sigmapi.u-net.com

This tutorial paper analyses the structure of a discrete dyadic wavelet analysis in a manner that facilitates its computation. A clear connection is maintained between the continuous function that is the object of the analysis and the discrete sequences that are its products.

The Dyadic Decomposition of a Space of Functions

A discrete wavelet analysis is based on the supposition that the elements of the data sequence $\{y_k; k = 0, 1, 2, \ldots, T-1\}$ have been sampled from a continuous function $f(t)$ with $t \in [0, T)$. It is presumed that the function can be reconstituted, to some degree of approximation, by associating a scaling function kernel or father wavelet $\phi_{0,k} = \phi(t-k)$ to each of these ordinates and by summing the result:

$$f(t) \simeq \sum_{k=0}^{T-1} y_k \phi(t-k). \tag{1}$$

The scaling functions are designed to constitute an orthonormal basis of the space $\mathcal{V}_0$ in which the function $f(t)$, or its approximation, resides. Therefore,

$$\int_t \phi(t-j)\phi(t-k)dt = \langle \phi(t-j), \phi(t-k) \rangle = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases} \tag{2}$$

Then, the elements of the data sequence, which are the amplitude coefficients of the associated scaling functions, are given by

$$y_k = \int_t f(t)\phi(t-k)dt = \langle f(t), \phi(t-k) \rangle. \tag{3}$$

The basis $\phi_{0,k} = \phi(t-k); k = 0, 1, \ldots, T-1$, which is ordered in time, may be described as the initial basis of scaling functions. Scaling functions have nominal frequency contents that extend from a limiting frequency down to the zero frequency.

In a dyadic wavelets analysis, the T amplitude coefficients of equation (1), which are associated with the initial basis, and which are the sampled values, are transformed into a hierarchy of T coefficients that are associated with an alternative basis, which is ordered both according to the temporal locations of

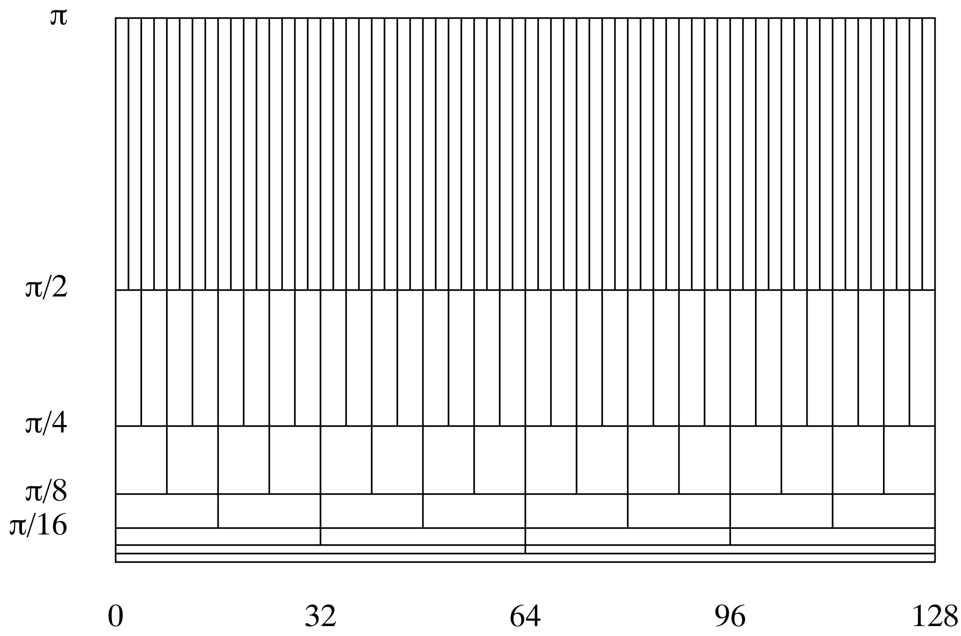

Figure 1. The partitioning of the time–frequency plane according to a multiresolution analysis of a data sequence of $128 = 2^7$ points.

the wavelets and according to their frequency contents. This constitutes the final basis.

The hierarchy of wavelets within the final basis can be described with reference to a so-called mosaic diagram that defines a partitioning of the time-frequency plane, which corresponds to the space $\mathcal{V}_0$. This is illustrated for a sample of size $T = 128 = 2^7$ by Figure 1. In the figure, the height of a cell corresponds to a bandwidth in the frequency domain, whereas its width denotes a temporal duration.

The highest frequency in the mosaic diagram is the Nyquist frequency of π radians per sample interval, which represents the maximum frequency that is detectable via the process of discrete sampling. Centred on each cell, but liable to extend beyond its temporal boundaries, there is a wavelet. The frequency contents of the wavelet is also liable to extend beyond the nominal bandwidth that is indicated in the figure.

Apart from the final cell, which stretches across the width of the diagram and which is bounded by the zero frequency, the cells within mosaic diagram and bounded above and below by positive frequencies. These cells are occupied by mother wavelets, which have a different form from that of the scaling functions.

The horizontal bands of the mosaic diagram are obtained by successive divisions of the frequency range. First, the range of frequencies $[0, \pi]$ of the space $\mathcal{V}_0$ is divided into the equal subintervals $[0, \pi/2]$ and $(\pi/2, \pi]$. The upper frequency interval will have $T/2$ wavelet functions, denoted by $\psi_{1,k}(t); k = 0, 1, \ldots, [T/2] - 1$, separated one from the next by two time intervals. These wavelets will constitute a basis for a space denoted by $\mathcal{W}_1$.

The lower frequency interval will have the same number $T/2$ of scaling functions, denoted by $\phi_{1,k}(t); k = 0, 1, \ldots, [T/2] - 1$, also separated by two intervals. These scaling functions will constitute a basis for a space denoted by $\mathcal{V}_1$. The division of $\mathcal{V}_0$ is such that its two subspaces $\mathcal{W}_1$ and $\mathcal{V}_1$ are mutually orthogonal. Their orthogonality, which can be denoted by writing $\mathcal{V}_1 \perp \mathcal{W}_1$, entails the fact that $\mathcal{V}_1 \cap \mathcal{W}_1 = 0$.

The direct sum of the two subspaces is $\mathcal{W}_1 \oplus \mathcal{V}_1 = \mathcal{V}_0$. This means that any element in $f(t) \in \mathcal{V}_0$ can be expressed as $f(t) = w_1(t) + v_1(t)$ with $w_1(t) \in \mathcal{W}_1$ and $v_1(t) \in \mathcal{V}_1$, which is the sum of two orthogonal functions.

In the next stage of the decomposition of $\mathcal{V}_0$, the lower interval is subdivided into the intervals $[0, \pi/4]$ and $(\pi/4, \pi/2]$, which are filled, respectively, with $T/4$ scaling functions, denoted by $\phi_{2,k}(t); k = 0, 1, \ldots, [T/4] - 1$, and $T/4$ wavelets, denoted by $\psi_{2,k}(t) =; k = 0, 1, \ldots, [T/4] - 1$, separated by four time intervals. These will constitute the basis functions, respectively, of the spaces $\mathcal{V}_2$ and $\mathcal{W}_2$, which are mutually orthogonal subspaces of $\mathcal{V}_1$

The process of subdivision continues, by dividing successively the lower subintervals, until it can go no further. If there are $T = 2^n$ points in the sample, then T can be divided n times, and there will be a total of $n + 1$ horizontal bands, with the cells of all but the final band filled with wavelets. The final band will contain a single scaling function.

The process of subdivision generates a nested sequence of vector spaces, each of which is spanned by a set of scaling functions:

$$\mathcal{V}_0 \supset \mathcal{V}_1 \supset \cdots \supset \mathcal{V}_n. \tag{4}$$

The jth stage of the process, which generates $\mathcal{V}_j$, also generates the accompanying space $\mathcal{W}_j$ of wavelet functions, which is its orthogonal complement within $\mathcal{V}_{j-1}$. The complete process can be summarised by displaying the successive decompositions of $\mathcal{V}_0$:

$$\begin{aligned} \mathcal{V}_0 &= \mathcal{W}_1 \oplus \mathcal{V}_1 \\ &= \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \mathcal{V}_2 \\ &\ \vdots \\ &= \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \cdots \oplus \mathcal{W}_n \oplus \mathcal{V}_n. \end{aligned} \tag{5}$$

The elements the final expression correspond to the successive horizontal bands of the mosaic diagram.

Given the decomposition of $\mathcal{V}_0$ as a sum of mutually orthogonal subspaces, represented by equation (5), and given that $f(t) \in \mathcal{V}_0$, it is possible to represent the function $f(t)$ as a sum of orthogonal components residing in the subspaces. Thus

$$f(t) = w_1(t) + w_2(t) + \cdots + w_n(t) + v_n(t), \tag{6}$$

with $w_j(t) \in \mathcal{W}_j$ for $j = 1, \ldots, n$ and with $v_n(t) \in \mathcal{V}_n$.

The generic component of this decomposition may be represented, relative to the basis functions $\psi_{j,k}(t); k = 0, 1, \ldots, [T/2^j] - 1$ of $\mathcal{W}_j$, by

$$w_j(t) = \sum_{k=0}^{[T/2^j]-1} \beta_{jk}\psi_{j,k}(t). \tag{7}$$

Here, β_{jk} is the amplitude coefficient of the kth wavelet function. The two final elements of the decomposition of (6) are the wavelet function $w_n(t) = \beta_{n0}\psi_{n,0}(t)$ and the scaling function $v_n(t) = \gamma_{n0}\phi_{n,0}(t)$, which have been scaled by the amplitude coefficients β_{n0} and γ_{n0}, respectively.

Given that the function $f(t)$ is represented, in practice, by its T sampled ordinates $y_k; k = 0, 1, 2, \ldots, T-1$, it is appropriate to express the components of the decomposition of (6) in terms of their ordinates sampled at the integer points. Thus, one purpose of the wavelet analysis is to generate the expression

$$y_k = w_{1k} + w_{2k} + \cdots + w_{nk} + v_{nk}; \quad k = 0, 1, \ldots, T-1, \tag{8}$$

which is the discrete-time counterpart of equation (6). This leads to the following expression for the continuous function:

$$f(t) = \sum_k y_k \phi_{0,k}(t) = \sum_{k=0}^{T-1} \Big\{ \sum_{j=1}^{n} w_{jk} + v_{nk} \Big\} \phi_{0,k}(t). \tag{9}$$

Given the decomposition of (8) and (9), it is possible to perform various operations that are designed to enhance the representation of the underlying signal.

A common purpose is to remove from $f(t)$ the traces of an additive noise contamination. If the noise resides within a limited set of wavelets bands, which are liable to be the high-frequency bands, then the signal can be enhanced by removing the corresponding components from the sum.

If the only part of the signal that is of interest resides within a limited set of adjacent bands, then it can be isolated in a straightforward way by forming the partial sum of the corresponding components.

Another common purpose in a wavelets analysis is to achieve a measure of data compression. If the absolute value of the amplitude coefficient β_{jk} associated with the wavelet basis function $\psi_{j,k}(t) \in \mathcal{W}_j$ is below a predetermined level of significance, then it can be set to zero. In this way, it may become possible to convey the essential information of the signal in far fewer than the T coefficients that are present in equation (1).

The Dilation Equations

The scaling functions and the wavelets in successive bands represent dilated or stretched versions of the functions in the bands above. Let $\mathcal{V}_0$ be the space spanned by the scaling functions $\phi_{0,k}(t) = \phi(t-k)$, which constitute an orthonormal basis, and let $\mathcal{V}_1 \subset \mathcal{V}_0$ be the subspace containing functions at half the resolution. Then, $\mathcal{V}_1$ will be spanned by the basis functions $\phi_{1,k}(t)$, which represent versions of the functions $\phi_{0,k}(t) = \phi(t-k)$ that have been stretched by a factor of 2. That is to say that, if $\phi_{0,k}(t)$ is supported on a finite interval, then $\phi_{1,k}(t)$ will be supported on an interval of double the length.

The relationship between the two sets of functions is such that

$$\phi_{1,k}(t) = 2^{-1/2}\phi(2^{-1}t - k). \tag{10}$$

Replacing t by $2^{-1}t$ means that a basis function of $\mathcal{V}_1$ will evolve at half the rate of a basis function of of $\mathcal{V}_0$. Multiplying the functions by the factor $2^{-1/2}$ ensures that their squares will continue to integrate to unity, which is a necessary normalisation. This follows from observing that

$$\int \phi_{1,0}^2(t)dt = \frac{1}{2}\int \phi^2(2^{-1}t)dt = \frac{1}{2}\int \phi^2(\tau)\frac{dt}{d\tau}d\tau = 1, \tag{11}$$

where $t = 2\tau$ and $dt/\tau = 2$, and where the integral of $\phi^2(\tau)$ is unity in consequence of (2).

Also observe that the basis functions of $\mathcal{V}_1$ are separated one from the next by intervals of 2 points. Thus, whereas $\phi_{1,k}(t) = 2^{-1/2}\phi(2^{-1}t - k)$ will have its centre at the point $t = 2k$, which is the solution of $2^{-1}t - k = 0$, $\phi_{1,k+1}(t)$ will have its centre at the point $t = 2k + 2$.

Equation (10) may be generalised to give

$$\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k), \tag{12}$$

which is a basis function of $\mathcal{V}_j$. It should be noted that, whereas the present notation has $\mathcal{V}_1 \subset \mathcal{V}_0$, it is common to reverse the order of the indices so that the space of higher dimension acquires the higher index.

It is possible to express the scaling function $\phi_{1,0}(t) \in \mathcal{V}_1$ as a linear combination of the elements of the basis of a space $\mathcal{V}_0$ of twice the resolution. The appropriate expression is

$$\phi_{1,0}(t) = 2^{-1/2}\phi(2^{-1}t) = \sum_k g_k\phi(t - k), \tag{13}$$

were

$$g_k = \langle\phi_{1,0}(t), \phi(t - k)\rangle = \int_{-\infty}^{\infty} \phi_{1,0}(t)\phi(t - k)dt \tag{14}$$

Equation (13) is the so-called dilation equation of the scaling function. The coefficients g_k of the dilation equation are also the coefficients of a lowpass filter. More generally, the relationship between the basis elements of $\mathcal{V}_j$ and those of $\mathcal{V}_{j-1}$ is indicated by

$$\phi_{j,0}(t) = \sum_k g_k\phi_{j-1,k}(t). \tag{15}$$

Various conditions must be imposed on the coefficients of the dilation equations. The first condition concerns the sum of the coefficients, which must be

$$\sum_k g_k = 2^{1/2}. \tag{16}$$

The necessity of this condition is established by integrating both sides of equation (13). Assuming that the integration and the summation can be commuted, this gives

$$2^{-1/2}\int \phi(2^{-1}t)dt = 2^{-1/2}\int \phi(\tau)\frac{dt}{d\tau}d\tau = \sum_k g_k \int \phi(t - k)dt. \tag{17}$$

Here, the variable of integration has been changed from t to $\tau = t/2$, which accounts for the factor $dt/d\tau = 2$. The value of the integral of $\phi(t-k)$ is independent of the translation k, which may be set to zero. Therefore, equation (17) leads directly to equation (16). This condition is regardless any orthogonality conditions that are imposed on the wavelets.

The next condition is that the sum of squares of the coefficients of the dilation is unity:

$$\sum_k g_k^2 = 1. \tag{18}$$

This follows from the fact that the scaling functions at all levels constitute orthonormal bases. Thus, at level 1, there is

$$\begin{aligned} 1 &= \int_t \phi_{1,0}^2(t)dt = \int_t \Big\{ \sum_k g_k \phi(t-k) \Big\}^2 dt \\ &= \sum_j \sum_k g_j g_k \int_t \phi(t-j)\phi(t-k)dt = \sum_k g_k^2, \end{aligned} \tag{19}$$

where the final equality follows from the conditions of orthonormality of (2).

A further important condition affecting the coefficients is that

$$\sum_k g_k g_{k+2m} = 0. \tag{20}$$

This also follows from the orthogonality of the scaling functions and from the dilation equation. The orthogonality of any two separate scaling functions at level 0 implies that, if $m \neq 0$, then, in view or (13), there is

$$\begin{aligned} 0 &= \int_t \phi(2^{-1}t)\phi(2^{-1}t-m)dt \\ &= 2\int_t \sum_j \sum_k g_j g_k \phi(t-j)\phi([t-2m]-k)dt \\ &= 2\sum_j \sum_k g_j g_k \int_t \phi(t-j)\phi(t-[2m+k])dt. \end{aligned} \tag{21}$$

The integral within the final expression will be zero-valued unless $j = k+2m$. In that case, the expression will deliver the term $2g_k g_{k+2m}$, which must be equal to zero. It follows that equation (20) is a necessary condition for the orthogonality of the scaling functions.

The orthogonal complement within $\mathcal{V}_0$ of the space $\mathcal{V}_1$ of scaling functions is the space $\mathcal{W}_1$ of wavelets functions. The $\mathcal{W}_1$ space is spanned by the wavelets functions $\psi_{1,k}(t); k = 0, 1, \ldots, [T/2]-1$, which constitute an orthonormal basis. Since $\mathcal{W}_1 \subset \mathcal{V}_0$, it is possible to express the wavelet function $\psi(t)$ as a linear combination of the elements of the basis of $\mathcal{V}_0$. The appropriate expression is

$$\psi_{1,0}(t) = 2^{-1/2}\psi(2^{-1}t) = \sum_k h_k \phi(t-k), \tag{22}$$

where

$$h_k = \langle \psi_{1,0}(t), \phi(t-k) \rangle = \int_{-\infty}^{\infty} \phi_{1,0}(t)\psi(t-k)dt \tag{23}$$

Equation (22) is the dilation equation of the wavelet function. The coefficients h_k of the equation are also the coefficients of a highpass filter that is complementary to the lowpass filter that entails the coefficients g_k.

More generally, there is

$$\psi_{j,0}(t) = \sum_k g_k \phi_{j-1,k}(t), \tag{24}$$

and, in parallel with equation (12), there is

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k). \tag{25}$$

The coefficients of the dilation equation of the wavelets must fulfil conditions that are equivalent to those that affect the scaling function dilation. Thus

$$p_0 = \sum_k h_k^2 = 1 \quad \text{and} \quad p_{2m} = \sum_k h_k h_{k+2m} = 0. \tag{26}$$

These are necessary conditions for the *sequential* orthogonality of separate wavelets within the band in question; and they represent restrictions on an autocovariance function. In addition, it is required that

$$\sum_k h_k = 0. \tag{27}$$

This is sufficient to ensure that the areas of the level-1 wavelets are zero.

Conditions must also be imposed to ensure that the wavelets are orthogonal to the scaling functions. To ensure that the scaling function $\phi(t)$ and the wavelet $\psi(t-m)$ that are at different displacements will be mutually orthogonal, it is sufficient to impose the condition that

$$\sum_k g_k h_{k+2m} = 0. \tag{28}$$

It is also necessary to ensure the orthogonality of wavelets and scaling functions that are at the same displacement. It is assumed that the two dilation equations contain the same number M of coefficients, and that this is an even number. Then, a sufficient condition for orthogonality is that

$$\sum_k g_k h_k = 0. \tag{29}$$

If the coefficients of the scaling function dilation equation are $g_0, g_1, \ldots, g_{M-1}$, then the conditions of (28) and (29) can be realised by setting

$$h_k = (-1)^k g_{M-1-k}, \quad \text{which implies that} \quad g_k = (-1)^{k+1} h_{M-1-k}. \tag{30}$$

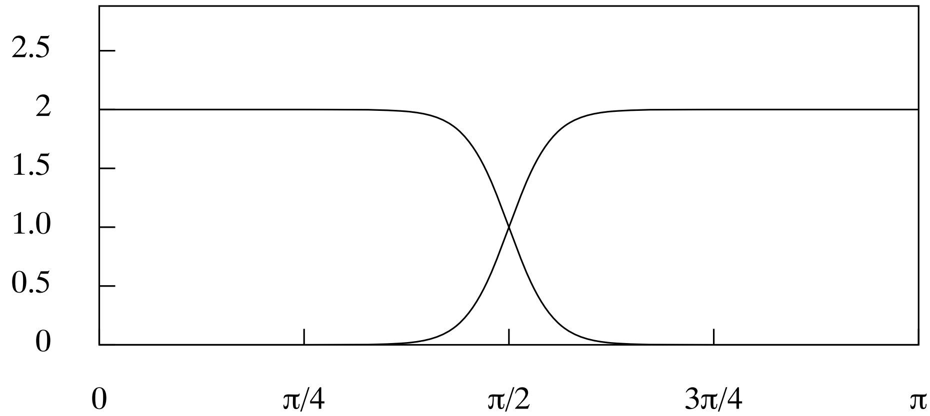

Figure 2. The squared gains of the complementary lowpass and highpass filters.

An example is provided by the case where $M = 4$. Then, there are

$$\begin{aligned} &g_0, \qquad && h_0 = g_3, \\ &g_1, && h_1 = -g_2, \\ &g_2, && h_2 = g_1, \\ &g_3, && h_3 = -g_0; \end{aligned} \tag{31}$$

and the conditiond of (28) and (29) are clearly satisfied.

It can be helpful to express these relationships in terms of the z-transforms of the coefficient sequences. These are

$$\begin{aligned} G(z) &= g_0 + g_1 z + g_2 z^2 + g_3 z^3 = z^3 H(-z^{-1}), \\ H(z) &= g_3 - g_2 z + g_1 z^2 - g_0 z^3 = -z^3 G(-z^{-1}). \end{aligned} \tag{32}$$

The autocovariance generating function formed from the coefficients of the dilation equation of the scaling function is

$$P(z) = G(z)G(z^{-1}). \tag{33}$$

The conditions of sequential orthogonality affecting the scaling function imply that the coefficients of $P(z)$ associated with the even powers of z must be zeros. The coefficients in question are comprised by the function $P(z) + P(-z)$, from which the odd powers of z are absent. On taking account of the condition that $p_0 = 1$, it can be seen that the condition for sequential orthogonality is that

$$\begin{aligned} P(z) + P(-z) &= G(z)G(z^{-1}) + G(-z)G(-z^{-1}) \\ &= G(z)G(z^{-1}) + H(z)H(z^{-1}) = 2. \end{aligned} \tag{34}$$

Equation (34) indicates the complementary nature of the highpass and lowpass filters that are derived from the coefficients of the dilation equations. Setting $z = \exp\{\mathrm{i}\omega\}$ with $\omega \in [-\pi, \pi]$ within $H(z)H(z^{-1})$ and $G(-z)G(-z^{-1})$

gives the squared gains of the filters. These are plotted in Figure 2 for the case of the Daubechies D4 filters that are to be specified in the section that follows.

The cross-covariance generating function formed from the coefficients of the highpass and lowpass filters is $Q(z) = G(z)H(z^{-1})$. The condition of (28), which imposes the mutual orthogonality of the wavelets and the scaling functions at displacements that are multiples of two points, is equivalent to the condition that

$$Q(z) + Q(-z) = G(z)H(z^{-1}) + G(-z)H(-z^{-1}) = 0. \tag{35}$$

Given that

$$\begin{aligned} G(-z) &= g_0 - g_1 z + g_2 z^2 - g_3 z^3 = -z^3 H(z^{-1}) \quad \text{and} \\ H(-z^{-1}) &= g_3 + g_2 z^{-1} + g_1 z^{-2} + g_0 z^{-3} = z^{-3} G(z), \end{aligned} \tag{36}$$

It follows that this condition is automatically satisfied by choosing $G(z)$ and $H(z)$ to be complementary filters.

The conditions that ensure the mutual orthogonality of the elements of the bases of $\mathcal{V}_1$ and $\mathcal{W}_1$, which are associated with the first round of the dyadic decomposition, will guarantee the mutually orthogonality of all of the elements of the final basis that reside in different bands. The orthogonality of such elements may be described as *lateral* orthogonality.

Generating Wavelets and Scaling Functions

In the majority of cases, there are no analytic functions to represent the wavelets and the scaling functions in the time domain. Therefore, iterative procedures must be used for generating graphical representations of these functions. Such iterative procedures are based on the appropriate dilation equations.

Consider the equation (13). This indicates that, moving in the direction of higher resolution, there is

$$\phi(t) = 2^{1/2} \sum_{k=0}^{M-1} g_k \phi(2t - k). \tag{37}$$

The scaling functions on the RHS of (37) are supported on intervals of half the width of the interval supporting $\phi(t)$ and they are separated one from the next by distances 1/2 a unit. The amplitude coefficients $g_0, g_1, \ldots, g_{M-1}$ form a discrete sequence of which the elements can be attributed to the central points of the corresponding wavelets.

The scaling functions on the RHS of (37) are themselves amenable to expansions via the dilation equation. Thus

$$\phi(2t - k) = 2^{1/2} \sum_{j=0}^{M-1} g_j \phi(2[2t - k] - j). \tag{38}$$

When equation (38) is substituted into the RHS of equation (37), for all values of k, the result is an expression for $\phi(t)$ that contains M^2 contracted scaling functions. Each of these is supported on an interval that has 1/4th of the length of the support of $\phi(t)$.

In the resulting expression for $\phi(t)$, the scaling functions on the RHS of (38) are separated one from the next by distances of 1/4 of a unit, as are the corresponding amplitude coefficients. Some of these contracted functions share the same supports; and, together, they are supported on the same interval as $\phi(t)$. The amplitude coefficients come in batches of M elements at a time. Successive batches, indexed by k, are separated by distances of 1/2 a unit.

The manner in which the coefficients of the second expansion are generated may be illustrated by the case where $M = 4$. The coefficients are the products of the following multiplications:

$$\begin{bmatrix} g_0 & 0 & 0 & 0 \\ g_1 & 0 & 0 & 0 \\ g_2 & g_0 & 0 & 0 \\ g_3 & g_1 & 0 & 0 \\ 0 & g_2 & g_0 & 0 \\ 0 & g_3 & g_1 & 0 \\ 0 & 0 & g_2 & g_0 \\ 0 & 0 & g_3 & g_1 \\ 0 & 0 & 0 & g_2 \\ 0 & 0 & 0 & g_3 \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} g_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ g_1 & g_0 & 0 & 0 & 0 & 0 & 0 \\ g_2 & g_1 & g_0 & 0 & 0 & 0 & 0 \\ g_3 & g_2 & g_1 & g_0 & 0 & 0 & 0 \\ 0 & g_3 & g_2 & g_1 & g_0 & 0 & 0 \\ 0 & 0 & g_3 & g_2 & g_1 & g_0 & 0 \\ 0 & 0 & 0 & g_3 & g_2 & g_1 & g_0 \\ 0 & 0 & 0 & 0 & g_3 & g_2 & g_1 \\ 0 & 0 & 0 & 0 & 0 & g_3 & g_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & g_3 \end{bmatrix} \begin{bmatrix} g_0 \\ 0 \\ g_1 \\ 0 \\ g_2 \\ 0 \\ g_3 \end{bmatrix}. \tag{39}$$

The expression on the LHS corresponds to the manner of forming the coefficients that has already been described. That is to say, four batches of the four coefficients, separated by a fixed interval, are shifted successively by a double interval before being multiplied in turn by the coefficients g_0, g_1, g_2 and g_3.

The expression on the RHS embodies the lower-triangular Toeplitz matrix of a linear filter. The sequence that is subject to the filter is obtained by interpolating zeros between the elements of the pre-existing vector of derived amplitude coefficients. This is described as an upsampling operation.

Successive expansions of the sum of wavelets can proceed in the manner indicated by the expression on the RHS of (39), by upsampling the sequence of amplitude coefficients derived in the previous expansion and then by subjecting the result to a process of filtering.

As the number of expansions increases, an increasing number of amplitude coefficients are mapped into the interval that supports $\phi(t)$. In the process, the supports of the wavelets associated with the coefficients are successively diminished. Eventually, the wavelets will be adequately represented by spikes of unit area based on a point, which are Dirac delta functions. By that stage, the profile of $\phi(t)$ will be well represented by the closely spaced sequence of the derived amplitude coefficients.

The dilation equation for the wavelets functions at level 0, which is given

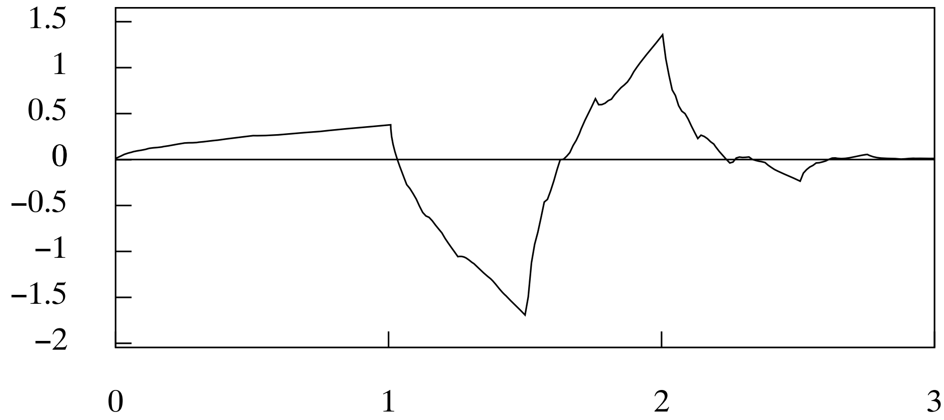

Figure 3. The Daubechies D4 wavelet function calculated via a recursive method.

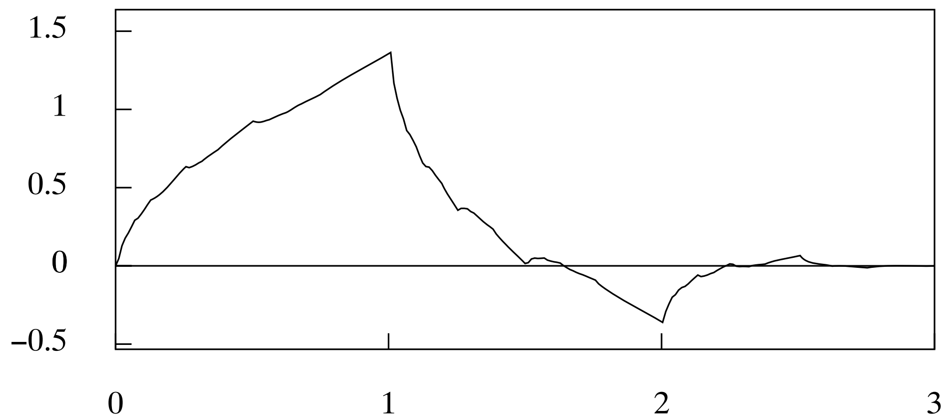

Figure 4. The Daubechies D4 scaling function calculated via a recursive method.

in (22), can be represented as follows:

$$\psi(t) = 2^{1/2} \sum_{k=0}^{M-1} h_k \phi(2t-k) \tag{40}$$

This can be expanded in the same way as the scaling function to generate a sequence of closely spaced coefficients that will represent the profile of the wavelet.

An example of a pair of wavelet and scaling functions that have dilation equations of four coefficients is provided by Daubechies' D4 functions. In this case, there are

$$\begin{aligned} g_0 &= (1+\sqrt{3})/(4\sqrt{2}), \qquad & g_1 &= (3+\sqrt{3})/(4\sqrt{2}), \\ g_2 &= (3-\sqrt{3})/(4\sqrt{2}), \qquad & g_3 &= (1-\sqrt{3})/(4\sqrt{2}), \end{aligned} \tag{41}$$

and there are $h_0 = g_3$, $h_1 = -g_2$, $h_2 = g_1$ and $h_3 = -g_0$, in accordance with (31). The profiles of the functions are represented in Figures 3 and 4.

The Decomposition of a Function in $\mathcal{V}_0$

If the approximation of equation (1) is replaced by an exact equality, then the equation will become

$$f(t) = \sum_k \big\langle \phi(t-k), f(t) \big\rangle \phi(t-j) = \sum_k y_k \phi(t-k). \tag{42}$$

This equation represents the projection of $f(t)$ on the basis vectors of $\mathcal{V}_0$. Since $\mathcal{V}_0 = \mathcal{W}_1 \oplus \mathcal{V}_1$, an alternative representation of $f(t)$ is obtained by projecting it on the conjunction of the basis vectors of $\mathcal{W}_1$ and $\mathcal{V}_1$, to generate the orthogonal components $v_1(t)$ and $w_1(t)$ of $f(t) = w_1(t) + v_1(t)$.

The projection on the level-1 scaling functions gives

$$v_1(t) = \sum_m \big\langle f(t), \phi_{1,m}(t) \big\rangle \phi_{1,m}(t). \tag{43}$$

But, in view of the dilation equation of (13), there is

$$\begin{aligned} \phi(2^{-1}t - m) &= 2^{1/2} \sum_k g_k \phi(2[2^{-1}t - m] - k) \\ &= 2^{1/2} \sum_k g_k \phi(t - [k + 2m]) = 2^{1/2} \sum_k g_{k-2m} \phi(t-k) \end{aligned} \tag{44}$$

which indicates that $\phi_{1,m}(t) = \sum_{k-2m} g_k \phi_{0,k}(t)$. Therefore, the coefficient associated with the basis function $\phi_{1,m}(t)$ is

$$\begin{aligned} \big\langle f(t), \phi_{1,m}(t) \big\rangle &= \sum_k g_{k-2m} \big\langle f(t), \phi_{0,k}(t) \big\rangle \\ &= \sum_k g_k y_{2m-k}. \end{aligned} \tag{45}$$

The projection of $f(t)$ on the level-1 wavelets gives

$$w_1(t) = \sum_m \big\langle f(t), \phi_{1,m}(t) \big\rangle \psi_{1,m}(t). \tag{46}$$

In this case, it can be shown, as in the case of the scaling functions, that the coefficient associated with $\psi_{1,m}(t)$ is

$$\big\langle f(t), \psi_{1,m}(t) \big\rangle = \sum_j h_k y_{2m-k}. \tag{47}$$

The equations

$$\gamma_{1m} = \sum_{k=0}^{T-1} g_k y_{2m-k}; \quad m = 0, 1, \ldots, [T/2] - 1, \tag{48}$$

$$\beta_{1m} = \sum_{k=0}^{T-1} h_k y_{2m-k}; \quad m = 0, 1, \ldots, [T/2] - 1, \tag{49}$$

of (45) and (47), which deliver the coefficients of the level-1 scaling functions and wavelets respectively, can be construed as the equations of a pair of complementary linear filters that are applied to a common data sequence $y_0, y_1, \ldots, y_{T-1}$ of T elements. Equation (48) describes a lowpass filter and equation (49) describes a highpass filter.

These filters move through the sample in step with the index $2m$, which is to say that they take steps of two points at a time. When this index is replaced by m, it becomes necessary to select alternate values of the filtered outputs via a process that is commonly described as down sampling.

Since the data sequence is finite, there will be problems in applying the filters at the ends of the sample where data are required that lie beyond the ends. To overcome the problem, the filter can be applied to the data via a process of circular convolution, which is equivalent to applying the filter to the periodic extension of the data.

To accommodate this adaptation within equation (48) and (49), it is sufficient to replace y_t by $y_{t \ mod \ T}$. When $t \in \{0, T-1\}$ there will be $y_{t \ mod \ T} = y_t$. Otherwise, when it appears to lie outside the sample, y_t will be replaced by a value from within the sample.

The second stage of the decomposition, as well as all subsequent stages, can be modelled on the first stage. Thus, the coefficients of the second stage are given by

$$\gamma_{2n} = \sum_{k=0}^{[T/2]-1} g_k \gamma_{1,2n-k}; \quad n = 0, 1, \ldots, [T/4]-1, \tag{50}$$

$$\beta_{2n} = \sum_{k=0}^{[T/2]-1} h_k \gamma_{1,2n-k}; \quad n = 0, 1, \ldots, [T/4]-1. \tag{51}$$

The complete process of decomposition is best represented using a matrix notation.

A Matrix Formulation of a Wavelets Analysis

Let $y = [y_0, \ldots, y_{T-1}]'$, where $T = 2^n$, represent the vector of observations, which are associated with the scaling functions of the initial basis, and let $\beta = [\beta_0, \ldots, \beta_{T-1}]'$ represent the vector of the coefficients associated with the wavelets of the final basis. Here, $\beta_{T-1} = \gamma_{n0}$ is the coefficient associated with the single scaling function in the ultimate subdivision of the frequency range. The mapping from y to β, denoted by $\beta = Q'y$, is effected by an orthonormal matrix Q such that $QQ' = Q'Q = I_T$.

Since $(Q')^{-1} = Q$, it follows that there is an inverse transformation from the wavelet coefficients to the data of the form $Q\beta = y$. This mapping from β to y effects a wavelet synthesis. If β contains a single nonzero element representing the amplitude coefficient of a solitary wavelet, then the mapping of β via Q will generate the vector, corresponding to a single column of Q, containing elements that approximate the ordinates of that wavelet, sampled at unit intervals.

The T elements of the vector β can be ordered in a manner that corresponds to a dyadic decomposition, such as is illustrated in Figure 1. Within β, there is a succession of subvectors, which contain the coefficients associated with the succession of the wavelet functions of the final basis. The subvectors of

$$\beta = [\beta'_{(1)}, \beta'_{(2)}, \ldots, \beta'_{(n)}, \gamma'_{(n)}]' \tag{52}$$

are

$$\begin{aligned} \beta_{(1)} &= [\beta_{10}, \beta_{11}, \ldots, \beta_{1,[T/2]-1}]', \\ \beta_{(2)} &= [\beta_{20}, \beta_{21}, \ldots, \beta_{2,[T/4]-1}]', \\ &\vdots \\ \beta_{(n-1)} &= [\beta_{n-1,0}, \beta_{n-1,1}]', \\ \beta_{(n)} &= [\beta_{n0},], \\ \gamma(n) &= [\gamma_{n0}]. \end{aligned} \tag{53}$$

A linear filter can be applied to a finite data sequence via a matrix transformation of the vector y of the data. Let the z-transform of a causal filter be represented by the polynomial $c(z) = c_0 + z_1 z + \cdots + c_{M-1}z^{M-1}$ and assume that the filter is applied to the data via a process of circular convolution.

Then, the matrix transformation that implements the filter can obtained by replacing the powers of z by powers of a circulant matrix $K_T = [e_1, e_2, \ldots, e_{T-1}, e_0]$. This matrix is formed from the identity matrix $I_T = [e_0, e_1, e_2, \ldots, e_{T-1}]$ by moving the leading vector e_0 to the end of the array. The resulting matrix is

$$c(K_T) = c_0 I_T + z_1 K_T + \cdots + c_{M-1}K_T^{M-1}, \tag{54}$$

and the filtered vector is given by $c(K_T)y$.

A process of down sampling can also be affected by a matrix transformation. The down sampling matrix is $\mathrm{V} = \Lambda' = [e_0, e_2, e_4, \ldots, e_{T-2}]'$, which is obtained by deleting alternate rows from the identity matrix I_T.

To see in detail how the wavelet amplitude coefficients can be generated in this manner, it is best to take a specific example. In the example, there are $T = 8 = 2^3$ data points and there are $M = 4$ coefficients in the dilation equations. Each stage of the process that converts the data into the wavelet coefficients involves the application of a linear filter followed by a process of down sampling.

The highpass filter that is to be applied to the data in the first round of the wavelets decomposition has the following matrix representation:

$$H_{(1)} = \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 \\ h_1 & h_0 & 0 & 0 & 0 & 0 & h_3 & h_2 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & h_3 \\ h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 \\ 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0 \\ 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0 \end{bmatrix}. \tag{55}$$

Premultiplying this by the down sampling matrix is a matter of deleting alternate rows:

$$\mathrm{V}H_{(1)} = \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & h_3 \\ 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 \end{bmatrix}. \tag{56}$$

When this matrix is combined with the matrix $\mathrm{V}G_{(1)}$, which is the down sampled version of the lowpass filter matrix, and when the data vector y is mapped through the combined matrix, the result is

$$\begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \hline \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & h_3 \\ 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 \\ \hline g_0 & 0 & 0 & 0 & 0 & g_3 & g_2 & g_1 \\ g_2 & g_1 & g_0 & 0 & 0 & 0 & 0 & g_3 \\ 0 & g_3 & g_2 & g_1 & g_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_3 & g_2 & g_1 & g_0 & 0 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix}. \tag{57}$$

The transformation can be represented, in summary notation, by

$$\begin{bmatrix} \beta_{(1)} \\ \gamma_{(1)} \end{bmatrix} = \begin{bmatrix} \mathrm{V}H_{(1)} \\ \mathrm{V}G_{(1)} \end{bmatrix} y. \tag{58}$$

In the second round of the wavelets decomposition, the coefficients associated with the level-1 wavelets are preserved and the coefficients associated with the level-1 scaling functions are subject to a further decomposition:

$$\begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \hline \beta_{20} \\ \beta_{21} \\ \hline \gamma_{20} \\ \gamma_{21} \end{bmatrix} = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & h_0 & h_3 & h_2 & h_1 \\ 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & h_3 \\ \hline 0 & 0 & 0 & 0 & g_0 & g_3 & g_2 & g_1 \\ 0 & 0 & 0 & 0 & g_2 & g_1 & g_0 & g_3 \end{array}\right] \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \hline \gamma_{10} \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \end{bmatrix}. \tag{59}$$

The summary notation for this is

$$\begin{bmatrix} \beta_{(1)} \\ \beta_{(2)} \\ \gamma_{(2)} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \mathrm{V}H_{(2)} \\ 0 & \mathrm{V}G_{(2)} \end{bmatrix} \begin{bmatrix} \beta_{(1)} \\ \gamma_{(1)} \end{bmatrix}. \tag{60}$$

The effect of the down sampling upon the circular filter can be see in equation (59). The two filters are defined on four points and, at this level, only four data points are available. There are no zeros remaining within the matrices $\mathrm{V}H'_{(2)}$ and $\mathrm{V}G'_{(2)}$.

In the next round of filtering, there are only two data points to be mapped through the filters. The consequence is that γ_{20} and γ_{21} must be used twice in the third and final transformation. This can represented equally by

$$\begin{bmatrix}\gamma_{30}\\ \beta_{30}\end{bmatrix} = \begin{bmatrix} h_0 & h_3 & h_2 & h_1\\ g_0 & g_3 & g_2 & g_1\end{bmatrix}\begin{bmatrix}\gamma_{20}\\ \gamma_{21}\\ \gamma_{20}\\ \gamma_{21}\end{bmatrix} \tag{61}$$

or by

$$\begin{bmatrix}\gamma_{30}\\ \beta_{30}\end{bmatrix} = \begin{bmatrix} h_0+h_2 & h_3+h_1\\ g_0+g_2 & g_3+g_1\end{bmatrix}\begin{bmatrix}\gamma_{20}\\ \gamma_{21}\end{bmatrix}. \tag{62}$$

On the LHS is a vector containing the amplitude coefficients, respectively, of a wavelet and a scaling function stretching the length of the data sequence.

A general expression can now be given for the set of amplitude coefficients associated with the projection of the function $f(t)$ onto the basis of the subspace $\mathcal{W}_j$. These coefficients are contained in the jth vector of the sequence of (53), which is given by

$$\beta_{(j)} = \mathrm{V}H_{(j)}\mathrm{V}G_{(j-1)}\cdots\mathrm{V}G_{(1)}y = Q'_{(j)}y. \tag{63}$$

In order to relieve the burden of notation, the subscripts have been omitted from the succession of down sampling matrices that would indicate their orders. Reading from right to left, the first down sampling matrix is $\mathrm{V}_{(1)}$ of order $T/2\times T$. The penultimate matrix is $\mathrm{V}_{(j-1)}$ of order $T/2^{j-1}\times T/2^{j-2}$ and the final matrix is $\mathrm{V}_{(j)}$ of order $T/2^j\times T/2^{j-1}$.

On the RHS of (63) is the matrix $Q'_{(j)}$, which represents a submatrix formed from a set of adjacent rows of the matrix Q', which is entailed in the mapping $\beta = Q'y$ from the sampled ordinates of $f(t)$ to the amplitude coefficients of the final basis. Given that $QQ' = I_T$, it follows that $Q\beta = y$ represents the synthesis of the vector y from the amplitude coefficients.

The vector $\beta_{(j)}$ of (63) is entailed in the synthesis of the component vector $w_j = [w_{0j}, w_{1j}, \ldots, w_{T-1,j}]'$ of the decomposition of $y = w_1 + \cdots + w_n + v_n$. The synthesis can be represented by

$$w_j = Q_{(j)}\beta_{(j)} = G'_{(1)}\Lambda\cdots G'_{(j-1)}\Lambda H'_{(j)}\Lambda\beta_{(j)}, \tag{64}$$

where $\Lambda = \mathrm{V}'$ represents the upsampling matrix, which interpolates zeros between the elements of any vector than it premultiplies.

The Two-Channel Quadrature Mirror Filter Bank

An understanding of the architecture of a dyadic wavelets analysis can be reaffirmed by considering the nature of two-channel quadrature mirror filter.

This should serve to highlight the symmetry and the essential simplicity of the design.

Consider, therefore, the following way of processing a signal. First, the signal is transmitted through two separate branches containing a lowpass filter G and a highpass filter H. Then, the filtered signals are down sampled by selecting alternate data points, indexed by even integers, and by discarding the points indexed by odd numbers. This operation is denoted by $(\downarrow 2)$. The two parts of the signal are transmitted separately, and an estimate of the original signal is produced by reassembling them.

Prior to the reassembly, zeros are interpolated between the elements of the component signals to replace the discarded elements. This operation is described as upsampling, and it is denoted by $(\uparrow 2)$. Then, the upsampled sequences are passed though separate smoothing filters E and D, designed to replace the interpolated zeros by estimates of the missing values. Finally, the two signals are added together.

Let the input signal be denoted by $x(t)$ and its z-transform by $x(z)$. Here, z is generally taken to be the complex exponential $e^{-\mathrm{i}\omega}$. In that case, $x(e^{-\mathrm{i}\omega})$ can be denoted more economically by $x(\omega)$. However, by using $x(z)$, a greater generality can be achieved at the same time as easing the burden of notation.

The path taken by the signal through the highpass branch of the network may be denoted by

$$x(z) \longrightarrow H(z) \longrightarrow (\downarrow 2) \longrightarrow \simeq \longrightarrow (\uparrow 2) \longrightarrow E(z) \longrightarrow w(z), \tag{65}$$

whereas the path taken through the lowpass branch may be denoted by

$$x(z) \longrightarrow G(z) \longrightarrow (\downarrow 2) \longrightarrow \simeq \longrightarrow (\uparrow 2) \longrightarrow D(z) \longrightarrow v(z). \tag{66}$$

The symbol $\simeq$ denotes the storage and transmission of the signals. The output signal, formed by merging the two branches, is $y(t) = v(t) + w(t)$.

The immediate objective is to find the z-transform of the reconstituted signal $y(t)$. Consider any signal $p(t)$ that has been subject to the processes of downsampling and upsampling to produce the sequence $q(t) = \{p(t \downarrow 2) \uparrow 2\}$.

Let $p(t) \longleftrightarrow p(\omega)$, which is to say that $p(\omega)$ is the Fourier transform of $p(t)$. Since $\omega \in [0, 2\pi]$, the domain of $p(\omega)$ is a circle. In the process of down sampling, the angular velocity ω is replaced by $\omega/2$ and the function evolves at half the previous rate. Therefore, it is wrapped twice around the circumference of the circle and the overlying ordinates are added. The effect is one of spectral aliasing.

The effect is summarised by writing $p(t \downarrow 2) \longleftrightarrow \frac{1}{2}\{p(\omega/2) + p(\pi + \omega/2)$. Since $e^{\pm \mathrm{i}\pi} = -1$ and $e^{-\mathrm{i}(\pi+\omega/2)} = -e^{-\mathrm{i}\omega/2}$, and, in terms of the z-transform, this becomes

$$p(t \downarrow 2) \longleftrightarrow \frac{1}{2}\{p(z^{1/2}) + p(-z^{1/2})\}. \tag{67}$$

Next, there is a process of upsampling, which doubles the value of the frequency argument. This gives $p\{(t \downarrow 2) \uparrow 2\} \longleftrightarrow \frac{1}{2}\{p(\omega) + p(\pi + \omega)\}$, which can also be written as

$$p\{(t \downarrow 2) \uparrow 2\} \longleftrightarrow \frac{1}{2}\{p(z) + p(-z)\}. \tag{68}$$

It follows that the signals that emerge from the two branches of the network are given by

$$\begin{aligned} w(z) &= \frac{1}{2}E(z)\{H(z)x(z) + H(-z)x(-z)\}, \\ v(z) &= \frac{1}{2}D(z)\{G(z)x(z) + G(-z)x(-z)\}. \end{aligned} \tag{69}$$

Combining the two signals gives

$$\begin{aligned} y(z) = \frac{1}{2}&\{D(z)G(-z) + E(z)H(-z)\}x(-z), \\ &+ \frac{1}{2}\{D(z)G(z) + E(z)H(z)\}x(z). \end{aligned} \tag{70}$$

The term in $x(-z)$ is due to aliasing and, by setting

$$D(z) = z^{-d}H(-z), \qquad E(z) = -z^{-d}G(-z), \tag{71}$$

it can be eliminated to give

$$\begin{aligned} y(z) &= \frac{1}{2}\{D(z)G(z) + E(z)H(z)\}x(z) \\ &= \frac{z^{-d}}{2}\{H(-z)G(z) - G(-z)H(z)\}x(z). \end{aligned} \tag{72}$$

The terms z^{-d}, which have been included in the definitions of $D(z)$ and $E(z)$, can serve to compensate for the time lags that have been induced by the initial the process of filtering via $G(z)$ and $H(z)$, and they may be omitted if it is required to depict a process that operates in real time.

It should be observed that the conditions of (71) guarantee the alias cancellation for any choice of the filters $G(z)$ and $H(z)$. More restricted choices are indicted if it is required that $y(t) = x(t)$. In that case, it can be said that the filters fulfil the condition of prefect reconstruction.

The restriction that the coefficients of the filters $G(z)$ and $H(z)$ should constitute mutually orthogonal vectors indicates a unique choice of the anti-aliasing filters. To illustrate this, we may consider the case where the filter span is $M = 4$. Then, the following relationships prevail, which can be generalised easily:

$$\begin{aligned} &\text{(i)} && G(z) = g_0 + g_1z + g_2z^2 + g_3z^3 = z^3H(-z^{-1}) = D(z^{-1}) \\ &\text{(ii)} && H(z) = g_3 - g_2z + g_1z^2 - g_0z^3 = -z^3G(-z^{-1}) = E(z^{-1}), \\ &\text{(iii)} && D(z) = g_0 + g_1z^{-1} + g_2z^{-2} + g_3z^{-3} = z^{-3}H(-z) = G(z^{-1}), \\ &\text{(iv)} && E(z) = g_3 - g_2z^{-1} + g_1z^{-2} - g_0z^{-3} = -z^{-3}G(-z) = H(z^{-1}). \end{aligned} \tag{73}$$

Using such relationships, equation (72) can be rendered as

$$\begin{aligned} y(z) &= \frac{1}{2}\{G(z^{-1})G(z) + H(z^{-1})H(z)\}x(z) \\ &= \frac{1}{2}\{D(z)G(z) + D(-z)G(-z)\}x(z) \\ &= \frac{1}{2}\{P(z) + P(-z)\}x(z). \end{aligned} \tag{74}$$

The condition that $y(t) = x(t)$ indicates that

$$P(z) + P(-z) = 2, \tag{75}$$

which is equation (29) again.

The coefficients of $P(z)$ and $P(-z)$ that are associated with odd powers of z and $-z$ will be cancelled within (75). To maintain the equality, the coefficients associated with positive powers must be zeros. These zero-valued coefficients of $P(z)$ and $P(-z)$ correspond to the orthogonality conditions of (17) and (22), respectively, and so the requirement is satisfied.

Finally, the normalisations (15) and (21), which relate to the sums of squares of the coefficients, provide the coefficients of unity that are associated with z^0 within $P(z)$ and $P(-z)$. Thus, the value on the RHS of (75) is confirmed.

The equation

$$\frac{1}{2}\big\{G(z^{-1})G(z) + H(z^{-1})H(z)\big\} = 1, \tag{76}$$

which is an alternative form of (75), summarises the structure of the filter bank with reference to the complementary highpass and low pass filters, represented by $H(z)$ and $G(z)$, respectively. It shows their overall effect, which is to deliver an output that is a perfect reconstruction of the input sequence.

It is straightforward to derive matrix representations of the filters from their z-transforms, by replacing the powers of the argument z by powers of the circulant matrix $K_T = [e_1, e_2, \ldots, e_{T-1}, e_0]$. Thus, for example, the filter matrix of (55) is obtained by setting $z = K_8$ within $H(z)$ of (73, ii).

The negative powers of z are to be replaced by powers of the transposed matrix $K_T' = [e_{T-1}, e_0, e_1, \ldots, e_{T-2}]$. Thus, the matrix associated with $H(z^{-1})$ corresponds to the transpose of the matrix associated with $H(z)$.