Orthogonal Wavelets via Filter Banks : Theory and Applications

Joab R. Winkler

The University of Sheffield
Department of Computer Science
Regent Court
211 Portobello Street
Sheffield S1 4DP
United Kingdom

`j.winkler@dcs.shef.ac.uk`

Abstract

Wavelets are used in many applications, including image processing, signal analysis and seismology. The critical problem is the representation of a signal using a small number of computable functions, such that it is represented in a concise and computationally efficient form. It is shown that wavelets are closely related to filter banks (subband filtering) and that there is a direct analogy between multiresolution analysis in continuous time and a filter bank in discrete time. This provides a clear physical interpretation of the approximation and detail spaces of multiresolution analysis in terms of the frequency bands of a signal. Only orthogonal wavelets, which are derived from orthogonal filter banks, are discussed. Several examples and applications are considered.

1 Introduction

The fundamental aim of signal processing is the construction of a set of basis (or more generally expansion) functions that allow a concise, efficient and informative representation of a signal. Clearly, the choice of basis functions depends on the signal; the Fourier basis (complex exponentials) is usually satisfactory for smooth signals but it is poor if the signal has discontinuities or has portions of both low and high frequencies. More specifically, since the Fourier basis functions have infinite support in the time domain, they have poor time resolution properties and therefore do not reveal the instants t_i at which changes in a signal $f(t)$ occur. The short time Fourier transform (STFT) improves this situation by placing a window round the exponential basis functions, thereby allowing individual portions of the signal to be analysed and revealing more local information about the signal. The main disadvantage of the STFT is that the width of the window is constant and must therefore capture information in both the low and high frequency portions of the signal. A wavelet basis is superior because it can adapt to local changes in the signal; narrow windows can be used in the high frequency regions of the signal and wider windows can be used in the low frequency regions. The wavelet transform of a function $f(t)$ is the representation of $f(t)$ in a wavelet basis, and the elements of the basis are scaled and translated forms of a mother wavelet $\psi(t)$.

It is shown that the calculation of the coefficients of the dilated and translated versions of $\psi(t)$ in the representation of $f(t)$ can be be performed recursively and implemented in a

perfect reconstruction filter bank, that is, a filter bank in which the output is exactly equal to the input. Two–channel filter banks are considered in section 2 and the conditions for perfect reconstruction are derived. An orthogonal filter bank is a particular class of filter bank, and the conditions for perfect reconstruction in this type of filter bank are obtained in section 3.

The multiresolution analysis of $f(t)$ into nested subspaces is considered in section 4 and it is shown that the equations that govern this wavelet decomposition of $f(t)$ are identical to the equations that define perfect reconstruction in an orthogonal filter bank. This analogy between filter banks, whose input is a discrete signal, and the multiresolution analysis of the continuous function $f(t)$ allows the coefficients of the representation of $f(t)$ in a wavelet basis to be implemented in a filter bank. Several examples of the applications of wavelets are considered in section 5.

Attention is restricted to orthogonal filter banks which lead to orthogonal wavelets, and the Daubechies wavelets are the best examples of this class of wavelet. More general biorthogonal filter banks have greater design freedom in the choice of filters that lead to perfect reconstruction, and two multiresolutions are required, one for the analysis bank and one for the synthesis bank.

2   Two–channel filter banks

A two–channel filter bank is shown in figure 1. It consists of two lowpass filters H_0 and F_0, two highpass filters H_1 and F_1, two downsamplers $\downarrow 2$, and two upsamplers $\uparrow 2$. The downsamplers remove all the odd–numbered samples, and the upsamplers insert a zero between every pair of samples. The filters H_0 and H_1 are called analysis filters, and the filters F_0 and F_1 are called synthesis filters.
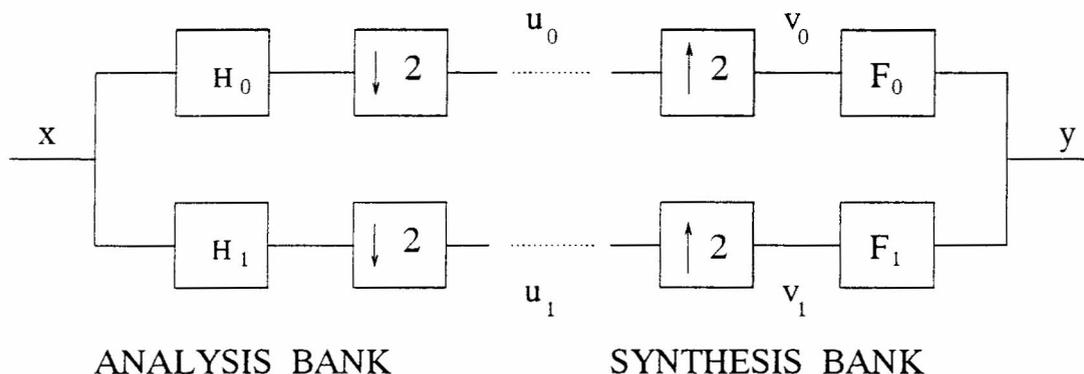

Figure 1: A two–channel filter bank.

The input signal $\mathbf{x} = \{x(n)\}$ is separated into two frequency bands, a low frequency band corresponding to the upper channel, and a high frequency band corresponding to the lower channel. The signal in each band between the analysis and synthesis stages may be coded for transmission or storage, and compression may occur in which case information is lost. Perfect reconstruction requires that the analysis bank be connected directly to the synthesis bank, and thus no compression occurs. The downsamplers improve the efficiency of the filter bank because each channel carries only half the data (information); the exclusion of the downsamplers implies that all the input goes into each channel, thus doubling the work that is performed by the filter bank, but with no increase in information.

The analysis of the filter bank in figure 1 requires that the upsamplers and downsamplers be formally defined. Thus if $\mathbf{r} = \{r(n)\}$, $\mathbf{s} = \{s(n)\}$, and $\mathbf{t} = \{t(n)\}$ then

1. The nth component of $\mathbf{r} = (\downarrow 2)\,\mathbf{t}$ is the $(2n)$th component of $\mathbf{t}$,

$$r(n) = t(2n). \tag{1}$$

It is shown in [11], chapter 3, that the z–domain form of (1) is

$$R(z) = \frac{1}{2}\left[T\left(z^{\frac{1}{2}}\right) + T\left(-z^{\frac{1}{2}}\right)\right], \tag{2}$$

where $R(z)$ and $T(z)$ are the z–transforms of $\mathbf{r}$ and $\mathbf{t}$ respectively.

2. The nth component of $\mathbf{s} = (\uparrow 2)\,\mathbf{t}$ is

$$
\begin{aligned}
s(2n) &= t(n) \\
s(2n+1) &= 0,
\end{aligned} \tag{3}
$$

and the z–domain form of (3) is [11], chapter 3,

$$S(z) = T\left(z^2\right). \tag{4}$$

Although the filters H_0, F_0, H_1 and F_1 are linear and time–invariant, it follows immediately from (1) and (3) that downsampling and upsampling are linear but not time–invariant. The downsamplers and upsamplers create aliases and images respectively, and these are removed by the filters. In particular, H_0 and H_1 must be band–limited to lower and upper halfbands respectively,

$$
\begin{aligned}
H_0(\omega) &= 0 \quad \text{if} \quad \frac{\pi}{2} \le |\omega| < \pi \\
H_1(\omega) &= 0 \quad \text{if} \quad 0 \le |\omega| < \frac{\pi}{2},
\end{aligned}
$$

to remove aliases. It is shown in the next section that the condition for perfect reconstruction implies that the synthesis filters are defined by the analysis filters.

2.1 Perfect reconstruction

The conditions for perfect reconstruction in a two–channel filter bank are derived, and the constraints that they impose on the coefficients of the filters are considered.

The input to the downsampler in the upper channel in the filter bank in figure 1 is $H_0(z)X(z)$ where $X(z)$ is the z–transform of the input sequence $\{x(n)\}$, and thus it follows from (2) that the z–transform of the output of the downsampler in the upper channel is

$$
\begin{aligned}
U_0(z) &= (\downarrow 2)\,H_0(z)X(z) \\
&= \frac{1}{2}\left[H_0\left(z^{\frac{1}{2}}\right)X\left(z^{\frac{1}{2}}\right) + H_0\left(-z^{\frac{1}{2}}\right)X\left(-z^{\frac{1}{2}}\right)\right],
\end{aligned}
$$

and hence the output of the upsampler in the upper channel is, from (4),

$$
\begin{aligned}
V_0(z) &= U_0(z^2) \\
&= \frac{1}{2}\left[H_0\left(z\right)X\left(z\right) + H_0\left(-z\right)X\left(-z\right)\right].
\end{aligned}
$$

Similarly, the output of the upsampler in the lower channel is

$$V_1(z) = \frac{1}{2}\left[H_1\left(z\right)X\left(z\right) + H_1\left(-z\right)X\left(-z\right)\right],$$

and thus the z–transform of the output $\mathbf{y} = \{y(n)\}$ of the filter bank is

$$\begin{aligned}
Y(z) &= F_0(z)V_0(z) + F_1(z)V_1(z) \\
&= \frac{1}{2}\left[F_0(z)H_0(z) + F_1(z)H_1(z)\right]X(z) + \frac{1}{2}\left[F_0(z)H_0(-z) + F_1(z)H_1(-z)\right]X(-z).
\end{aligned}$$

$$(5)$$

Equation (5) shows that $Y(z)$ is a function of $X(z)$ and $X(-z)$, and since perfect reconstruction requires that $Y(z) = X(z)$, it follows that

$$F_0(z)H_0(z) + F_1(z)H_1(z) = 2, \tag{6}$$
$$F_0(z)H_0(-z) + F_1(z)H_1(-z) = 0. \tag{7}$$

Equation (6) is the condition for the absence of distortion and (7) is the condition for alias cancellation, and the two equations define the conditions that must be satisfied by the filters for perfect reconstruction. They can be stated in matrix form by making the substitution $z \to -z$, in which case they become

$$F_0(-z)H_0(-z) + F_1(-z)H_1(-z) = 2, \tag{8}$$
$$F_0(-z)H_0(z) + F_1(-z)H_1(z) = 0, \tag{9}$$

and (6)–(9) can be combined to yield

$$\begin{bmatrix} F_0(z) & F_1(z) \\ F_0(-z) & F_1(-z) \end{bmatrix}\begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \tag{10}$$

or

$$\mathbf{F}_m(z)\mathbf{H}_m(z) = 2\mathbf{I}, \tag{11}$$

where $\mathbf{F}_m(z)$ is the synthesis modulation matrix and $\mathbf{H}_m(z)$ is the analysis modulation matrix. The frequency domain forms of (6) and (7) are obtained by substituting $z = e^{i\omega}$, and these equations require that $F_0(\omega = 0)H_0(\omega = 0) = 2$ and $F_1(\omega = \pi)H_1(\omega = \pi) = 2$. It follows that the lowpass and highpass filters must be normalised so that

$$H_0(\omega = 0) = F_0(\omega = 0) = \sqrt{2} \quad \text{and} \quad H_1(\omega = \pi) = F_1(\omega = \pi) = \sqrt{2}. \tag{12}$$

It is noted that their more usual values are

$$H_0(\omega = 0) = F_0(\omega = 0) = 1 \quad \text{and} \quad H_1(\omega = \pi) = F_1(\omega = \pi) = 1.$$

It follows from (10) that

$$\det \mathbf{H}_m(z) = H_0(z)H_1(-z) - H_1(z)H_0(-z) = -\det \mathbf{H}_m(-z),$$

and thus $\det \mathbf{H}_m(z)$ is an odd function of z. It also follows from (10) that

$$\begin{bmatrix} F_0(z) & F_1(z) \end{bmatrix} = \frac{1}{\det \mathbf{H}_m(z)}\begin{bmatrix} 2 & 0 \end{bmatrix}\begin{bmatrix} H_1(-z) & -H_0(-z) \\ -H_1(z) & H_0(z) \end{bmatrix},$$

and thus

$$F_0(z) = \frac{2H_1(-z)}{\det \mathbf{H}_m(z)} \quad \text{and} \quad F_1(z) = \frac{-2H_0(-z)}{\det \mathbf{H}_m(z)}.$$

If $H_0(z)$ and $H_1(z)$ are finite impulse response (FIR) filters, then $F_0(z)$ and $F_1(z)$ are also FIR filters if $\det \mathbf{H}_m(z)$ is of the form cz^{-l} where l is an odd integer and c is an arbitrary constant. The choice $c = 2$ leads to

$$F_0(z) = z^l H_1(-z) \quad \text{and} \quad F_1(z) = -z^l H_0(-z), \tag{13}$$

4

and these definitions of the synthesis filters in terms of the analysis filters guarantee that the alias cancellation condition (7) is satisfied. It is noted that the frequency domain form of (13) is

$$F_0(\omega) = e^{i\omega l}H_1(\omega + \pi) \qquad \text{and} \qquad F_1(\omega) = -e^{i\omega l}H_0(\omega + \pi),$$

and thus $F_0(\omega)$ and $F_1(\omega)$ are lowpass and highpass filters respectively if $H_0(\omega)$ and $H_1(\omega)$ are lowpass and highpass filters respectively, as required.

The condition for perfect reconstruction can be cast in a more convenient form by substituting (13) into (8), which yields

$$P_0(z) + P_0(-z) = 2, \qquad (14)$$

where

$$P_0(z) = F_0(z)H_0(z), \qquad (15)$$

is the lowpass product filter. It follows from (14) that all the even coefficients, apart from the constant coefficient, of $P_0(z)$ are zero,

$$P_0(z) = \sum_n p_0(n)z^{-n} = 1 + \sum_n p_0(2n+1)z^{-(2n+1)}.$$

The design of a perfect reconstruction two–channel filter bank has been reduced to the following three steps :

1. Design a lowpass filter $P_0(z)$ that satisfies (14). The design variables are the odd coefficients of $P_0(z)$.

2. Factorise $P_0(z)$ into $F_0(z)H_0(z)$.

3. Calculate the highpass filters from the alias cancellation condition (13).

There are several methods that can be used to achieve the first step, and the factorisation in the second step is not unique because if $P_0(z) = \prod(z - \alpha_i)$, the assignment of the factors $z - \alpha_i$ to $F_0(z)$ and $H_0(z)$ must still be determined, and thus there is considerable scope for the design of filters that satisfy additional properties. The next section considers the alternating flip, which simplifies the design procedure because only the lowpass filter $H_0(z)$ is designed and the remaining filters are then defined from it. Furthermore, the alternating flip leads to orthogonal filter banks.

2.1.1 The alternating flip

The alternating flip defines the coefficients $h_1(n)$ of $H_1(z)$ in terms of the coefficients $h_0(n)$ of $H_0(z)$,

$$h_1(n) = (-1)^n h_0(N - n), \qquad n = 0 \ldots N, \qquad (16)$$

where $H_0(z)$ and $H_1(z)$ have $N+1$ coefficients. Furthermore, N must be odd since it guarantees double shift orthogonality between $\mathbf{h}_0 = \{h_0(n)\}_{n=0}^N$ and $\mathbf{h}_1 = \{h_1(n)\}_{n=0}^N$,

$$\sum_{n=0}^N h_0(n - 2k)h_1(n) = 0, \qquad (17)$$

for all integers k. It will be shown that this double shift orthogonality provides the link between perfect reconstruction in an orthogonal filter bank and multiresolution analysis.

5

Example 2.1 If $N = 5$, then from (16),

$$\begin{array}{rcllllll}
\mathbf{h}_0 & = & h_0(0) & h_0(1) & h_0(2) & h_0(3) & h_0(4) & h_0(5) \\
\mathbf{h}_1 & = & h_0(5) & -h_0(4) & h_0(3) & -h_0(2) & h_0(1) & -h_0(0),
\end{array}$$

and it is clear that

$$\sum_{n=0}^{5} h_0(n)h_1(n) = \sum_{n=2}^{5} h_0(n-2)h_1(n) = \sum_{n=4}^{5} h_0(n-4)h_1(n) = 0,$$

and thus (17) is satisfied for all the double shifts of a filter with six coefficients. However it is readily verified that (17) not satisfied if N is even. $\square$

It follows from (16) that

$$H_1(z) = -z^{-N} H_0(-z^{-1}), \tag{18}$$

in the z–domain, and

$$H_1(\omega) = -e^{-i\omega N}\overline{H_0(\omega + \pi)}, \tag{19}$$

in the frequency domain, and thus $H_1(\omega)$ is highpass if $H_0(\omega)$ is lowpass.

The simplifications in the design of a perfect reconstruction filter bank that follow from the application of the alternating flip are obtained by substituting (18) into (13),

$$F_0(z) = z^{l-N} H_0\left(z^{-1}\right),$$

where both l and N are odd integers. The choice $l = N$ yields $F_0(z) = H_0\left(z^{-1}\right)$ and thus the lowpass product filter (15) is given by

$$P_0(z) = H_0(z)H_0\left(z^{-1}\right), \tag{20}$$

which must still satisfy (14). The frequency domain form of (20) is

$$P_0(\omega) = H_0(\omega)\overline{H_0(\omega)} = |H_0(\omega)|^2,$$

and thus the frequency domain form of (14) is

$$|H_0(\omega)|^2 + |H_0(\omega + \pi)|^2 = 2. \tag{21}$$

Furthermore, it is easily verified from (19) that

$$|H_1(\omega)|^2 + |H_1(\omega + \pi)|^2 = 2, \tag{22}$$

$$H_1(\omega)\,\overline{H_0(\omega)} + H_1(\omega + \pi)\,\overline{H_0(\omega + \pi)} = 0, \tag{23}$$

$$H_0(\omega)\,\overline{H_1(\omega)} + H_0(\omega + \pi)\,\overline{H_1(\omega + \pi)} = 0. \tag{24}$$

It is shown in section 3 that (21) is the fundamental design equation of an orthogonal filter bank. Equations (22)–(24) follow directly from the alternating flip and therefore do not yield extra information. However they are included because they will be referenced in section 3 where it will be shown that the alternating flip leads to an orthogonal filter bank. Comparison of (15) and (20) shows that the application of the alternating flip implies that $P_0(z)$ is a function of $H_0(z)$ rather than $H_0(z)$ and $F_0(z)$, and thus the design of an orthogonal filter bank reduces to the design of the lowpass analysis filter $H_0(z)$; the highpass analysis filter $H_1(z)$ is defined by the alternating flip and the synthesis filters $F_0(z)$ and $F_1(z)$ are defined by (13).

3 Orthogonal filter banks

Equation (11) states that perfect reconstruction in an arbitrary filter bank is achieved if the analysis bank is inverted by the synthesis bank. Restrictions on the class of matrices limit the class of filters, and in particular, an orthogonal filter bank is derived by restricting $\mathbf{H}_m(z)$ to be unitary, up to a scalar multiplier, on the unit circle in the z–plane. This leads to paraunitary matrices, which are extensions of unitary matrices [16].

Definition 3.1 The matrix $\mathbf{H}(z)$ is paraunitary on the unit circle $|z| = 1$ if

$$\mathbf{H}^T \left(e^{-i\omega}\right) \mathbf{H} \left(e^{i\omega}\right) = d\mathbf{I} \qquad \text{for all } \omega, \tag{25}$$

where d is an arbitrary positive constant. This extends to all $z \neq 0$ by

$$\mathbf{H}^T \left(z^{-1}\right) \mathbf{H} \left(z\right) = \tilde{\mathbf{H}} \left(z\right) \mathbf{H} \left(z\right) = d\mathbf{I}. \tag{26}$$

$\square$

The positive constant d is included so that the analysis modulation matrix, for which $d = 2$, can be constrained to be paraunitary. It follows from (25) that $\mathbf{H}(z)$ is paraunitary on the unit circle in the z–plane, that is, in the frequency domain, and that its extension (26) to the rest of the z–plane (apart from the origin) follows from $\overline{z} = z^{-1} = e^{-i\omega}$ if $|z| = 1$.

It follows from (26) that the analysis modulation matrix $\mathbf{H}_m(z)$ is paraunitary if

$$\left[\begin{array}{cc} H_0\left(z^{-1}\right) & H_1\left(z^{-1}\right) \\ H_0\left(-z^{-1}\right) & H_1\left(-z^{-1}\right) \end{array} \right] \left[\begin{array}{cc} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{array} \right] = \left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array} \right],$$

and since the right hand side is a multiple of the identity matrix, the order of multiplication on the left hand side can be reversed,

$$\left[\begin{array}{cc} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{array} \right] \left[\begin{array}{cc} H_0\left(z^{-1}\right) & H_1\left(z^{-1}\right) \\ H_0\left(-z^{-1}\right) & H_1\left(-z^{-1}\right) \end{array} \right] = \left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array} \right]. \tag{27}$$

The frequency domain form of (27) is

$$\left[\begin{array}{cc} H_0(\omega) & H_0(\omega + \pi) \\ H_1(\omega) & H_1(\omega + \pi) \end{array} \right] \left[\begin{array}{cc} \overline{H_0\left(\omega\right)} & \overline{H_1\left(\omega\right)} \\ \overline{H_0\left(\omega + \pi\right)} & \overline{H_1\left(\omega + \pi\right)} \end{array} \right] = \left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array} \right], \tag{28}$$

and thus (21)–(24) are reproduced. This establishes that if the synthesis filters are defined by (13), then the alternating flip definition of the analysis highpass filter (16) necessarily leads to an orthogonal filter bank.

It follows from (27) that

$$H_0\left(z\right) H_0\left(z^{-1}\right) + H_0\left(-z\right) H_0\left(-z^{-1}\right) = 2, \tag{29}$$

$$H_1\left(z\right) H_1\left(z^{-1}\right) + H_1\left(-z\right) H_1\left(-z^{-1}\right) = 2, \tag{30}$$

$$H_1\left(z\right) H_0\left(z^{-1}\right) + H_1\left(-z\right) H_0\left(-z^{-1}\right) = 0, \tag{31}$$

$$H_0\left(z\right) H_1\left(z^{-1}\right) + H_0\left(-z\right) H_1\left(-z^{-1}\right) = 0, \tag{32}$$

and it is seen that (32) follows from (31) by the substitution $z \rightarrow z^{-1}$. It is instructive to consider the inverse z–transform of each of these equations because it will reveal the conditions

that the coefficients of $H_0(z)$ and $H_1(z)$ must satisfy in an orthogonal filter bank. It is adequate to consider the general function

$$P(z)Q(z^{-1}) + P(-z)Q(-z^{-1}),$$

because the left hand sides of (29) and (30) are obtained by setting $P(z) = Q(z)$. In particular, if $P(z) \longleftrightarrow p(n)$ and $Q(z) \longleftrightarrow q(n)$, then

$$P(z)Q(z^{-1}) + P(-z)Q(-z^{-1}) \longleftrightarrow 2\sum_n p(n)q(n-2k),$$

and thus the inverse z-transforms of (29)–(32) are

$$\sum_n h_0(n)h_0(n-2k) = \delta(k), \tag{33}$$

$$\sum_n h_1(n)h_1(n-2k) = \delta(k), \tag{34}$$

$$\sum_n h_1(n)h_0(n-2k) = 0. \tag{35}$$

These equations define the double shift orthogonality conditions, and their satisfaction guarantees that the filter bank is orthogonal. Equations (33) and (34) are the normalisation condition on the 2–norm of the coefficients, and (35) is exactly the same as (17), as expected. It is important to emphasize that only (33) need be solved; the alternating flip necessarily implies that if this equation is satisfied, then (34) and (35) are also satisfied. Although the coefficient forms (33)–(35) are equivalent to the z-domain forms (29)–(32), the former are preferred because it will be shown that the equivalence between multiresolution and perfect reconstruction in an orthogonal filter bank is clearer.

Comparison of (10) and (27) shows that the synthesis filters are given by

$$F_0(z) = H_0(z^{-1}) \quad \text{and} \quad F_1(z) = H_1(z^{-1}), \tag{36}$$

or

$$f_0(n) = h_0(-n) \quad \text{and} \quad f_1(n) = h_1(-n),$$

and thus the synthesis filters are anticausal if the analysis filters are causal. These results are summarised in the following theorem.

Theorem 3.1 The coefficients of the filters in an orthogonal filter bank are given by

$$\begin{aligned}
h_1(n) &= (-1)^n h_0(N-n), \qquad n = 0\ldots N, \\
f_0(n) &= h_0(-n), \\
f_1(n) &= h_1(-n).
\end{aligned}$$

The coefficients $h_0(n)$ satisfy the double shift orthogonality condition,

$$\sum_n h_0(n)h_0(n-2k) = \delta(k).$$

□

Example 3.1 The simplest example of an orthogonal filter bank is obtained with

$$H_0(z) = \frac{1}{\sqrt{2}}(1 + z^{-1}),$$

8

whose output is a scaled form of the moving average of the input, where the scale factor $\sqrt{2}$ is included to satisfy (12). The filter $H_1(z)$ is a scaled from of the moving difference,

$$H_1(z) = \frac{1}{\sqrt{2}}(1 - z^{-1}),$$

and the synthesis filters are, from (36),

$$F_0(z) = \frac{1}{\sqrt{2}}(1 + z) \qquad \text{and} \qquad F_1(z) = \frac{1}{\sqrt{2}}(1 - z).$$

These filters, which satisfy (33)–(35) trivially, give rise to the Haar wavelet. It is easily verified that $\det \mathbf{H}_m(z) = 2z^{-1}$, as required since $N = 1$. $\square$

Example 3.2 The coefficients

$$\mathbf{h}_0 = \frac{1}{4\sqrt{2}} \left[\ 1 + \sqrt{3}\quad 3 + \sqrt{3}\quad 3 - \sqrt{3}\quad 1 - \sqrt{3}\ \right],$$

satisfy the double shift orthogonality condition (33). They define the Daubechies wavelet D4, which is a member of the Daubechies family of wavelets. This family of filters and wavelets has maximum flatness at $\omega = \pi$ and therefore good approximation properties. It is easily verified that $\det \mathbf{H}_m(z) = 2z^{-3}$, as required since $N = 3$. $\square$

Theorem 3.1 shows that all the filters must be the same length $(N + 1)$, but this may be a disadvantage in some applications. This constraint is relaxed in biorthogonal wavelets, which leads to greater design freedom. Furthermore, some applications require that the filter have linear phase, which implies that the filter coefficients are symmetric or anti–symmetric, but this property can only be achieved by a very restricted class of orthogonal filter, as shown in the following example.

Example 3.3 Consider an anti–symmetric filter for which $N = 5$,

$$\mathbf{h}_0 \quad = \quad h_0(0)\quad h_0(1)\quad h_0(2)\quad -h_0(2)\quad -h_0(1)\quad -h_0(0).$$

Since this vector must be orthogonal to its shift by two and shift by four, it follows that

$$h_0(0)h_0(2) - h_0(1)h_0(2) = 0 \qquad \text{and} \qquad h_0(0)h_0(1) = 0.$$

Since $N = 5$, it follows that $h_0(0) \neq 0$, and thus $h_0(1) = 0$, which implies that $h_0(2) = 0$, and hence only the first and last coefficients of the filter are non–zero. $\square$

The next section considers multiresolution analysis, which is defined for a continuous rather than discrete function. It will be shown that a multiresolution analysis yields the double shift orthogonality conditions (33)–(35), and this enables the link between subband filtering and multiresolution analysis to be considered in greater detail.

4 Multiresolution analysis

Every measurement or observation is made on a particular scale, and this defines the amount of information that is contained in the measurement. For example, when a distant object is viewed, only the low level (coarse) detail can be observed and hence a small scale is used. However as the distance between the observer and object decreases, higher level (finer) details are observed and the scale increases. Another example of scale occurs in maps, for which a small scale is used to cover a large area and thus local detail is absent, but as the scale increases,

9

a smaller physical area is covered and more detail is included. These two simple examples show that measurements can be made on different scales. Multiresolution is a mathematical description of an object (measurement, image, etc.) in terms of scale, or level of detail.

The scale in which an object is viewed is intrinsically related to sampling a function $f(t)$. Specifically, if $f(t)$ is sampled at 2^j samples per unit time, the scale of the discrete signal $f(n)$ is defined because details that occur at a frequency that is higher than that specified by the sampling theorem are not represented in $f(n)$. However if the sampling frequency increases to 2^{j+1}, then finer detail can be captured in the discrete signal. The changes in detail that occur as the sampling frequency changes must be defined quantitatively, and this leads to a multiresolution decomposition of $f(t)$.

4.1 Scale spaces

The relation between scale and detail is formalised by defining a subspace $V_j \subset L^2(\mathbb{R})$ that contains all functions $f_j(t)$ that are represented at a scale 2^j, that is, functions that can be reconstructed by sampling at 2^j samples per unit time. Similarly, the subspace V_{j+1} contains all functions $f_{j+1}(t)$ that can be represented at a scale 2^{j+1}, and since the space V_{j+1} contains functions $\Delta f_j(t)$ that are not in V_j, it follows that there exists a detail space W_j such that

$$V_{j+1} = V_j \oplus W_j \qquad \text{and} \qquad V_j \cap W_j = \{0\}. \tag{37}$$

where $\Delta f_j(t) \in W_j$. The zero intersection condition guarantees that only the zero vector lies in both V_j and W_j. By definition, $V_j \subset V_{j+1}$, and this defines a family of nested subspaces,

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots. \tag{38}$$

Multiresolution decomposes a function $f(t)$ into a sum of functions $f_j(t) \in V_j$ where each space V_j is associated with a scale or level of detail. In particular, the recursive application of (37) leads to

$$V_0 \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots \oplus W_j = V_{j+1}, \tag{39}$$

or in terms of functions,

$$f_0(t) + \Delta f_0(t) + \Delta f_1(t) + \Delta f_2(t) + \cdots + \Delta f_j(t) = f_{j+1}(t). \tag{40}$$

Equations (39) and (40) show that multiresolution analyses a signal $f(t)$ in terms of a low level approximation that lies in a space V_0 and successively finer details that lie in the spaces $W_k, k \geq 0$. In the frequency domain, (40) states that the function $f_0(t) \in V_0$ is a low frequency approximation of $f(t)$, and higher frequencies are added as more detail spaces W_j are included, and thus the approximation of $f(t)$ by $f_j(t)$ increases as j increases. This frequency domain interpretation of scale spaces is shown in figure 2 for $j = 3$,

$$\begin{aligned} f_3(t) \in V_3 &= V_2 \oplus W_2 \\ &= V_1 \oplus W_1 \oplus W_2 \\ &= V_0 \oplus W_0 \oplus W_1 \oplus W_2. \end{aligned}$$

It is seen that the function $f_0(t) \in V_0$ has a lowpass frequency spectrum but that each of the functions $\Delta f_j(t) \in W_j$ has a bandpass frequency spectrum.

It follows from (39) that W_k is contained in V_j if $k < j$, and thus if it is assumed that W_j is orthogonal to V_j, it follows that W_j is orthogonal to W_k :

$$V_j \perp W_j \implies W_j \perp W_k. \tag{41}$$

This condition implies that if $f_j(t)$ is orthogonal to $\Delta f_j(t)$, then $\Delta f_j(t)$ is orthogonal to $\Delta f_k(t)$.

Two properties of the spaces V_j are obtained by considering changes of scale and translation of a function $f_j(t) \in V_j$:
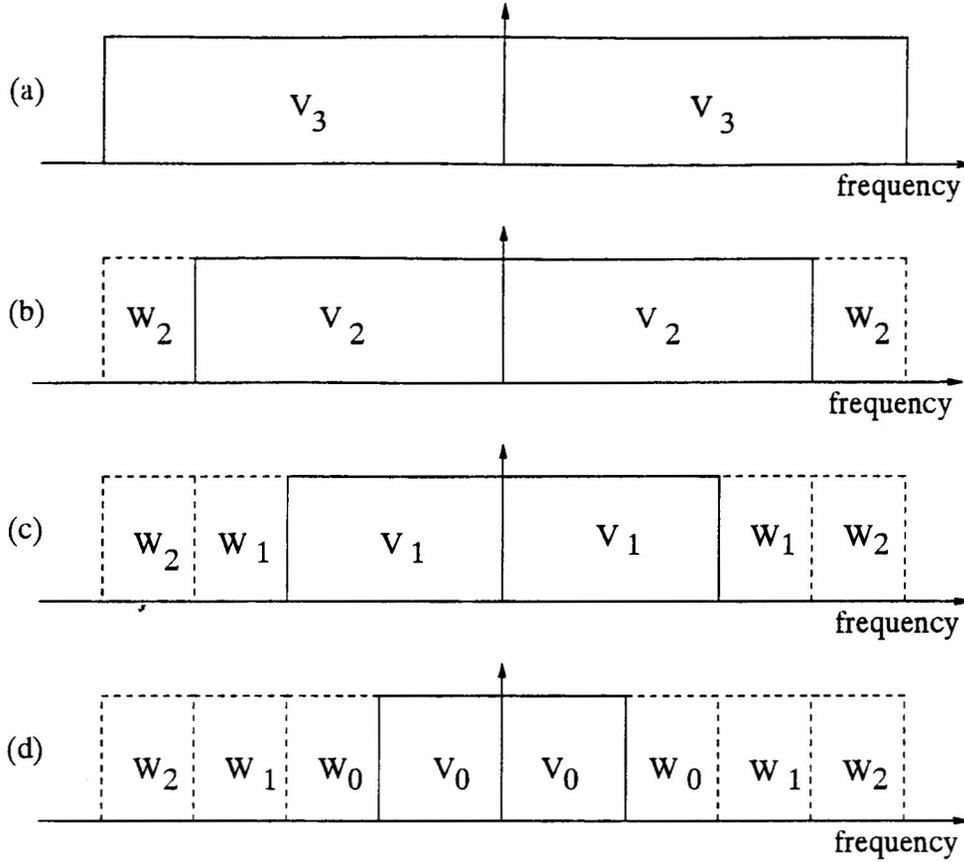
10

Figure 2: The frequency domain interpretation of multiresolution : (a) V_3, (b) $V_3 = V_2 \oplus W_2$, (c) $V_3 = V_1 \oplus W_1 \oplus W_2$, and (d) $V_3 = V_0 \oplus W_0 \oplus W_1 \oplus W_2$.

1. Consider the effect of a change in scale,

$$f_j(t) \in V_j \iff f_j(2t) \in V_{j+1},$$

or more generally,

$$f(t) \in V_0 \iff f(2^j t) \in V_j, \tag{42}$$

since scaling in time causes an inverse scaling in frequency,

$$f(t) \leftrightarrow F(\omega) \iff f\left(2^j t\right) \leftrightarrow \frac{1}{2^j} F\left(\frac{\omega}{2^j}\right),$$

where $F(\omega)$ is the Fourier transform of $f(t)$.

2. The translation of $f_j(t)$ to $f_j(t - k)$ does not cause a change in the space V_j to which it belongs,

$$f_j(t) \in V_j \iff f_j(t - k) \in V_j, \tag{43}$$

because the translation of a function only causes a change in the phase of its Fourier transform,

$$f(t) \leftrightarrow F(\omega) \iff f\left(t - k\right) \leftrightarrow e^{-i\omega k} F\left(\omega\right).$$

The two extreme limits on the space V_j are obtained by considering $j \to \pm\infty$.

11

1. As $j \to -\infty$, the space V_j contains the constant function, and since $V_j \subset L^2(\mathbb{R})$, the constant function must be the zero function,

$$\lim_{j \to -\infty} V_j = \{0\} \qquad \text{or} \qquad \bigcap_{j=-\infty}^{\infty} V_j = \{0\}. \qquad (44)$$

This is called the emptiness condition and implies that $\|f_j(t)\| \to 0$ as $j \to -\infty$.

2. As $j \to \infty$, all possible scales or levels of detail are contained in V_j and the completeness condition,

$$\lim_{j \to \infty} V_j = L^2(\mathbb{R}) \qquad \text{or} \qquad \text{Closure} \left(\bigcup_{j=-\infty}^{\infty} V_j \right) = L^2(\mathbb{R}), \qquad (45)$$

is satisfied. The closure guarantees that the limits of all vectors in the subspaces V_j are included [8], page 52. In particular, the union $V_{-\infty}$ of all the nested subspaces V_j is not the same as the space $L^2(\mathbb{R})$ but is dense in $L^2(\mathbb{R})$ since for every function $f(t) \in L^2(\mathbb{R})$, there exists a function in $V_{-\infty}$ that is arbitrarily close to $f(t)$.

It follows from recursive application of (37) that

$$V_{j+1} = V_{-j} \oplus W_{-j} \oplus W_{-j+1} \oplus \cdots \oplus W_{j-1} \oplus W_j,$$

and the emptiness condition (44) and completeness condition (45) yield

$$L^2(\mathbb{R}) = \bigoplus_{j=-\infty}^{\infty} W_j. \qquad (46)$$

It follows that if the orthogonality condition (41) of the subspaces W_j is satisfied, then these spaces form an orthogonal decomposition of the space of square integrable functions.

This section has considered some properties that the spaces V_j must satisfy but there exists one more requirement that must be considered for a complete description of a multiresolution analysis. In particular, it is assumed that there exists a function $\phi_0(t) = \phi(t)$ such that the integer translates $\{\phi(t - k) : k \in \mathbb{Z}\}$ of $\phi(t)$ form an orthonormal basis for V_0. It will be shown that this enables an orthonormal basis for the space V_j to be developed.

These requirements of a multiresolution analysis are collected together in the following definition.

Definition 4.1 A multiresolution analysis consists of a sequence of nested subspaces (38) such that the following conditions are satisfied :

1. Scale invariance : Equation (42).

2. Translation invariance : Equation (43).

3. Emptiness : Equation (44).

4. Completeness : Equation (45).

5. There exists a function $\phi_0(t) = \phi(t)$ such that $\{\phi(t - k) : k \in \mathbb{Z}\}$ is an orthonormal basis of V_0. $\square$

The next sections consider the dilation and wavelet equations and the decomposition relation, which enable expressions for $\phi(t)$ and an orthonormal basis for W_j to be developed.

4.2　The dilation equation

The nested property of the subspaces V_j implies that if $f(t) = \phi(t - k) \in V_0$, where $\{\phi(t - n) : n \in \mathbb{Z}\}$ is an orthonormal basis for V_0, then $f\left(2^j t\right) = \phi\left(2^j t - k\right) \in V_j$, and thus

$$\left\{\phi_j(t) = 2^{\frac{i}{2}}\phi\left(2^j t - k\right) : k \in \mathbb{Z}\right\}, \tag{47}$$

is an orthonormal basis for V_j. The abbreviated notation

$$\phi_{jk}(t) = 2^{\frac{i}{2}}\phi\left(2^j t - k\right),$$

denotes the kth translate of the jth basis function. The scale factor $2^{\frac{i}{2}}$ is included so that the basis functions are normalised for all scales j and translations k,

$$\|\phi_{jk}(t)\|^2 = \int_{-\infty}^{\infty}\left|2^{\frac{i}{2}}\phi\left(2^j t - k\right)\right|^2\,\mathrm{dt} = \int_{-\infty}^{\infty}|\phi(t)|^2\,\mathrm{dt}.$$

By assumption, $V_0 \subset V_1$ and since $\phi(t) \in V_0$, it follows that $\phi(t) \in V_1$ and thus there exist constants $c(n)$ such that

$$\phi(t) = \sqrt{2}\sum_k c(k)\phi(2t - k). \tag{48}$$

This is the dilation equation, also called the two–scale or refinement equation because it relates the basis functions at level $j = 0$ to the basis functions at level $j = 1$. It will be shown in section 4.5 that the coefficients $c(n)$ are exactly equal to the coefficients $h_0(n)$ of the lowpass filter $H_0(z)$ in an orthogonal filter bank, and thus the dilation equation must be solved for the scaling function $\phi(t)$ for given constants $c(n)$. This equation is usually solved by the cascade algorithm, which is an iterative method.

The equivalence between the constants $c(n)$ and filter coefficients $h_0(n)$ is suggested by considering the conditions that these constants must satisfy to guarantee that the functions $\{\phi_{0k}(t)\} = \{\phi(t - k)\}$ are orthonormal. In particular, it follows from (48) that

$$\phi(t - m) = \sqrt{2}\sum_k c(k)\phi(2t - 2m - k),$$

and hence

$$
\begin{aligned}
\langle\,\phi(t), \phi(t - m)\,\rangle &= \int_{-\infty}^{\infty}\phi(t)\phi(t - m)\,\mathrm{dt} \\
&= 2\sum_{k,l}c(k)c(l)\int_{-\infty}^{\infty}\phi(2t - k)\phi(2t - 2m - l)\,\mathrm{dt} \\
&= \sum_{k,l}c(k)c(l)\int_{-\infty}^{\infty}\phi(t)\phi(t + k - 2m - l)\,\mathrm{dt} \\
&= \sum_{k,l}c(k)c(l)\delta(2m + l - k) \\
&= \sum_l c(l)c(2m + l). \tag{49}
\end{aligned}
$$

Since the set of functions $\{\phi(t - k)\}$ are orthonormal, it follows that

$$\langle\,\phi(t), \phi(t - m)\,\rangle = \sum_k c(k)c(k - 2m) = \delta(m), \tag{50}$$

13

and thus the coefficients $c(n)$ have double shift orthogonality, which characterises the conditions on the coefficients of an orthogonal bank, (33)–(35). The full equivalence between orthogonal filter banks and multiresolution will be established by considering the wavelet equation and decomposition relation.

Equation (48) is obtained by considering $V_0 \subset V_1$, but it is valid for all subspaces, that is, $V_j \subset V_{j+1}$. This follows by making the substitution $t \to 2^j t - m$ and multiplying both sides of the equation by $2^{\frac{i}{2}}$,

$$2^{\frac{i}{2}}\phi(2^j t - m) = 2^{\frac{i+1}{2}} \sum_k c(k)\phi(2^{j+1}t - 2m - k). \tag{51}$$

The function $f_j(t)$, defined by

$$f_j(t) = \sum_k a_{jk}\phi_{jk}(t), \tag{52}$$

is the orthogonal projection of a function $f(t) \in V_m$, $j < m$, onto the space V_j. This projection is a low level approximation of $f(t)$, and it will be shown in section 4.6 that a fast recursion enables the coefficients a_{jk} in the space V_j to be calculated.

4.3 The wavelet equation

The derivation of the wavelet equation is similar to that of the dilation equation, except that the space V_0 is replaced by the space W_0. In particular, since $W_0 \subset V_1$, it follows that if $\psi(t) \in W_0$, then $\psi(t) \in V_1$ and thus there exist constants $d(n)$ such that

$$\psi(t) = \sqrt{2} \sum_k d(k)\phi(2t - k). \tag{53}$$

It will be shown that the constants $d(n)$ are equal to the coefficients of the highpass filter $H_1(z)$ in an orthogonal filter bank, in the same way as the coefficients $c(n)$ of the dilation equation are equal to the coefficients of the lowpass filter in the filter bank. Thus the wavelet equation (53) yields the mother wavelet $\psi(t)$ directly from the scaling function. The condition (50) on the coefficients $c(n)$ arises from the orthogonality of the basis functions $\{\phi(t - k)\}$, and similarly, a condition on the coefficients $d(k)$ can be deduced by imposing the orthogonality condition (41). If the inner product of both sides of (53) with $\phi(t - m)$ is taken, then following (49) identically, the result

$$\sum_k c(k)d(k - 2m) = 0,$$

is obtained, and it is noted that this equation has the same form as (35). Although $\psi(t) \in W_0$, an orthogonal basis for W_0 has not yet been developed, but the decomposition relation of $\phi(t)$ and $\psi(t)$, which is considered in the next section, ensures that $\{\psi(t - k) : k \in \mathbb{Z}\}$ generates the whole space W_0.

4.4 The decomposition relation

It follows from (37) that if $\{\psi(t - k) : k \in \mathbb{Z}\}$ generates the whole space W_0, there must exist constants $a(2k)$ and $b(2k)$ such that

$$\sqrt{2}\phi(2t) = \sum_k a(2k)\phi(t - k) + \sum_k b(2k)\psi(t - k), \tag{54}$$

is satisfied. The even indexed coefficients of the unknown coefficients are used because this simplifies the following analysis. Let $2k \to 2k - l$, in which case (54) becomes

$$\sqrt{2}\phi(2t) = \sum_k a(2k - l)\phi\left(t + \frac{l}{2} - k\right) + \sum_k b(2k - l)\psi\left(t + \frac{l}{2} - k\right),$$

and the substitution $t \to t - \frac{l}{2}$ yields

$$\sqrt{2}\phi(2t - l) = \sum_k a(2k - l)\phi(t - k) + \sum_k b(2k - l)\psi(t - k), \tag{55}$$

and thus

$$\langle \phi(t), \phi(2t - l) \rangle = \frac{1}{\sqrt{2}}a(-l).$$

From the dilation equation,

$$\langle \phi(t), \phi(2t - l) \rangle = \frac{1}{\sqrt{2}}c(l),$$

and hence

$$a(l) = c(-l), \tag{56}$$

which is the solution for the coefficients $a(2k)$ in (54). The expression for the coefficients $b(2k)$ follows by taking the inner product of (55) with $\psi(t)$,

$$\sqrt{2}\langle \phi(2t - l), \psi(t) \rangle = \sum_k b(2k - l)\langle \psi(t - k), \psi(t) \rangle,$$

and from the wavelet equation,

$$\sqrt{2}\langle \phi(2t - l), \psi(t) \rangle = d(l),$$

and thus

$$d(l) = \sum_k b(2k - l)\langle \psi(t - k), \psi(t) \rangle.$$

If $\psi(t)$ is orthogonal to all its integer translates and normalised to unit magnitude, then it follows from this equation that

$$b(l) = d(-l). \tag{57}$$

Equations (56) and (57) are the solutions of the decomposition relation (54) if $\psi(t)$ is orthogonal to all its integer translates. The substitution of these solutions into (55) yields

$$\sqrt{2}\phi(2t - l) = \sum_k c(l - 2k)\phi(t - k) + \sum_k d(l - 2k)\psi(t - k),$$

and the transformation $t \to 2^j t$ followed by multiplication by 2^j yields

$$2^{\frac{j+1}{2}}\phi(2^{j+1}t - l) = 2^{\frac{j}{2}}\sum_k c(l - 2k)\phi(2^j t - k) + 2^{\frac{j}{2}}\sum_k d(l - 2k)\psi(2^j t - k),$$

which is the decomposition relation between levels j and $j + 1$, that is, $V_{j+1} = V_j \oplus W_j$. Thus (54) has a solution between levels $j = 0$ and $j = 1$ and hence there exists a solution between any two levels. It follows that

$$\left\{ \psi_j(t) = 2^{\frac{j}{2}}\psi\left(2^j t - k\right) : k \in \mathbb{Z} \right\}, \tag{58}$$

15

is an orthonormal basis for the space W_j.

The substitution $t \to 2^j t - m$ in the wavelet equation and multiplying both sides by $2^{\frac{i}{2}}$ yields

$$2^{\frac{i}{2}}\psi(2^j t - m) = 2^{\frac{i+1}{2}}\sum_k d(k)\phi(2^{j+1}t - 2m - k), \tag{59}$$

which is a generalisation of the wavelet equation from $W_0 \subset V_1$ to $W_j \subset V_{j+1}$. This is equivalent to (51), which is the dilation equation between V_j and V_{j+1}. The abbreviated notation

$$\psi_{jk}(t) = 2^{\frac{i}{2}}\psi\left(2^j t - k\right),$$

denotes the kth translate of the jth basis function. Since $W_j \perp W_k$, it follows that

$$\langle\, \psi_{jk}(t), \psi_{mn}(t)\, \rangle = \left\{ \begin{array}{ll} 1 & \text{if } j = m \text{ and } k = n \\ 0 & \text{otherwise.} \end{array} \right.$$

The implications on the wavelet equation of the orthonormality of the functions (58) must be considered. Following the procedure in (49), it is easily shown that

$$\langle\, \psi(t), \psi(t - m)\, \rangle = \sum_k d(k)d(k - 2m) = \delta(m),$$

and hence the orthonormality of the basis functions (58) implies a double shift orthogonality of the coefficients $d(n)$.

Equation (46) shows that the spaces W_j form an orthogonal decomposition of $L^2(\mathbb{R})$, and thus there exist constants b_{jk} such that

$$f(t) = \sum_j \sum_k b_{jk}\psi_{jk}(t). \tag{60}$$

The functions (58) are the wavelet basis functions at level j, and (53) shows that the mother wavelet $\psi_{00}(t) = \psi(t)$ is derived directly from the wavelet equation. The physical interpretation of (60) follows from figure 2, in which it is shown that each basis function in the set (58) has a bandpass frequency spectrum, and thus the level of detail or approximation increases as j increases. Finally, it will be shown in section 4.6 the coefficients b_{jk} can be computed recursively and quickly from the coefficients a_{jk} in (52).

The next section combines the results on orthogonal filter banks and multiresolution analysis, and this enables their equivalence to be considered.

4.5 Multiresolution and orthogonal filter banks

Three equations, the dilation equation, the wavelet equation and the decomposition relation have generated three conditions on the coefficients $c(n)$ and $d(n)$. These conditions are summarised and the identification of the coefficients $c(n)$ and $d(n)$ with the filter coefficients $h_0(n)$ and $h_1(n)$ is established.

1. The dilation equation and the orthonormality of the functions $\{\phi(t - k) : k \in \mathbb{Z}\}$:

$$\sum_k c(k)c(k - 2m) = \delta(m). \tag{61}$$

2. The wavelet equation and the orthogonality of the spaces V_0 and W_0 :

$$\sum_k c(k)d(k - 2m) = 0. \tag{62}$$

16

3. The decomposition relation and the orthonormality of the functions $\{\psi(t-k) : k \in \mathbb{Z}\}$:

$$\sum_k d(k)d(k-2m) = \delta(m). \tag{63}$$

The link between orthogonal filter banks and multiresolution analysis follows directly by comparing (33)–(35) and (61)–(63); both sets of equations involve a double shift orthogonality, and thus the coefficients $c(n)$ and $d(n)$ in a multiresolution analysis are also the filter coefficients in an orthogonal filter bank. Moreover, the alternating flip guarantees that if (33) is satisfied, then it necessarily follows that (34) and (35) are also satisfied, and hence the alternating flip can also be used to define the coefficients $d(n)$ in terms of the coefficients $c(n)$,

$$d(k) = (-1)^k c(N-k), \qquad k = 0 \ldots N, \tag{64}$$

where N is odd. The only remaining issue is the identification of $c(n)$ with the lowpass filter coefficients $h_0(n)$ rather than the highpass filter coefficients $h_1(n)$. Since a necessary (but not sufficient) condition for the convergence of the cascade algorithm for the scaling function $\phi(t)$ is that the filter $C(\omega) = \sum c(k)e^{-i\omega k}$ satisfies $C(\omega = \pi) = 0$, [11] page 234, it follows that the coefficients $c(n)$ must be identified with the lowpass filter, and thus the coefficients $d(n)$ are identified with the highpass filter.

4.6 The fast wavelet transform

It has been shown that a multiresolution analysis of a function requires two sets of spaces, the approximation spaces V_j and the detail spaces W_j. A function in V_{j+1} can be written as the sum of a function in V_j and a function in W_j, and this process can be repeated in a recursive manner. The function $f_{j-1}(t) \in V_{j-1}$ is a low level approximation to $f_j(t) \in V_j$, and it is necessary to calculate the coefficients $a_{j-1,k}$ and $b_{j-1,k}$ from a_{jk}. This calculation can be done in a fast recursive manner called the fast wavelet transform (FWT), which is derived in theorem 4.1. This is an analysis transform because it enables the coefficients of a low level approximation of $f_j(t)$ to be calculated. By contrast, the inverse fast wavelet transform (IFWT), which is derived in theorem 4.2, is used for the synthesis of $f_j(t)$ from its low level approximations. It will be shown that these recursions can be implemented in a filter bank in which the coefficients a_{jk} are generated by the lowpass filter with coefficients $c(-n)$ and the wavelet coefficients b_{jk} are generated by the highpass filter with coefficients $d(-n)$, that is, aniticausal filters are required. Two important issues that must be considered for the implementation of the fast wavelet transform are the initialisation of the recursion and the effect of signals of finite length. Both these points will be discussed after the FWT and IFWT have been derived.

Theorem 4.1 A function

$$f_{j+1}(t) = \sum_l a_{j+1,l}\phi_{j+1,l}(t) \in V_{j+1} = V_j \oplus W_j, \tag{65}$$

has coefficients a_{jk} and b_{jk},

$$a_{jk} = \sum_l c(l-2k)a_{j+1,l} \qquad \text{and} \qquad b_{jk} = \sum_l d(l-2k)a_{j+1,l}, \tag{66}$$

in the spaces V_j and W_j respectively.

Proof Consider the dilation equation (51),

$$\phi_{jm}(t) = \sum_l c(l-2m)\phi_{j+1,l}(t). \tag{67}$$

The equivalent equation that is derived from the wavelet equation is (59),

$$\psi_{jm}(t) = \sum_l d(l - 2m)\phi_{j+1,l}(t). \tag{68}$$

The recursion for the coefficients a_{jk} in (66) is derived by taking the inner product of both sides of (67) with $f_{j+1}(t)$, which is defined in (65),

$$
\begin{aligned}
\langle\, f_{j+1}(t), \phi_{jm}(t)\, \rangle &= \sum_l c(l - 2m) \,\langle\, f_{j+1}(t), \phi_{j+1,l}(t)\, \rangle \\
&= \sum_l c(l - 2m) \,\langle\, \sum_k a_{j+1,k}\,\phi_{j+1,k}(t), \phi_{j+1,l}(t)\, \rangle \\
&= \sum_l c(l - 2m) \sum_k a_{j+1,k}\,\langle\, \phi_{j+1,k}(t), \phi_{j+1,l}(t)\, \rangle \\
&= \sum_l c(l - 2m) \sum_k a_{j+1,k}\delta(k - l) \\
&= \sum_l c(l - 2m)a_{j+1,l}. \tag{69}
\end{aligned}
$$

The left hand side of (69) is simplified by using the decomposition of $f_{j+1}(t) \in V_{j+1}$ into a portion in V_j and a portion in W_j,

$$f_{j+1}(t) = \sum_l a_{j+1,l}\phi_{j+1,l}(t) = \sum_l a_{jl}\phi_{jl}(t) + \sum_l b_{jl}\psi_{jl}(t), \tag{70}$$

and since V_j is orthogonal to W_j,

$$
\begin{aligned}
\langle\, f_{j+1}(t), \phi_{jm}(t)\, \rangle &= \langle\, \sum_l a_{jl}\phi_{jl}(t) + \sum_l b_{jl}\psi_{jl}(t), \phi_{jm}(t)\, \rangle \\
&= \langle\, \sum_l a_{jl}\phi_{jl}(t), \phi_{jm}(t)\, \rangle \\
&= \sum_l a_{jl}\delta(l - m) \\
&= a_{jm}.
\end{aligned}
$$

The combination of this result and (69) establishes the recursion (66) for the coefficients a_{jk}. The recursion for the wavelet coefficients follows identically, except that (68) is used instead of (67). $\square$

The recursions (66) represent the convolution of $\{a_{j+1}\}$ with the anticausal filters $c(-n)$ and $d(-n)$, followed by downsampling, as shown in figure 3.

Theorem 4.2 The coefficients $a_{j+1,l}$ in (65) can be reconstructed from their values at lower levels by

$$a_{j+1,l} = \sum_k c(l - 2k)a_{jk} + \sum_k d(l - 2k)b_{jk}. \tag{71}$$

Proof By definition,

$$
\begin{aligned}
\phi_{jk}(t) &= 2^{\frac{i}{2}}\phi(2^j t - k) \\
&= 2^{\frac{i+1}{2}} \sum_n c(n)\phi(2^{j+1}t - 2k - n) \\
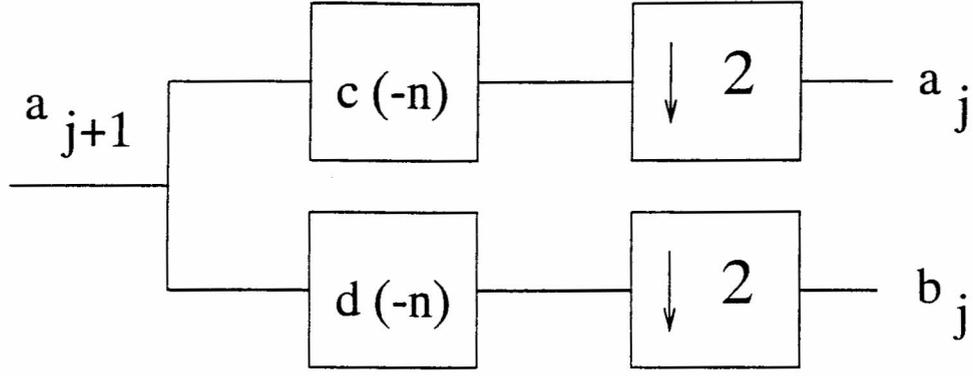&= \sum_n c(n)\phi_{j+1,2k+n}(t), \tag{72}
\end{aligned}
$$

18

Figure 3: The implementation of the fast wavelet transform as the analysis stage of a two–channel filter bank.

and similarly,

$$\psi_{jk}(t) = \sum_n d(n)\phi_{j+1,2k+n}(t). \tag{73}$$

The inner product of (70) with $\phi_{j+1,m}(t)$ and the application of (72) and (73) yields

$$
\begin{aligned}
\sum_l a_{j+1,l}\langle \phi_{j+1,l}(t), \phi_{j+1,m}(t) \rangle &= \sum_l a_{jl} \sum_n c(n)\langle \phi_{j+1,2l+n}(t), \phi_{j+1,m}(t) \rangle \\
&\quad + \sum_l b_{jl} \sum_n d(n)\langle \phi_{j+1,2l+n}(t), \phi_{j+1,m}(t) \rangle \\
&= \sum_l a_{jl} \sum_n c(n)\delta(m - 2l - n) \\
&\quad + \sum_l b_{jl} \sum_n d(n)\delta(m - 2l - n) \\
&= \sum_l c(m - 2l)a_{jl} + \sum_l d(m - 2l)b_{jl}.
\end{aligned}
$$

The left hand side of this equation is $a_{j+1,m}$ and thus the result (71) is established. $\square$

The inverse fast wavelet transform (71) is implemented by upsampling $\{a_j\}$ and $\{b_j\}$ and then filtering with, respectively, filters with coefficients $c(n)$ and $d(n)$, as shown in figure 4.

The combination of figures 3 and 4 shows that the approximation of $f_{j+1}(t)$ by $f_j(t)$, and then its synthesis from $f_j(t)$ and $\Delta f_j(t)$ is implemented by an orthogonal filter bank. More importantly, the fast wavelet transform is implemented by a tree of filter banks, as shown in figure 5. It is seen that the coefficients $\{a_j\}$ are defined by the outputs of the lowpass filters and the coefficients $\{b_j\}$ are defined by the outputs of the highpass filters.

The filters $c(-n)$ and $d(-n)$ are the anticausal forms of the filters $h_0(n)$ and $h_1(n)$ in figure 1, and thus a multiresolution analysis is implemented as a tree of analysis filter banks, where the filters satisfy the conditions of an orthogonal filter bank. There exist two further issues, the initialisation of the fast wavelet transform and the effect of signals of finite length, that must be considered before the transform can be implemented.

The initialisation of the fast wavelet transform needs special attention because the assumption of multiresolution analysis of a function $f(t)$ requires that it lie in a space V_j. The simplest assumption is that the sampled values $f(n)$ of $f(t)$ are used as input to the filter bank. Assuming that the transform is initiated at level $j = 0$, it follows that there must exist
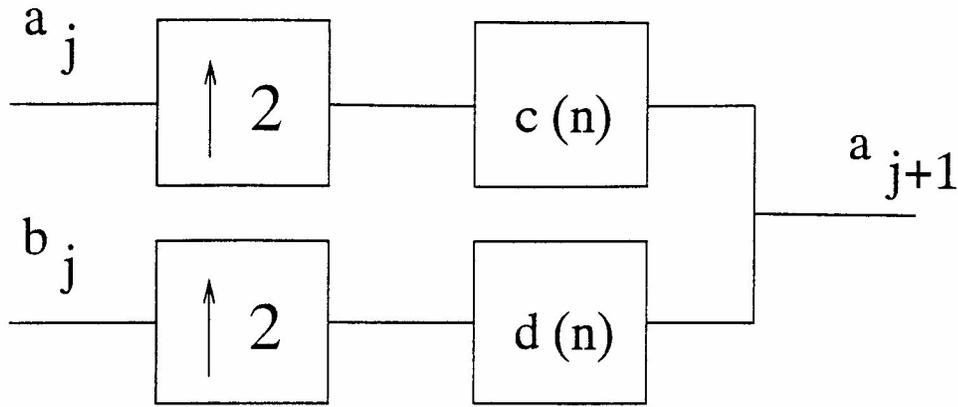
Figure 4: The implementation of the inverse fast wavelet transform as the synthesis stage of a two–channel filter bank.
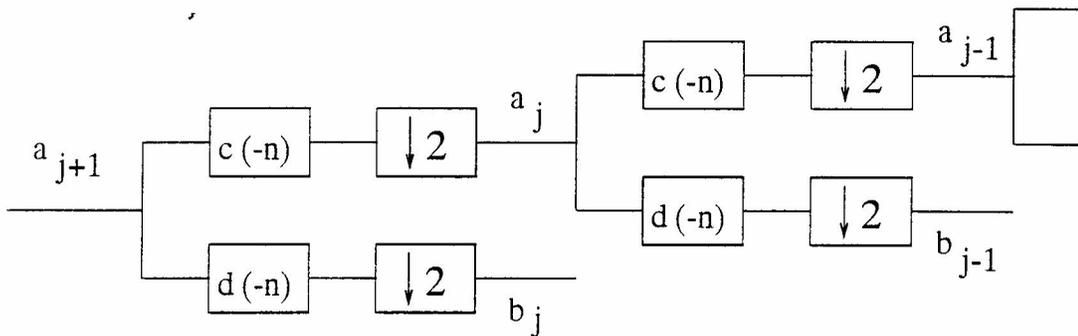

Figure 5: The fast wavelet transform implemented as a tree of filter banks.

a scaling function $\phi(t)$ such that

$$f_s(t) = \sum_n f(n)\phi(t - n),$$

where $f_s(t)$ is the underlying continuous function that is defined by the samples $f(n)$. This equation is satisfied by the delta function, $\phi(t) = \delta(t)$, but most scaling functions do not satisfy it. A better approach requires that the coefficients a_{jk} interpolate exactly the function values [2], chapter 6, [14], [15].

Consider now the implementation of the fast wavelet transform for signals of finite length. A basic assumption of wavelet theory is that the signals are of infinite duration, but all signals are of finite duration. If the signal is sufficiently long, the assumption of a signal of infinite length is adequate in the middle of the signal where boundary effects need not be considered. However problems occur at the boundary of the signal, where the sample values $x(n)$ for $n < 0$ and $n > L$ are not defined for a signal of length $L+1$. This finite length must be considered, and the problem is filtering this class of signal because the convolution $\sum h(k)x(n-k)$ is not always defined. It is necessary to extend the signal $x(n)$ such that the convolution is defined, and there exist two possibilities, periodicity (wraparound) and extension by reflection (symmetric extension) [11], [13]. It is noted that extension by zeros (extrapolation by zeros) yields poor results and is not considered satisfactory.

5 Examples

Wavelets are used extensively in many areas of applied mathematics, and a few of these applications are listed :

1. Image and video compression [8], [11].

2. Speech analysis, audio and ECG compression [4], [11].

3. Medicine and biology [1].

4. Statistics [6],[9].

5. Computational linear algebra and the solution of Poisson's equation [12].

6. The analysis of economic and financial data [7].

7. Image processing [3].

8. Computer graphics [10].

9. Vibrations and acoustics [5].

References

[1] Aldroubi A., and Unser M., *Wavelets in Medicine and Biology*, Boca Raton, FL : CRC Press, (1996).

[2] Chui C., *Wavelets : A Mathematical Tool for Signal Analysis*, Philadelphia, PA: SIAM, (1997).

[3] Kingsbury N., "Image processing with complex wavelets", *Phil. Trans. R. Soc. Lond. A.*, **357**, pp. 2543–2560, (1999).

[4] Kobayashi M., Sakamoto M., Saito T., Hashimoto Y., Nishimura M., and Suzuki K., "Wavelet analysis for a text–to–speech (TTS) system", in: *Wavelets and Their Applications : Case Studies* (Kobayashi M., Ed.), Philadelphia, PA: SIAM, pp. 75–100, (1998).

[5] Newland D., "Harmonic wavelets in vibrations and acoustics", *Phil. Trans. R. Soc. Lond. A.*, **357**, pp. 2607–2625, (1999).

[6] Ogden R., *Essential Wavelets for Statistical Applications and Data Analysis*, Boston, MA: Birkhäuser, (1997).

[7] Ramsey J., "The contribution of wavelets to economic and financial data", *Phil. Trans. R. Soc. Lond. A.*, **357**, pp. 2593–2606, (1999).

[8] Rao R., and Bopardikar A., *Wavelet Transforms : Introduction to Theory and Applications*, Reading, MA: Addison–Wesley, (1998).

[9] Silverman B., "Wavelets in statistics : beyond the standard assumptions", *Phil. Trans. R. Soc. Lond. A.*, **357**, pp. 2459–2473, (1999).

[10] Stollnitz E., DeRose T., and Salesin D., *Wavelets for Computer Graphics : Theory and Applications*, San Francisco, CA: Morgan Kaufmann, (1996).

[11] Strang G., and Nguyen T., *Wavelets and Filter Banks*, Wellesley, MA: Wellesley–Cambridge Press, (1997).

[12] Tanata N., "A wavelet based conjugate gradient method for solving Poisson equations", in: *Wavelets and Their Applications : Case Studies* (Kobayashi M., Ed.), Philadelphia, PA: SIAM, pp. 45–73, (1998).

[13] Unser M.,"A practical guide to the implementation of the wavelet transform", in: *Wavelets in Medicine and Biology* (A. Aldroubi and M. Unser, Eds.), CRC Press, pp. 37–73, (1996).

[14] Unser M., Aldroubi A., and Eden M., "A family of polynomial spline wavelet transforms", *Signal Process.*, **30**, pp. 141–162, (1993).

[15] Unser M., Aldroubi A., and Eden M., "The L_2 polynomial spline pyramid", *IEEE Trans. Pattern Anal. Mach. Intell.*, **15**, pp. 364–379, (1993).

[16] Vaidyanathen P., *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice Hall, (1993).