Eigenvalues of $(\downarrow 2)H$ and Convergence of the Cascade Algorithm

Gilbert Strang

Department of Mathematics

Massachusetts Institute of Technology

gs@math.mit.edu

Abstract

This paper is about the eigenvalues and eigenvectors of $(\downarrow 2)H$. The ordinary FIR filter H is convolution with a vector $h = (h(0), \ldots, h(N))$, the impulse response. The operator $(\downarrow 2)$ downsamples the output $y = h * x$, keeping the even-numbered components $y(2n)$. Where H is represented by a constant-diagonal matrix — this is a Toeplitz matrix with $h(k)$ on its kth diagonal — the odd-numbered rows are removed in $(\downarrow 2)H$. The result is a double shift between rows, yielding a block Toeplitz matrix with 1×2 blocks.

Iteration of the filter is governed by the eigenvalues. If the transfer function $H(z) = \sum h(k)z^{-k}$ has a zero of order p at $z = -1$, corresponding to $\omega = \pi$, then $(\downarrow 2)H$ has p special eigenvalues $\frac{1}{2}, \frac{1}{4} \ldots, \left(\frac{1}{2}\right)^p$. We show how each additional "zero at π" divides all eigenvalues by 2 and creates a new eigenvector for $\lambda = \frac{1}{2}$. This eigenvector solves the dilation equation $\phi(t) = 2\sum h(k)\phi(2t-k)$ at the integers $t = n$. The left eigenvectors show how $1, t, \ldots, t^{p-1}$ can be produced as combinations of $\phi(t-k)$.

The dilation equation is solved by the cascade algorithm, an infinite iteration of $M = (\downarrow 2)2H$. Convergence in L^2 is governed by the eigenvalues of $T = (\downarrow 2)2HH^T$, corresponding to the response $2H(z)H(z^{-1})$. We find a simple proof of the necessary and sufficient condition for convergence.

SP-8030

Key words: wavelets, cascade algorithm, scaling function

1 Introduction

The key operator in multirate filtering is $(\downarrow 2)H$. The input signal x is filtered by H and then downsampled. We keep the even-numbered components of Hx. This is the lowpass channel in the analysis half of a filter bank. When $(\downarrow 2)H$ is iterated, either finitely often in practice or infinitely often in the passage to scaling functions and wavelets, its eigenvalues become all-important. These eigenvalues are intimately related to the number of "zeros at π" in the frequency response, and to the equal number of "vanishing moments" in the wavelets. This note studies the eigenvalues and eigenvectors (left as well as right). We also determine when the cascade algorithm converges to the scaling function.

The cascade algorithm is the iteration $\phi^{(i+1)} = M\phi^{(i)} = (\downarrow 2)2H\phi^{(i)}$. The extra factor 2 maintains constant area for the sequence of functions $\phi^{(i+1)}(t) = \sum 2h(k)\phi^{(i)}(2t-k)$, when the filter coefficients are normalized by $\sum h(k) = 1$. With double-shift orthogonality, $\sum 2h(k)h(k+2l) = \delta(l)$, convergence to $\phi(t)$ is almost (but not quite) certain. In a *biorthogonal* filter bank, with a more general lowpass filter H, this convergence is not at all assured. Nevertheless these non-orthogonal filters are giving the best results in compression, and their condition numbers are often quite moderate. We determine when they lead to wavelets.

Part 3 presents a simple proof of the necessary and sufficient condition for convergence to $\phi(t)$. Since we work in the L^2 norm (convergence in energy), inner products play a decisive part. They lead to the "transition operator" $T = (\downarrow 2)2HH^T$, whose eigenvalues control the convergence. Those eigenvalues also determine the smoothness of the scaling function and wavelets.

We summarize now the main conclusions. Some are already in the literature, with different proofs, and some are new. The (real) filter coefficients $h(0), \ldots, h(N)$ yield the transfer function $H(z) = \sum h(n)z^{-n}$. The input signal transforms to $X(z) = \sum x(n)z^{-n}$, and the filtered signal is $H(z)X(z)$. In the z-domain, the key operators $M = (\downarrow 2)2H$ and $T = (\downarrow 2)2HH^T$ involve multiplication by $H(z)$ from the filter and an aliasing term (identified by $-z$) from the downsampling:

$$(MX)(z^2) = H(z)X(z) + H(-z)X(-z) \tag{1}$$

$$(TX)(z^2) = H(z)X(z)H(z^{-1}) + H(-z)X(-z)H(-z^{-1}). \tag{2}$$

Note the argument z^2. We are dealing with the even part, because $(\downarrow 2)$ removes the odd terms.

In the time domain, the i, k entry of M is $2h(2i-k)$. It is $2i$ that reflects the double shift from $(\downarrow 2)$. The entries of T are $2p(2i-k)$, where $P(z) = H(z)H(z^{-1})$ corresponds to HH^T. The calculations involve finite matrices, in which i and k range from 0 to $N-1$ for M and from $1-N$ to $N-1$ for T. The frequency responses of interest have p zeros at π. Thus $H(z)$ has p zeros at $z = e^{j\pi} = -1$:

$$H(z) = \left(\frac{1+z^{-1}}{2}\right)^p Q(z) \qquad \text{with} \qquad Q(-1) \neq 0.$$

We state the conclusions in the time domain (for matrix eigenvalues), where they are easiest to check. We establish those conclusions in the z-domain, where they are easiest to prove.

Theorem 1 *Each time $H(z)$ is multiplied by $\frac{1+z^{-1}}{2}$, all the eigenvalues of M are multiplied by $\frac{1}{2}$ and a new eigenvalue $\lambda = 1$ is introduced. Thus the eigenvalues of M are*

$$1, \frac{1}{2}, \ldots, \left(\frac{1}{2}\right)^{p-1} \quad \textit{together with} \quad \frac{1}{2^p} \quad \textit{times the eigenvalues for} \quad (\downarrow 2)2Q. \tag{3}$$

Theorem 2 *When $H(z)$ is multiplied by $\left(\frac{1+z^{-1}}{2}\right)$, the new eigenvectors $\tilde{x}$ are the* differences *of the previous eigenvectors and the new left eigenvectors $\tilde{y}$ are the* sums *of the previous left eigenvectors:*

$$\tilde{X}(z) = \left(1 - z^{-1}\right) X(z) \quad \textit{and} \quad \tilde{Y}(z) = \frac{Y(z)}{(1-z)}. \tag{4}$$

The extra eigenvalue $\lambda = 1$ has left eigenvector $e = [1 \;\; 1 \;\; \cdots \;\; 1]$. The right eigenvector gives the new values of the scaling function at the integers.

Theorem 3 *If $H(-1) = \sum (-1)^k h(k) = 0$, the periodized functions $P^{(i)}(t) = \sum \phi^{(i)}(t-n)$ satisfy the identity*

$$P^{(i+1)}(t) = P^{(i)}(2t) \quad \textit{and thus} \quad P^{(i)}(t) = P^{(0)}(2^i t).$$

The cascade algorithm cannot converge to $\phi(t)$ unless $P^{(0)}(t) \equiv 1$. Othervise $P^{(0)}(2^i t)$ will oscillate faster and faster. Thus Theorem 3 defines the acceptable class I of initial functions; they must satisfy $\sum \phi^{(0)}(t-n) \equiv 1$. This is equivalent to the so-called Strang-Fix condition on the Fourier transform: $\widehat{\phi}^{(0)}(2\pi n) = \delta(n)$. In that form, the requirement on $\phi^{(0)}(t)$ was discovered and proved necessary by Durand [7] and by Meyer and Paiva [14]. Our proof uses the identity $P^{(1)}(t) = P^{(0)}(2t)$.

The central question is convergence from these acceptable $\phi^{(0)}(t)$, and this is governed by the eigenvalues of T.

Theorem 4 *The cascade algorithm $\phi^{(i+1)}(t) = \sum 2h(k)\phi^{(i)}(2t-k)$ converges in L^2 for all $\phi^{(0)}$ in I if and only if the eigenvalues of T satisfy Condition* E*:*

$$\lambda = 1 \quad \textit{is a simple eigenvalue and all other eigenvalues have} \quad |\lambda| < 1. \tag{5}$$

Condition E is also the Cohen-Daubechies requirement [2] for the translates $\phi(t-k)$ to be strongly independent. Jia [10] has studied convergence and independence very carefully also in L^p.

The scaling function and wavelets are smoother by one more derivative, pointwise and in L^2, for every additional factor $(1 + z^{-1})$ in $H(z)$. Splines come from the special choice $H(z) = \left(\frac{1+z^{-1}}{2}\right)^p$. They have no orthogonality, except in Haar's piecewise constant case $p = 1$, but they have maximum smoothness. $H(z)$ has binomial coefficients $h(k)$ divided by 2^p. For $p = 2$, 3, 4 we indicate the matrix $M = (\downarrow 2)2H$ and the eigenvalues predicted by Theorem 1:

$$\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \qquad \frac{1}{4}\begin{bmatrix} 1 & 0 & 0 \\ 3 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix} \qquad \frac{1}{8}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 6 & 4 & 1 & 0 \\ 1 & 4 & 6 & 4 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\lambda = 1,\ \tfrac{1}{2} \qquad \lambda = 1,\ \tfrac{1}{2},\ \tfrac{1}{4} \qquad \lambda = 1,\ \tfrac{1}{2},\ \tfrac{1}{4},\ \tfrac{1}{8}$$

The operator $(\downarrow 2)$ produces the double-shift between rows. The matrix on the right comes from the coefficients 1, 4, 6, 4, 1. This also illustrates the matrix T for the sequence $h(k) = 1, 2, 1$ because $h * h = 1, 4, 6, 4, 1$. (Actually the first row and column are dropped in T, so the eigenvalues are $1, \frac{1}{2}, \frac{1}{4}$.) Condition E in Theorem 4 is fully satisfied, and the cascade algorithm for the $\frac{1}{4}(1, 2, 1)$ filter converges quickly to the hat function—the linear spline.

In summary, the factor $(1 + z^{-1})^p$ gives the zeros at π that produce flatness of $H(z)$ and smoothness of $\phi(t)$. This pth order zero has important effects:

- p vanishing moments for the wavelets
- p sum rules for the coefficients $h(k)$
- pth order accuracy in approximation $f(t) \approx \sum_k a_k \phi(t-k)$
- pth order decay of wavelet coefficients for a smooth $f(t) = \sum b_{jk} w_{jk}(t)$
- All polynomials of degree $< p$ are combinations of the translates $\phi(t-k)$.

The smoothness of $\phi(t)$ is measured in the L^2 norm by using Parseval's equality. The scaling function has s derivatives when $|\omega|^s \hat{\phi}(\omega)$ has finite energy. The supremum $s_{\max}$ depends on p and the largest eigenvalue $|\lambda_{\max}|$ of $(\downarrow 2)2QQ^T$:

$$s_{\max} = p - \log_4 |\lambda_{\max}| \,. \tag{6}$$

Villemoes [18,19] has given a particularly neat analysis of this smoothness formula. It is the eigenvalue $\lambda_{\max}$ from $Q(z)$ that has no simple expression (but is easily computed). Then Theorem 1 shows how the factor $(\frac{1+z^{-1}}{2})^{2p}$ in HH^T divides it by $2^{2p} = 4^p$. Smaller eigenvalues of T mean more smoothness of $\phi(t)$ and the wavelets.

2 Eigenvalues and Eigenvectors of M

Theorems 1 and 2 will be proved together. By identifying the change in eigenvectors when $H(z)$ is multiplied by $\left(\frac{1+z^{-1}}{2}\right)$, we also confirm that the eigenvalues are cut in half. Starting from $Q(z)$ with no zeros at π, this multiplication occurs p times to reach the final $H(z)$ with p zeros at π. We go one step at a time, monitoring the eigenvectors. By equation (1), $(\downarrow 2)2Hx = \lambda x$ means

$$H(z)X(z) + H(-z)X(-z) = \lambda X(z^2). \tag{7}$$

Theorem 2 states that the step to $\left(\frac{1+z^{-1}}{2}\right)H(z)$ produces the new eigenfunction $\tilde{X}(z)$ $= \left(1 - z^{-1}\right)X(z)$ with eigenvalue $\tilde{\lambda} = \frac{1}{2}\lambda$. If this is true, then equation (7) will hold for $\tilde{H}(z)$ and $\tilde{X}(z)$ and $\tilde{\lambda}$:

$$\left(\frac{1+z^{-1}}{2}\right)H(z)\left(1 - z^{-1}\right)X(z) + \left(\frac{1-z^{-1}}{2}\right)H(-z)\left(1 + z^{-1}\right)X(-z) = \frac{\lambda}{2}\left(1 - z^{-2}\right)X\left(z^2\right). \tag{8}$$

To verify (8), multiply (7) by $\frac{1}{2}\left(1 - z^{-2}\right)$. That is the only step in the proof. Daubechies [5, p.228] proved in a different way that $M = (\downarrow 2)2H$ has eigenvalues $1, \frac{1}{2}, \ldots, (\frac{1}{2})^{p-1}$.

Now consider the left eigenvectors. These are right eigenvectors of $M^T = 2H^T(\downarrow 2)^T$. In the z-domain this transposed operator takes $Y(z)$ into $2H(z^{-1})Y(z^2)$. Thus $yM = \lambda y$ means

$$2H(z^{-1})Y(z^2) = \lambda Y(z). \tag{9}$$

Theorem 2 says that equation (9) remains correct for $\tilde{H}(z) = \left(\frac{1+z^{-1}}{2}\right)H(z)$ and $\tilde{\lambda} = \frac{1}{2}\lambda$ when the eigenvector transforms to $\tilde{Y}(z) = \frac{Y(z)}{1-z}$. The left side is multiplied by $\frac{1+z}{2}$ and divided by $1 - z^2$. The right side is divided by $2(1 - z)$. This agreement completes the proof of Theorem 2.

In the time domain, note how $\frac{1}{1-z} = -z^{-1} - z^{-2} - z^{-3} - \cdots$ yields *minus* the sum (with a delay). There is also a constant of summation, whose role will become clear. Thus the pattern for left eigenvectors with each new zero at $z = -1$ is

$$y = [y(0) \;\; y(1) \;\; \cdots] \;\longrightarrow\; \tilde{y} = [C \;\; C - y(0) \;\; C - y(0) - y(1) \;\; \cdots]. \tag{10}$$

We can illustrate these eigenvectors in specific examples. A sequence of frequently studied filters begins with $Q(z) = -\frac{1}{2} + 2z^{-1} - \frac{1}{2}z^{-2}$. Each time we add a zero at π, we get a potentially useful filter. After four steps we reach the maxflat halfband filter H_4,

$$H_4(z) = \frac{1}{16}\left(-1 + 9z^{-2} + 16z^{-3} + 9z^{-4} - z^{-6}\right).$$

Its spectral factor is the Daubechies filter D_4. Thus H_4 has four zeros at π and D_4 has $p = 2$.

The first matrix is $M_0 = (\downarrow 2)2Q$. When diagonalized by $M_0 = S\Lambda S^{-1}$, the columns of S hold the eigenvectors x and the rows of S^{-1} display the left eigenvectors y:

$$M_0 = \begin{bmatrix} -1 & 0 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ .2 & 1 \end{bmatrix} \begin{bmatrix} -1 & \\ & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -.2 & 1 \end{bmatrix} \begin{matrix} \leftarrow y_1 \\ \leftarrow y_2 \end{matrix} \tag{11}$$

The biorthogonality $y_j x_k = \delta_{jk}$ comes from $S^{-1}S = I$. For the *infinite* matrix $(\downarrow 2)2Q$, the right eigenvectors are extended by zeros (but not the left eigenvectors):

$$\begin{bmatrix} \cdot & & & & \\ \cdot & 4 & -1 & & \\ & & -1 & 4 & -1 \\ & & & & -1 & \cdot \\ & & & & & \cdot \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -.2 \\ 0 \\ \cdot \end{bmatrix} = (-1) \begin{bmatrix} 0 \\ 1 \\ -.2 \\ 0 \\ \cdot \end{bmatrix}.$$

The odd rows have been removed by $(\downarrow 2)$, and M_0 is the central 2×2 submatrix.

Now multiply $Q(z)$ by $\left(\frac{1+z^{-1}}{2}\right)$. The new filter H_1 has coefficients $\frac{1}{4}(-1, 3, 3, -1)$. The new matrix M_1 is $(\downarrow 2)2H_1$. It is one size larger, with the earlier eigenvalues cut in half. The new eigenvalue $\lambda = 1$ is guaranteed by the fact that *each column of the matrix adds to* 1. This comes directly from the zero at $z = -1$:

$$H(1) = 1 \quad \text{and} \quad H(-1) = 0 \qquad \Longleftrightarrow \qquad \sum_{\text{even } k} 2h(k) = \sum_{\text{odd } k} 2h(k) = 1.$$

When the columns of a matrix add to 1, the vector $[1 \;\; 1 \;\; \cdots \;\; 1]$ is a left eigenvector y_1 for $\lambda = 1$.

$$M_1 = \frac{1}{2}\begin{bmatrix} -1 & 0 & 0 \\ 3 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ .5 & -.8 & 1 \\ .5 & -.2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -.5 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ .3 & .5 & -.5 \end{bmatrix} \begin{matrix} \leftarrow y_1 \\ \leftarrow y_2 \\ \leftarrow y_3 \end{matrix}$$

The right eigenvector $(0, .5, .5)$ for $\lambda = 1$ gives the values of the scaling function $\phi(t)$ at the integers. Recall that $\phi(t)$ solves the dilation equation with coefficients from $2H$:

$$\phi(t) = \frac{1}{2}\left[-\phi(2t) + 3\phi(2t-1) + 3\phi(2t-2) - \phi(2t-3)\right].$$

Set $t = 0$, 1, and 2. Then the dilation equation becomes an eigenvalue problem for $\phi(0)$, $\phi(1)$, $\phi(2)$. *This eigenvalue problem is exactly* $x = M_1 x$.

The eigenvectors of M_1 for $\lambda = -\frac{1}{2}$ and 2 should be the differences of the eigenvectors of M_0 for $\lambda = -1$ and 4. The component $x_{\text{new}}(k)$ is the difference $x_{\text{old}}(k) - x_{\text{old}}(k-1)$ for all k (extend x by zeros). This step $x_{\text{old}} \to x_{\text{new}}$ agrees in the z-domain with $\tilde{X}(z) = (1 - z^{-1})\, X(z)$:

$$\begin{bmatrix} 1 \\ .2 \end{bmatrix} \to \begin{bmatrix} 1 \\ -.8 \\ -.2 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

The left eigenvectors of M_1 are minus the sums of the left eigenvectors of M_0, plus a "constant of summation"—which is a multiple of $[1 \;\; 1 \;\; 1]$:

$$y = [1 \;\; 0] \quad \to \quad \tilde{y} = -[0 \;\; 1 \;\; 1] + [1 \;\; 1 \;\; 1] = [1 \;\; 0 \;\; 0].$$

$$y = [-.2 \;\; 1] \quad \to \quad \tilde{y} = -[0 \;\; -.2 \;\; .8] + [.3 \;\; .3 \;\; .3] = [.3 \;\; .5 \;\; -.5].$$

Those constants of integration, $C = 1$ and $C = .3$, assure orthogonality to the eigenvector $(0, .5, .5)$. Thus the eigenvectors of M_1 illustrate the pattern established by Theorem 2.

The next matrix M_2 has its filter H_2 proportional to $(-1, 2, 6, 2, -1)$. The largest eigenvalue drops from 2 to 1. But a new $\lambda = 1$ enters as always, so 1 is repeated! In this exceptional case, the matrix M_2 is not diagonalizable. There is no eigenvector with $\sum x(k) = 1$.

The consequences for the scaling function $\phi(t)$ were examined by Ragozin, Bruce, and Gao [15]. *This function is infinite at all dyadic points* $t = m/2^n$. The eigenvector x that normally gives the values of $\phi(t)$ at the integers is missing for this matrix. In the (failed) diagonalization $M_2 = S\Lambda S^{-1}$, we indicate missing eigenvectors by z's and compress the diagonal matrix Λ to a column:

$$\begin{bmatrix} z & 0 & 1 & 0 \\ z & .5 & -1.8 & 1 \\ z & 0 & .6 & -2 \\ z & -.5 & .2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ .5 \\ -.25 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ z & z & z & z \end{bmatrix} \begin{matrix} y_1 \\ y_2 \\ y_3 \\ \end{matrix}$$

The left eigenvectors $[1 \quad 1 \quad 1 \quad 1]$ for $\lambda = 1$ and $[2 \quad 1 \quad 0 \quad -1]$ for $\lambda = \frac{1}{2}$ are constant and linear. This conforms to the rule that the left eigenvector of M for $\left(\frac{1}{2}\right)^k$ is a discrete polynomial of degree $k - 1$, for each $k < p$.

We have described elsewhere [17] the consequences in continuous time. The polynomials $1, t, \ldots, t^{p-1}$ are linear combinations of the scaling functions $\phi(t - k)$. The combinations are given by the components of y! These left eigenvectors are *not* extended by zeros, but rather by the requirement that they are discrete polynomials. It is the presence of $1, t, \ldots, t^{p-1}$ that assures approximation of order p in V_0, and vanishing moments for the biorthogonal wavelets $\tilde{w}(t)$.

The repeated eigenvalue of M_2 is important. Two levels higher, with $p = 4$ zeros at $z = -1$, the matrix M_4 has a repeated eigenvalue $\lambda = \frac{1}{4}$ (and eigenvector still missing). M_4 will reappear as the matrix T for the Daubechies filter D_4 that has $p = 2$. Then the familiar Daubechies scaling function $\phi(t)$ *almost* has one derivative in L^2. Since $\lambda_{\max}$ was 4 for M_0, coming from Q, the upper bound on smoothness of $\phi(t)$ and its wavelets is $s_{\max} = p - \log_4 |\lambda_{\max}| = 2 - 1$.

3 Convergence of the Cascade Algorithm in L^2

This section turns from a pointwise analysis of the dilation equation and its solution $\phi(t)$ to an analysis based on inner products. The pointwise analysis involved $M = (\downarrow 2) 2H$. The inner products $a(k) = \int \phi(t)\phi(t+k)\,dt$ will lead to a different matrix $T = (\downarrow 2) 2HH^T$. This L^2 theory is simpler, because it involves only powers of T. We will establish the necessary and sufficient "Condition E" for convergence of the cascade algorithm (Theorem 4).

The pointwise theory involves M for the integer values $\phi(n)$ and a *shift* of M for the half-integer values. At other dyadic points, $\phi(t)$ comes from a product of those two matrices. The order of matrices for $\phi\left(\frac{19}{32}\right)$ is the order of 0's and 1's in the binary expansion .10011 of $t = \frac{19}{32}$. Since eigenvalues of products are notoriously more difficult than eigenvalues of powers, the conditions of Daubechies-Lagarias [6] and Heil-Colella [4] and others can be delicate to test [13]. The L^2 theory is based on the eigenvalues of *one* matrix T—and those eigenvalues follows the pattern established in Theorems 1 and 2.

In place of the filter H and its transfer function $H(z)$ we have HH^T and the "squared" function $H(z)H(z^{-1})$. In the frequency domain this is $\left|\sum h(k)e^{-jk\omega}\right|^2$. The zero at π now has order $2p$. Therefore T has the special eigenvalues $1, \frac{1}{2}, \ldots, \left(\frac{1}{2}\right)^{2p-1}$, which encourage convergence of the cascade algorithm and smoothness of the limit function $\phi(t)$.

The matrix T also has eigenvalues corresponding to $Q(z)Q(z^{-1})$, which has no zeros at π. After $2p$ steps of Theorem 1, those eigenvalues are divided by $2^{2p} = 4^p$. When they are less than 1 (Condition E), the cascade algorithm converges in L^2. Theorem 3 will describe the admissible starting functions $\phi^{(0)}(t)$.

Each iteration filters the current function $\phi^{(i)}(t)$ and rescales by 2, compressing time and dilating the function to maintain $\int \phi^{(i+1)}(t)dt = 1$:

$$\textit{Cascade algorithm}: \quad \phi^{(i+1)}(t) = \sum_0^N 2h(k)\phi^{(i)}(2t-k). \tag{12}$$

Normally $\phi^{(0)}(t)$ is the box function $\chi_{[0,1]}$. If convergence holds, the limit $\phi(t)$ solves the dilation equation and its graph is usually drawn by means of this iteration.

Note that the shifted box $\phi^{(0)}(t-1) = \chi_{[1,2]}$ has the same convergence properties. The shift becomes 2^{-i} after i iterations, so convergence is the same from both starting boxes. A smoother start, when $\phi^{(0)}(t)$ is the hat function on $[0,2]$, will give smoother convergence. But not all initial functions are admissible. A referee observed that the hat function on $[0,1]$, piecewise linear with $\phi^{(0)}(\frac{1}{2}) = 2$, becomes a *double hat* after one Haar iteration:

$$\phi^{(0)}(2t) + \phi^{(0)}(2t-1) \ \text{ is linear between } \ 0,2,0,2,0 \ \text{ at } \ t = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1.$$

Each cascade step scales t by 2 and doubles the number of hats. There is only weak convergence to the limit function which is $\phi(t) = \chi_{[0,1]}$. There is no L^2 convergence from this narrow hat.

A remarkable fact is that this example applies to all starting functions and all filters with a zero at π, when we periodize the cascade algorithm. The periodic functions are compressed and doubled, $P^{(i+1)}(t) = P^{(i)}(2t)$, exactly as the hat function was. This identity is Theorem 3:

$$P^{(i+1)}(t) \equiv \sum_n \phi^{(i+1)}(t-n) = \sum_n \sum_k 2h(k)\phi^{(i)}(2t-2n-k)$$

$$= \sum_{\text{even } k} 2h(k) \sum_m \phi^{(i)}(2t-2m) + \sum_{\text{odd } k} 2h(k) \sum_m \phi^{(i)}(2t-2m-1)$$

$$= \sum_m \phi^{(i)}(2t-2m) + \sum_m \phi^{(i)}(2t-2m-1) = P^{(i)}(2t).$$

After i steps the periodized function is $P^{(i)}(t) = P^{(0)}(2^i t)$. This oscilates faster and faster with no convergence, unless $P^{(0)}(t)$ is constant.

One acceptable $\phi^{(0)}(t)$ gives an exact start at the integers, using the eigenvector $x = Mx$ for the values $\phi^{(0)}(n)$. Then $\phi^{(i)}(t)$ is exact at each dyadic point $t = n/2^i$. This is the "recursive algorithm" with piecewise constant $\phi^{(i)}(t)$. In all cases, if $\phi^{(0)}(t)$ is zero outside the interval $[0, N]$, the inner products

$$a^{(0)}(k) = \int_{-\infty}^{\infty} \phi^{(0)}(t)\phi^{(0)}(t+k)\,dt$$

are identically zero for $|k| \geq N$. We always normalize by $\int \phi(t)dt = 1$ and $\int \phi^{(0)}(t)dt = 1$.

Our analysis is based on monitoring the inner products $a^{(i)}(k)$ as $i \to \infty$. The component $a^{(i)}(0)$ is the energy $\left\|\phi^{(i)}(t)\right\|^2$, and the step to $a^{(i+1)}(0)$ has been analyzed by Eirola [8], Villemoes [18], Cohen-Daubechies [2], and Hervé [9]. Lemmas 1 and 2 show that the matrix T controls the evolution of all inner products. Then Theorem 4 will establish "Condition E" for convergence.

Lemma 1 *The vectors $a^{(i)}$ of inner products $a^{(i)}(k) = \int_{-\infty}^{\infty} \phi^{(i)}(t)\phi^{(i)}(t+k)\,dt$ satisfy*

$$a^{(i+1)} = Ta^{(i)} = (\downarrow 2)2HH^Ta^{(i)}. \tag{13}$$

Proof It is very convenient to compute all inner products at once, by working with vectors:

$$a^{(i)} = \int_{-\infty}^{\infty} \phi^{(i)}(t)\Phi^{(i)}(t)\,dt \quad \text{with} \quad \Phi^{(i)}(t) = \begin{bmatrix} \cdot \\ \phi^{(i)}(t-1) \\ \phi^{(i)}(t) \\ \phi^{(i)}(t+1) \\ \cdot \end{bmatrix}.$$

The next vector in the cascade is $\Phi^{(i+1)}(t) = (\downarrow 2)2H\Phi^{(i)}(2t)$. The inner products become

$$\begin{aligned} a^{(i+1)} &= \int_{-\infty}^{\infty} \phi^{(i+1)}(t)\Phi^{(i+1)}(t)\,dt \\ &= \int_{-\infty}^{\infty} \left[2\sum h(k)\phi^{(i)}(2t-k)\right]\left[(\downarrow 2)2H\Phi^{(i)}(2t)\right]dt. \end{aligned} \tag{14}$$

Bring the operator $(\downarrow 2)2H$ outside the integral. Change variables in the kth term to $u = 2t - k$. That term becomes (with $du = 2dt$)

$$\int_{-\infty}^{\infty} h(k)\phi^{(i)}(u)\Phi^{(i)}(u+k)\,du = h(k)S^{-k}a^{(i)}. \tag{15}$$

The k-step shift S^{-k} allowed us to write $\Phi^{(i)}(u+k)$ as $S^{-k}\Phi^{(i)}(u)$. Then the integration with respect to u produced $a^{(i)}$. Now sum equation (15) on k to reach the matrix $\sum h(k)S^{-k}$, which is H^T as Lemma 1 requires:

$$a^{(i+1)} = (\downarrow 2)2HH^Ta^{(i)} = Ta^{(i)}. \tag{16}$$

Lemma 2 *The vectors $b^{(i)}$ of inner products $b^{(i)}(k) = \int_{-\infty}^{\infty} \phi^{(i)}(t)\,\phi(t+k)\,dt$ satisfy $b^{(i+1)} = Tb^{(i)}$.*

Proof The steps are the same, with no superscript on $\Phi(t)$:

$$b^{(i+1)} = \int_{-\infty}^{\infty} \phi^{(i+1)}(t)\Phi(t)\,dt \quad \text{with} \quad \Phi(t) = \begin{bmatrix} \cdot \\ \phi(t-1) \\ \phi(t) \\ \phi(t+1) \\ \cdot \end{bmatrix}. \tag{17}$$

The cascade formula is substituted for $\phi^{(i+1)}(t)$, as before. This time we use the dilation equation for $\Phi(t)$. It is useful to see this equation in vector form:

$$\phi(t) = \sum 2h(k)\,\phi(2t-k) \quad \text{is exactly} \quad \Phi(t) = M\Phi(2t) = (\downarrow 2)2H\Phi(2t). \tag{18}$$

After these substitutions (17) matches (14), with $\Phi^{(i)}$ replaced by Φ. Change variables in the kth term to $u = 2t - k$, and that term matches (15)—with a replaced by b. Then sum on k to find $b^{(i+1)} = Tb^{(i)}$.

Note 1 The same steps give the inner products of translates of *any functions that satisfy two-scale equations*. The inner products are computed from the coefficients in the equations, and not from the functions themselves [16]. The inner products of $\phi(t)$ and $w(t)$ with the wavelets $w(t+k)$ are given by $(\downarrow 2)2HH_1^T a$ and $(\downarrow 2)2H_1H_1^T a$. Here H_1 is the highpass filter. We have simple formulas for inner products but not for $\phi(t)$ and $w(t)$. A very useful C++ code has been created by Kunoth to compute integrals of products of several scaling functions (and derivatives). It is accessed by `ftp.igpm.rwth-aachen.de`.

An extreme case is the two-scale relation $\delta(t) = 2\delta(2t)$ for the delta function. The inner products $\int \delta(t)\phi(t+k)\,dt$ are the values $\Phi(0) = (\ldots, \phi(0), \phi(1), \phi(2), \ldots)$. The transition m Matrix T becomes $(\downarrow 2)2HI$; this is M. The eigenvalue problem $\Phi(0) = M\Phi(0)$ gives $\phi(n)$ at the integers.

Note 2 The computation of $b^{(i)}$ assumed the existence of $\phi(t)$ in L^2. We could prove this existence, by demonstrating that $\left\|\phi^{(m)}(t) - \phi^{(n)}(t)\right\|$ is a Cauchy sequence [17]. The requirement is the same Condition E, and the completeness of L^2 guarantees a limit $\phi(t)$. For simplicity we omit this step, and prove Theorem 4—that $\left\|\phi(t) - \phi^{(i)}(t)\right\|$ converges to zero.

Proof of convergence (Theorem 4) At the ith step of the cascade algorithm, the squared distance from $\phi^{(i)}(t)$ to $\phi(t)$ is

$$\left\|\phi^{(i)} - \phi\right\|^2 = \left\langle \phi^{(i)}, \phi^{(i)} \right\rangle - 2\left\langle \phi^{(i)}, \phi \right\rangle + \langle \phi, \phi \rangle. \tag{19}$$

Those are the zeroth components $a^{(i)}(0)$ and $-2b^{(i)}(0)$ and $a(0)$. By Lemmas 1–2, they come from multiplication i times by T. Now apply Condition E, which is exactly the requirement for convergence of the "power method":

$$a^{(i)} = T^i a^{(0)} \quad \text{and} \quad b^{(i)} = T^i b^{(0)} \quad \text{both converge to } a.$$

This immediately gives the convergence of the cascade algorithm, from (19):

$$\left\|\phi^{(i)} - \phi\right\|^2 = a^{(i)} - 2b^{(i)}(0) + a(0) \to a(0) - 2a(0) + a(0) = 0. \tag{20}$$

Although we only needed the zeroth components, the argument was made simple by working with the vectors $a^{(i)}$ and the matrix T.

Recall how Condition E enters in $T^i a^{(0)} \to a$. When the initial vector $a^{(0)}$ is a combination $a + c_2 x_2 + c_3 x_3 + \cdots$ of eigenvectors of T, multiplication by T^i will introduce factors λ^i. For $|\lambda| < 1$ those factors approach zero. The limit is the eigenvector a with $\lambda = 1$. In case T has a shortage of eigenvectors, the argument is based on its Jordan form and the conclusion $T^i a^{(0)} \to a$ is still correct. Similarly we have $T^i b^{(0)} \to a$, because all these vectors have the same normalization:

$$\begin{aligned} \textstyle\sum a(n) &= \textstyle\int \phi(t) \sum \phi(t-n)dt &&= 1 \\ \textstyle\sum a^{(i)}(n) &= \textstyle\int \phi^{(i)}(t) \sum \phi^{(i)}(t-n)dt &&= 1 \\ \textstyle\sum b^{(i)}(n) &= \textstyle\int \phi(t) \sum \phi^{(i)}(t-n)dt &&= 1. \end{aligned} \tag{21}$$

The periodized functions are identically 1 and $\int \phi(t)dt = \int \phi^{(i)}(t) = 1$. Note that the requirement $\sum(-1)^k h(k) = 0$ in Theorem 3, which gives one zero at least ($p \geq 1$) at the point $z = -1$, is necessary for convergence (but not sufficient).

Finally we suppose that the cascade algorithm is convergent in L^2. The limit function $\phi(t)$ is the unique solution [6] to the dilation equation, normalized by $\int \phi(t)dt = 1$. Therefore the inner product vectors $a^{(i)}$ for $\phi^{(i)}(t)$ converge to the inner product vector a for $\phi(t)$, and $Ta = a$. We now prove that Condition E must hold for the matrix T.

Suppose that $Tv = \lambda v$ for some nonzero v with $\lambda \neq 1$. Then v is perpendicular to the row vector $e = [1 \;\; 1 \;\; \ldots \;\; 1]$ of all ones. (Reason: We have $eT = e$ and thus $eTv = ev$. But also $eTv = \lambda ev$. Therefore $ev = 0$.) We can choose an initial function $\phi^{(0)}(t)$ in the acceptable set I such that $a^{(0)} = a + cv$. Then $T^i a^{(0)} = a + c\lambda^i v^i$. The condition $|\lambda| < 1$ is necessary for convergence to a.

The argument is the same if a second eigenvector $v = Tv$ comes from a repeated $\lambda = 1$. We can choose v so that $ev = 0$, and then choose $\phi^{(0)}(t)$ in I so that $a^{(0)} = a + cv$. Convergence to a

would fail because $T^i a^{(0)} = a + cv$. Therefore this v cannot exist.

Note that every acceptable $\phi^{(0)}(t)$ (in I) gives $ea^{(0)} = 1$ by (21). The set of piecewise constant functions $\phi^{(0)}(t)$ in I can produce any inner product vector $a^{(0)}$ that has $ea^{(0)} = 1$ and $A(\omega) = \sum a^{(0)}(k)e^{ik\omega} \geq 0$ for all ω. Spectral factorization of $A(\omega)$ gives the constants in $\phi^{(0)}(t)$. This provides a ball around a on the hyperplane $ea^{(0)} = 1$. Then if $ev = 0$, we *can* choose $\phi^{(0)}(t)$ to achieve $a^{(0)} = a + cv$.

To complete the proof that condition E is necessary, we must show that convergence fails when $\lambda = 1$ is a repeated eigenvalue of T with only one eigenvector. This is exactly the case illustrated by the matrix M_2 given earlier. (M_2 is the matrix T for the "square root" of the filter $\frac{1}{8}(-1, 2, 6, 2, -1)$. That square root has $H(z)H(z^{-1}) = \frac{1}{8}(-z^2+2z+6+2z^{-1}-z^{-2})$. We are proving that the cascade algorithm cannot converge for this $H(z)$.) If convergence did hold, the limit $\phi(t)$ would have an inner product vector with $Ta = a$ and $ea = 1$. But the eigenvector for a defective eigenvalue $\lambda = 1$ is perpendicular to the left eigenvector e. In the M_2 example this right eigenvector is $(0, .5, 0, -.5)$. It is not an acceptable a, and the cascade algorithm could not converge

Note 3 In the orthonormal case, there is no danger that T has an eigenvalue with $|\lambda| > 1$. The norm of T is $\sup\left(\left|H\left(e^{j\omega}\right)\right|^2 + \left|H\left(-e^{j\omega}\right)\right|^2\right) = 1$. Condition E reduces to the Cohen-Lawton condition [2], [11] that $\lambda = 1$ is a simple eigenvalue of T. Then $\{\phi(t+k)\}$ is an orthonormal basis.

Note 4 If all eigenvalues of T with $|\lambda| = 1$ are nondefective, the powers T^i remain bounded. Then the vectors $a^{(i)}$ are bounded. S. L. Lee observed that there is at least *weak* convergence to an L^2 solution of the dilation equation. (See [12] also for multidimensional application of Condition E.) The familiar example $h = (\frac{1}{2}, 0, 0, \frac{1}{2})$ gives weak convergence to the stretched box $\phi(t) = \frac{1}{3}\chi_{[0,3]}$. In this case T has eigenvalues $1, 1, -1, \ldots$ and Condition E is violated. The cascade algorithm cannot converge strongly in L^2. It is conjectured that this weak form of Condition E, allowing nondefective $|\lambda| = 1$, is also necessary for the existence of $\phi(t)$ in L^2.

References

[1] A. S. Cavaretta, W. Dahmen, and C. A. Micchelli, Stationary Subdivision, *Amer. Math. Soc. Memoirs*, vol. 453, Providence RI (1991).

[2] A. Cohen and I. Daubechies, A stability criterion for biorthogonal wavelet bases and their related subband coding scheme, *Duke Math. J.* **68** (1992) 313–335.

[3] A. Cohen, I. Daubechies, and J. C. Feauveau, Biorthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **45** (1992) 485–560.

[4] C. Heil and D. Colella, *Dilation equations and the smoothness of compactly supported wavelets*, Wavelets: Mathematics and Applications, J. Benedetto and M. Frazier, eds., CRC Press (1993).

[5] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.

[6] I. Daubechies and J. Lagarias, Two-scale difference equations: I. Existence and global regularity of solutions, *SIAM J. Math. Anal.*, **22** (1991), 1388–1410; Two-scale difference equations: II. Local regularity, infinite products of matrices and fractals, *SIAM J. Math. Anal.*, **23** (1992), 1031–1079.

[7] S. Durand, Convergence of cascade algorithms introduced by I. Daubechies, Numer. Alg. **4** (1993) 307-322.

[8] T. Eirola, Sobolev characterization of solutions of dilation equations, *SIAM J. Math. Anal.* **23** (1992) 1015–1030.

[9] L. Hervé, Construction et regularité des fonctions d'échelle, preprint (1992).

[10] R.–Q. Jia, Subdivision schemes in L^p spaces, *Advances in Computational Math.* **3** (1995) 309-341.

[11] W. Lawton, Necessary and sufficient conditions for constructing orthogonal wavelets, *J. Math. Phys.* **32** (1991) 52–61.

[12] W. Lawton, S. L. Lee, and Z. Shen, Convergence of multidimensional cascade algorithm, preprint.

[13] M.Maesumi, Optimum unit ball for joint spectral radius, in Approximation Theory VIII, C. K. Chui and L. Schumaker, eds., World Scientific (1995).

[14] Y. Meyer and F. Paiva, Remarques sur la construction des ondelettes orthogonales, J. d'Analyse Math. **60** (1993) 227-240.

[15] D. Ragozin, A. Bruce, and H. Gao, Non-smooth wavelets unbounded on every interval, *Proc. Maratea NATO Conference*, Kluwer (1995).

[16] G. Strang, Inner products and condition numbers for wavelets and filter banks, preprint (1994).

[17] G. Strang and T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley MA (1995).

[18] L. F. Villemoes, Energy moments in time and frequency for two-scale difference equation solutions and wavelets, *SIAM J. Math. Anal.* **23** (1992) 1519–1543.

[19] L. F. Villemoes, Wavelet analysis of refinement equations, *SIAM J. Math. Anal.* **25** (1994) 1433–1460.