NUMERICAL STABILITY OF BIORTHOGONAL WAVELET TRANSFORMS

FRITZ KEINERT
DEPARTMENT OF MATHEMATICS
IOWA STATE UNIVERSITY
AMES, IA 50011
KEINERT@IASTATE.EDU

May 8, 1995

ABSTRACT. For orthogonal wavelets, the discrete wavelet and wave packet transforms and their inverses are orthogonal operators with perfect numerical stability. For biorthogonal wavelets, numerical instabilities can occur. We derive bounds for the 2-norm and average 2-norm of these transforms, including efficient numerical estimates if the number L of decomposition levels is small, as well as growth estimates for $L \to \infty$. These estimates allow easy determination of numerical stability directly from the wavelet coefficients. Examples show that many biorthogonal wavelets are in fact numerically well behaved.

1. INTRODUCTION

The discrete wavelet transform and wave packet transform have become well established in many applications, such as signal processing. Originally derived for orthogonal wavelets, they can equally well be based on biorthogonal wavelets.

Numerical experiments with various types of biorthogonal wavelets show that in some cases considerable roundoff error is accumulated during decomposition and reconstruction. Strangely enough, these wavelets still perform well in certain applications such as signal compression (see [8]).

Other types of biorthogonal wavelets have excellent stability behavior. That is, the roundoff error remains close to machine accuracy even after extensive calculations.

We denote by W_L, P_L the wavelet and wave packet transforms through L levels of decomposition, respectively. We estimate error growth by the matrix 2-norm $\|.\|_2$ (worst case) and the average 2-norm $\|.\|_{2,\text{avg}}$ defined in (2.14) (average case).

The main goal of this paper is to calculate explicit bounds on these norms for a given set of wavelet coefficients $\{h_k\}$, $\{g_k\}$, where no conditions other than the perfect reconstruction property are imposed. Omitting the technical details and rigor found in sections 4, 5, the main result states the following:

Given wavelet coefficients $\{h_k\}$, $\{g_k\}$, we can form two matrices A and B (given in equations (5.1), (5.2)) and calculate their largest eigenvalues, denoted by ρ_W and ρ_P, respectively.

1991 *Mathematics Subject Classification.* primary 15A12, 65F35, secondary 42C15.
Key words and phrases. wavelets, biorthogonal wavelets, stability.

Then, for $L \to \infty$,

$$\|W_L\|_2 \le c \cdot \max(\sqrt{\rho_W}, 1)^L$$
$$\|P_L\|_2 \le c \cdot (\sqrt{\rho_P})^L$$
$$\|W_L\|_{2,\text{avg}} \le c \cdot \max(\sqrt{\rho_W/2}, 1)^L$$
$$\|P_L\|_{2,\text{avg}} \le c \cdot \left(\sqrt{\rho_P/2}\right)^L$$

The first two inequalities may be strict (i.e. the actual growth factor could be less than $\sqrt{\rho}$).

Under the mild additional restrictions $\sum h_{2k} = \sum h_{2k+1} = \sum(-1)^k g_{2k} = \sum(-1)^{k+1} g_{2k+1} = \sqrt{2}/2$ (which are usually satisfied), $\rho_W \ge 1$ and $\rho_P > 2$. In this case, both standard and average norms for the wave packet decomposition grow without bound. The wavelet decomposition is stable in L^2 (i.e. $\|W_L\|_2$ remains bounded as $L \to \infty$) exactly if $\rho_W = 1$, which is the case for many wavelets of practical interest. The average case norm is even more likely to remain bounded. In the examples we considered, the norms that remained bounded became constant rather quickly, and at very small levels (see table 6).

Wavelets form a Riesz basis if and only if the norms of W_L and W_L^{-1} remain bounded as $L \to \infty$. Conditions under which wavelets and wave packets form a Riesz basis have been investigated previously in [1], [2]. Our results for the 2-norm, with the restrictions on wavelet coefficients added, are already contained in these papers.

A secondary goal of this paper is to explain why even wavelets with severe stability problems perform well in certain applications. We to do this in section 6, where we interpret the norm estimates for three settings.

The outline of the paper is as follows:

In section 2 we introduce the necessary background from wavelet theory and define the norms we plan to calculate.

In section 3, we introduce methods for calculating the norms for small L. These techniques are more efficient than forming the matrices themselves, and also address the dependence on vector length N (or rather, independence of N).

Section 4 contains the proofs of two technical lemmas needed in section 5. They are presented separately so as not to interrupt the flow of ideas later.

Section 5 (summarized in this introduction) gives upper bounds on the norms as $L \to \infty$.

In section 6 we give some illustrations on how to apply these estimates in practice.

2. Background

The fundamentals of wavelet theory have been described in many publications, so we refer to the literature for most of the background, and only introduce the specific notation and results we need. A standard reference for wavelets is [5]. Biorthogonal wavelets are discussed in [3].

Biorthogonal wavelets are given by two dual sets of coefficients $\{h_k\}$, $\{g_k\}$ and $\{\tilde{h}_k\}$, $\{\tilde{g}_k\}$. For orthogonal wavelets, both sets are identical. For simplicity, we assume these sequences have finite length and are real.

Let

$$\mathbf{s}^0 = \{s_0^0, s_1^0, \ldots, s_{N-1}^0\}^T$$

be a vector of length N, where N is divisible by 2^L for some $L \geq 1$. The *discrete wavelet transform* or *wavelet decomposition* of $\mathbf{s}^0$ through L levels is the vector $(\mathbf{s}^L, \mathbf{d}^L, \mathbf{d}^{L-1}, \ldots, \mathbf{d}^1)^T$, where each $\mathbf{s}^n$, $\mathbf{d}^n$ is a subvector of length $N/2^n$ which is calculated recursively by

$$\begin{aligned} s_k^{n+1} &= \sum_j h_{j-2k} s_j^n, \\ d_k^{n+1} &= \sum_j g_{j-2k} s_j^n. \end{aligned} \tag{2.1}$$

Subscripts are interpreted modulo $N/2^n$, the length of $\mathbf{s}^n$.

The *inverse discrete wavelet transform* or *wavelet reconstruction* is based on the formula

$$s_j^n = \sum_k \left\{ \tilde{h}_{j-2k} s_k^{n+1} + \tilde{g}_{j-2k} d_k^{n+1} \right\}, \tag{2.2}$$

where subscripts are interpreted modulo $N/2^{n+1}$.

Throughout this paper, L denotes the number of levels of decomposition, and N denotes the length of the original vector $\mathbf{s}^0$.

The roles of $\{h_k\}$, $\{g_k\}$, and $\{\tilde{h}_k\}$, $\{\tilde{g}_k\}$ in (2.1) and (2.2) can be reversed, leading to a dual transform.

The *discrete wave packet transform* is defined in the same manner as the wavelet transform, except that the recursive decomposition (2.1) is applied to all intermediate vectors at each level, instead of just the $\mathbf{s}^n$. A full description would require a more complicated notation; we refer to [4] for details.

In terms of the auxiliary functions

$$\begin{aligned} h(\theta) &= \frac{1}{\sqrt{2}} \sum_j h_j e^{ij\theta}, \\ g(\theta) &= \frac{1}{\sqrt{2}} \sum_j g_j e^{ij\theta}, \end{aligned} \tag{2.3}$$

(and likewise for $\tilde{h}$, $\tilde{g}$), perfect reconstruction is equivalent to

$$\begin{pmatrix} h(\theta) & h(\theta+\pi) \\ g(\theta) & g(\theta+\pi) \end{pmatrix} \begin{pmatrix} \overline{\tilde{h}(\theta)} & \overline{\tilde{g}(\theta)} \\ \overline{\tilde{h}(\theta+\pi)} & \overline{\tilde{g}(\theta+\pi)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.4}$$

where the bar denotes complex conjugation. We impose no further conditions until the end of section 5, which means there may not be any actual scaling or wavelet functions associated with these coefficients. The decomposition and reconstruction operations work for purely algebraic reasons.

For later reference we note that the determinant of the first matrix in (2.4) satisfies

$$h(\theta)g(\theta+\pi) - h(\theta+\pi)g(\theta) = \Delta \cdot e^{(2k+1)i\theta} \tag{2.5}$$

for some $\Delta \in \mathbb{C}$, $\Delta \neq 0$, $k \in \mathbb{Z}$. The scaling is usually chosen so that $\Delta = 1$.

The discrete Fourier transform $\hat{\mathbf{s}}^n$ of $\mathbf{s}^n$ is given by

$$\hat{s}_k^n = \frac{1}{\sqrt{N/2^n}} \sum_j s_j^n e^{-2^n jk \cdot \frac{2\pi i}{N}}.$$

$\hat{s}_k^n$ is the value at $2^n k \cdot \frac{2\pi}{N}$ of the 2π-periodic function

$$\hat{s}^n(\theta) = \frac{1}{\sqrt{N/2^n}} \sum_j s_j^n e^{-ij\theta}.$$

We use the 2-norms

$$\begin{aligned} \|\mathbf{x}\|_2 &= \sqrt{\sum_j |x_j|^2}, \\ \|f\|_2 &= \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|^2 \, d\theta} \end{aligned} \tag{2.6}$$

to measure the size of vectors and 2π-periodic functions, respectively. Our normalizations imply

$$\begin{aligned} \|\hat{\mathbf{s}}^n\|_2 &= \|\mathbf{s}^n\|_2, \\ \|\hat{s}^n\|_2 &= \frac{1}{\sqrt{N/2^n}} \|\mathbf{s}^n\|_2. \end{aligned} \tag{2.7}$$

In terms of $\hat{s}$, $\hat{d}$, formulas (2.1) and (2.2) read (see Daubechies [5], section 5.6)

$$\begin{aligned} \hat{s}^{n+1}(\theta) &= h(\frac{\theta}{2})\hat{s}^n(\frac{\theta}{2}) + h(\frac{\theta}{2}+\pi)\hat{s}^n(\frac{\theta}{2}+\pi), \\ \hat{d}^{n+1}(\theta) &= g(\frac{\theta}{2})\hat{s}^n(\frac{\theta}{2}) + g(\frac{\theta}{2}+\pi)\hat{s}^n(\frac{\theta}{2}+\pi), \end{aligned} \tag{2.8}$$

$$\hat{s}^n(\theta) = 2\left[\tilde{h}(\theta)\hat{s}^{n+1}(2\theta) + \tilde{g}(\theta)\hat{d}^{n+1}(2\theta)\right]. \tag{2.9}$$

For a fixed set of wavelet coefficients $\{h_k\}$, $\{g_k\}$, let W_L be the matrix corresponding to a wavelet transform through L levels, i.e.

$$\begin{pmatrix} \mathbf{s}^L \\ \mathbf{d}^L \\ \cdots \\ \mathbf{d}^1 \end{pmatrix} = W_L \mathbf{s}^0. \tag{2.10}$$

The dependence of W_L on the vector length N is not explicitly shown, to keep the notation simple. $\tilde{W}_L$ denotes the matrix corresponding to the dual transform using $\{\tilde{h}_k\}$, $\{\tilde{g}_k\}$, while the matrices for the wave packet transform are called P_L and $\tilde{P}_L$.

The inverse transforms correspond to the matrices

$$\begin{aligned} W_L^{-1} &= \tilde{W}_L^T, & \tilde{W}_L^{-1} &= W_L^T, \\ P_L^{-1} &= \tilde{P}_L^T, & \tilde{P}_L^{-1} &= P_L^T. \end{aligned} \tag{2.11}$$

where the superscript T denotes transposition.

Applying discrete Fourier transforms in (2.10) (on each $\mathbf{s}^n$, $\mathbf{d}^n$ separately) leads to

$$\begin{pmatrix} \hat{\mathbf{s}}^L \\ \hat{\mathbf{d}}^L \\ \cdots \\ \hat{\mathbf{d}}^1 \end{pmatrix} = \hat{W}_L \hat{\mathbf{s}}^0. \tag{2.12}$$

If F_N stands for a discrete Fourier transform of length N, W_L and $\hat{W}_L$ are related by

$$\hat{W}_L = \begin{pmatrix} F_{N/2^L} & & & \\ & F_{N/2^L} & & \\ & & \ddots & \\ & & & F_{N/2} \end{pmatrix} W_L F_N^{-1}.$$

Assume now that some error vector $\mathbf{e}$ is added to $\mathbf{s}^0$. The error in $W_L\mathbf{s}^0$ is bounded by

$$\|W_L(\mathbf{s}^0 + \mathbf{e}) - W_L\mathbf{s}^0\|_2 = \|W_L\mathbf{e}\|_2 \leq \|W_L\|_2\|\mathbf{e}\|_2, \tag{2.13}$$

where the matrix 2-norm is defined by

$$\|A\|_2 = \max_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2.$$

This bound is sharp, i.e. it is achieved for some error $\mathbf{e}$.

In addition to this worst case estimate, we are also interested in the "average" case of error growth. For this we assume that

$$\mathbf{e} = \|\mathbf{e}\|_2 \cdot X,$$

where X is a random variable uniformly distributed on the unit sphere S^{N-1} in $\mathbb{R}^N$. The expected value of the error in $W_L\mathbf{s}^0$ is

$$E\left[\|W_L\mathbf{e}\|_2\right] = \|W_L\|_{2,\mathrm{avg}} \cdot \|\mathbf{e}\|_2,$$

where the *average 2-norm* of a matrix A is defined by

$$\|A\|_{2,\mathrm{avg}} = E\left[\|AX\|_2\right]. \tag{2.14}$$

For a given $N \times N$ matrix A let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N \geq 0$ be the singular values of A, i.e. the square roots of the eigenvalues of A^TA. Then

$$\begin{aligned} \|A\|_2 &= \sigma_1, \\ \|A\|_{2,\mathrm{avg}} &= \int_{S^{N-1}} \sqrt{\sum_k x_k^2\sigma_k^2}\, dS^{N-1}, \end{aligned} \tag{2.15}$$

where dS^{N-1} is the normalized surface measure on S^{N-1}.

The upper bound

$$\|A\|_{2,\mathrm{avg}} \leq \sqrt{\frac{\sum_k \sigma_k^2}{N}} \tag{2.16}$$

is derived in [6]. The authors of [6] furthermore provided a computer program that can calculate the exact value in (2.15) numerically.

The goal of this paper is to derive efficient methods for calculating $\|.\|_2$, $\|.\|_{2,\mathrm{avg}}$ for W_L, $\tilde{W}_L$, P_L, $\tilde{P}_L$ and their inverses for small L, and to give bounds on the growth of these norms as $L \to \infty$. We only need to do this for W_L, P_L; the calculations for $\tilde{W}_L$, $\tilde{P}_L$ are analogous, using the dual coefficients, and

$$\|W_L^{-1}\| = \|\tilde{W}_L^T\| = \|\tilde{W}_L\|$$

for both norms under consideration.

The results are illustrated on a number of different wavelets:

- The dual binomial and complementary wavelets with M vanishing moments, introduced in [8], for $M = 2, \dots, 6$.
- Several of the Cohen wavelets, introduced in [3], section 6.A.
- the Barlaud wavelet, introduced in [3], section 6.C.1.

The coefficients of these wavelets are listed in table 1.

We note that the Barlaud wavelet and all Cohen wavelets with minimal regularity are known to form Riesz bases, which implies stability.

(Table 1 goes near here)

3. Numerical Norm Estimates For Small L

For small values of L, N, the norms of W_L, P_L could be calculated by forming the matrices explicitly and calculating their norm. We present here some techniques that are more efficient, address the dependence on N, and introduce some estimates needed later.

3.1. 2-Norms. By (2.15), we need to calculate the largest singular values of W_L, P_L, or equivalently those of $\hat{W}_L$, $\hat{P}_L$. (The Fourier transform is a unitary operator which leaves singular values invariant).

We start with $\hat{W}_1$. For a vector of length N, it has the explicit form (see (2.12) and (2.8))

$$\hat{W}_1 = \begin{pmatrix} h(\theta_0) & & & & h(\theta_{N/2}) & & & \\ & h(\theta_1) & & & & h(\theta_{N/2+1}) & & \\ & & \ddots & & & & \ddots & \\ & & & h(\theta_{N/2-1}) & & & & h(\theta_{N-1}) \\ g(\theta_0) & & & & g(\theta_{N/2}) & & & \\ & g(\theta_1) & & & & g(\theta_{N/2+1}) & & \\ & & \ddots & & & & \ddots & \\ & & & g(\theta_{N/2-1}) & & & & g(\theta_{N-1}) \end{pmatrix} \tag{3.1}$$

where

$$\theta_j = j \cdot \frac{2\pi}{N}. \tag{3.2}$$

After further reordering of rows and columns in $\hat{W}_1$ (again a unitary operation), we conclude that W_1 has the same singular values as the block diagonal matrix

$$\begin{pmatrix} \hat{W}(\theta_0) & & & \\ & \hat{W}(\theta_1) & & \\ & & \ddots & \\ & & & \hat{W}(\theta_{N/2-1}) \end{pmatrix},$$

where

$$\hat{W}(\theta) = \begin{pmatrix} h(\theta) & h(\theta+\pi) \\ g(\theta) & g(\theta+\pi) \end{pmatrix}.$$

The singular values of each block $\hat{W}(\theta)$ can be explicitly calculated. They are

$$\sqrt{\frac{F(\theta) \pm \sqrt{F(\theta)^2 - 4|\Delta|^2}}{2}},$$

where $F(\theta)$ is the square of the Frobenius norm of $\hat{W}(\theta)$

$$F(\theta) = |h(\theta)|^2 + |h(\theta+\pi)|^2 + |g(\theta)|^2 + |g(\theta+\pi)|^2,$$

and Δ is defined in (2.5).

Standard calculus arguments show that the largest singular value of $\hat{W}_1$ occurs in the block $\hat{W}(\theta_j)$ where $F(\theta_j)$ achieves its maximum, and we find

$$\sigma_1 \le \sqrt{\frac{F_{\max} + \sqrt{F_{\max}^2 - 4|\Delta|^2}}{2}}, \tag{3.3}$$

where

$$F_{\max} = \max_\theta F(\theta).$$

For any reasonably large N, we expect one of the points θ_j to be at or very near the maximum, so this upper bound should be quite accurate. In any case, the bound (3.3) is valid for all N. In some simple cases, the largest singular value can be found explicitly. In general it is easy to determine numerically.

Example: For the dual B-spline wavelets with 2 vanishing moments (see table 1) we find

$$F(\theta) = \frac{5}{2} - \frac{1}{2}\cos(2\theta),$$

and $\Delta = 1$, so $F_{\max} = 3$, $\sigma_1 \approx 1.618$. The maximum is at $\theta = \pi/2$ and is achieved for any N divisible by 4. □

Reordering $\hat{W}_2$ results in a block diagonal matrix with 4×4 blocks calculated by multiplying

$$\begin{pmatrix} h(2\theta) & h(2\theta+\pi) & 0 & 0 \\ g(2\theta) & g(2\theta+\pi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h(\theta) & 0 & h(\theta+\pi) & 0 \\ 0 & h(\theta+\frac{\pi}{2}) & 0 & h(\theta+\frac{3\pi}{2}) \\ g(\theta) & 0 & g(\theta+\pi) & 0 \\ 0 & g(\theta+\frac{\pi}{2}) & 0 & g(\theta+\frac{3\pi}{2}) \end{pmatrix},$$

for $\theta = \theta_j$, $j = 1, \ldots, N/4$.

For $\hat{P}_2$, the blocks are

$$\begin{pmatrix} h(2\theta) & h(2\theta+\pi) & 0 & 0 \\ g(2\theta) & g(2\theta+\pi) & 0 & 0 \\ 0 & 0 & h(2\theta) & h(2\theta+\pi) \\ 0 & 0 & g(2\theta) & g(2\theta+\pi) \end{pmatrix} \begin{pmatrix} h(\theta) & 0 & h(\theta+\pi) & 0 \\ 0 & h(\theta+\frac{\pi}{2}) & 0 & h(\theta+\frac{3\pi}{2}) \\ g(\theta) & 0 & g(\theta+\pi) & 0 \\ 0 & g(\theta+\frac{\pi}{2}) & 0 & g(\theta+\frac{3\pi}{2}) \end{pmatrix},$$

Reordering $\hat{W}_L$, $\hat{P}_L$ similarly produces blocks of size $2^L \times 2^L$.

We know of no closed form solution for the singular values of these blocks for $L > 1$, but a simple numerical maximization can find the largest possible singular value for any N in $O(2^{3L})$ calculations.

Some sample results for the dual binomial wavelet with $M = 2$ are shown in the first two columns in table 2. Column 1 shows the actual norm for $N = 64$, calculated by forming the matrix explicitly; column 2 shows the upper bound for any N. It appears that for $L = 1 \ldots 5$, the maximum occurs at one of the θ_j, for $L = 6, 7$ the maximum falls between the θ_j.

(Table 2 is referred to in both sections 3.1 and 3.2; it could go anywhere between here and the end of section 3.2)

3.2. Average 2-Norms. Using estimate (2.16), we need to calculate $\sum \sigma_k^2/N$. Again, we work with $\hat{W}_L$, $\hat{P}_L$ rather than W_L, P_L.

The sum of all eigenvalues of a matrix is equal to its trace. The σ_k^2 are the eigenvalues of $\hat{W}_L^H \hat{W}$ (where the superscript H denotes the Hermitian transpose), so

$$\sum_k \sigma_k^2 = \sum_i |(\hat{W}_L^H \hat{W}_L)_{ii}| = \sum_{ij} |(\hat{W}_L)_{ij}|^2.$$

From (3.1), we find for $\hat{W}_1$

$$\frac{\sum_k \sigma_k^2}{N} = \frac{1}{N} \left[\sum_{j=0}^{N-1} |h(\theta_j)|^2 + \sum_{j=0}^{N-1} |g(\theta_j)|^2 \right], \tag{3.4}$$

with θ_j given in (3.2).

Assuming that $\{h_k\}$, $\{g_k\}$ both contain m coefficients, the trigonometric polynomials $|h(\theta)|^2$, $|g(\theta)|^2$ have indices ranging from $(-m+1)$ to $(m-1)$. If $N > (m-1)$, then

$$\sum_{j=0}^{N-1} e^{jk \cdot 2\pi i/N} = \begin{cases} 0 & \text{if } k \neq 0, \\ N & \text{if } k = 0. \end{cases}$$

Thus, for $N > (m-1)$ all terms on the right-hand side of (3.4) sum to 0, except the constant terms, and we get

$$\frac{\sum_k \sigma_k^2}{N} = \text{ constant term in } |h(\theta)|^2 + |g(\theta)|^2.$$

By the same argument, we find for two decomposition steps and $N > 3(m-1)$

$$\frac{\sum_k \sigma_k^2}{N} = \text{ constant term in } |h(\theta)|^2 \left[|h(2\theta)|^2 + |g(2\theta)|^2 \right] + |g(\theta)|^2,$$

and in general for $N > (2^L - 1)(m-1)$

$$\frac{\sum_k \sigma_k^2}{N} = \text{ constant term in } f_L(\theta), \tag{3.5}$$

where f_L is defined by the recursion

$$\begin{aligned} f_1(\theta) &= |h(\theta)|^2 + |g(\theta)|^2, \\ f_L(\theta) &= |h(\theta)|^2 f_{L-1}(2\theta) + |g(\theta)|^2. \end{aligned} \tag{3.6}$$

For $N \leq (2^L - 1)(M-1)$, we can still estimate the average 2-norm from the coefficients of f_L, as the sum of coefficients of all terms whose index is congruent to 0 modulo N. The trigonometric polynomials f_L can be calculated recursively very rapidly.

Similarly for $\hat{P}_2$

$$\frac{\sum_k \sigma_k^2}{N} = \text{ constant term in } \left[|h(\theta)|^2 + |g(\theta)|^2 \right] \left[|h(2\theta)|^2 + |g(2\theta)|^2 \right],$$

and in general for $N > (2^L - 1)(m-1)$

$$\frac{\sum_k \sigma_k^2}{N} = \text{ constant term in } p_L(\theta), \tag{3.7}$$

where p_L is defined by the recursion

$$\begin{aligned} p_1(\theta) &= |h(\theta)|^2 + |g(\theta)|^2, \\ p_L(\theta) &= \left[|h(\theta)|^2 + |g(\theta)|^2\right] p_{L-1}(2\theta). \end{aligned} \tag{3.8}$$

What we have really calculated here is not the average 2-norm, but the upper bound (2.16) for it. However, it is noted in [6] that in many cases the upper bound is quite close to the true value. This observation is supported by our own experiments.

Table 2 shows the results for the dual binomial wavelet with $M = 2$. Column 3 shows the actual values for $N = 64$, calculated by forming W_L explicitly, determining its singular values, and evaluating the integral in (2.15) numerically. Column 4 shows the upper bound for $N = 64$, calculated by (2.14) using the same singular values. The upper bounds are barely larger than the true values. Column 5 shows the upper bound for large enough N calculated from (3.5). The condition $N > (2^L - 1)(m - 1)$ is satisfied for $L \leq 5$.

4. Some Technical Results

In preparation for estimating the growth of norms as $L \to \infty$, we prove two technical lemmas. Throughout this section, c denotes an unspecified constant, not necessarily the same from case to case. We denote by $\mathbf{1}$ the function which is identically equal to 1.

Throughout this section, p is an even, non-negative trigonometric polynomial of degree d, i.e.

$$p(\theta) = \sum_{k=-d}^{d} p_k e^{ik\theta}, \qquad p_{-k} = p_k.$$

The mappings

$$\begin{aligned} M_p f(\theta) &= p(\theta) f(2\theta), \\ T_p f(\theta) &= p(\frac{\theta}{2}) f(\frac{\theta}{2}) + p(\frac{\theta}{2} + \pi) f(\frac{\theta}{2} + \pi) \end{aligned}$$

are bounded linear operators on $L^1[0, 2\pi]$, the space of integrable, 2π-periodic functions on $\mathbb{R}$ with norm

$$\|f\| = \frac{1}{2\pi} \int_0^{2\pi} |f(\theta)| \, d\theta.$$

Let ρ_p be the spectral radius of the matrix P given by

$$P_{ij} = 2p_{2i-j}, \qquad i, j = -d \ldots d.$$

Lemma 4.1. *Let f_0 be a bounded function in $L^1[0, 2\pi]$, and define f_n recursively by*

$$f_n = M_p f_{n-1}.$$

Then

$$\limsup_{n \to \infty} \|f_n\|^{1/n} \leq \rho_p / 2. \tag{4.1}$$

Equivalently, for any $\epsilon > 0$ there exists a constant c so that

$$\|f_n\| \leq c(\rho_p/2 + \epsilon)^n. \tag{4.2}$$

Proof. The equivalence of (4.1) and (4.2) is easy to establish.

Directly by induction, or by using lemma 3.2 in [1] with $f = \mathbf{1}$, it is easy to show that

$$\|M_p^n \mathbf{1}\| = 2^{-n} \|T_p^n \mathbf{1}\|.$$

For positive p, theorem 3.1 in [7] shows that $T_p^n \mathbf{1} / \|T_p^n \mathbf{1}\|$ converges to a d-degree trigonometric polynomial which is an eigenvector of T_p, with eigenvalue equal to the spectral radius of T_p. This $(2d+1)$-dimensional eigenvalue problem reduces to the matrix P (which also appears in [1] and other places).

REMARK: Eirola's definition of T_p is different from ours, but for the present purpose both are equivalent. He uses functions on $[0, \pi]$, we use even functions on $[0, 2\pi]$. □

So far, we have that for $f_0 = \mathbf{1}$ and positive p we obtain in fact

$$\lim_{n\to\infty} \|f_n\|^{1/n} = \rho_p / 2.$$

For non-negative p and $f_0 = \mathbf{1}$, choose any $\epsilon > 0$, and choose $\delta > 0$ small enough so that

$$\rho_{p+\delta} \leq \rho_p + 2\epsilon.$$

Then

$$\|f_n\| = \|M_p^n \mathbf{1}\| \leq \|M_{p+\delta}^n \mathbf{1}\| \leq c(\rho_{p+\delta}/2 + \epsilon)^n \leq c(\rho_p/2 + 2\epsilon)^n.$$

Finally, for arbitrary bounded f_0, we apply

$$\|f_n\| = \|M_p^n f_0\| \leq \max_\theta |f(\theta)| \cdot \|M_p^n \mathbf{1}\|.$$

□

Lemma 4.2. *Let f_0 be a bounded function in $L^1[0, 2\pi]$, and define f_n recursively by*

$$f_n = M_p f_{n-1} + q,$$

where q is a fixed bounded function in $L^1[0, 2\pi]$.

(a) If $\rho_p < 2$, then

$$f_n \to Q = \sum_{j=0}^{\infty} M_p^j q \qquad \text{as } n \to \infty;$$

in particular

$$\lim_{n\to\infty} \|f_n\| = const.$$

(b) If $\rho_p \geq 2$, then

$$\limsup_{n\to\infty} \|f_n\|^{1/n} \leq \rho_p/2.$$

Proof. f_n has the explicit representation

$$f_n = M_p^n f_0 + \sum_{j=0}^{n-1} M_p^j q. \tag{4.3}$$

(a) If $\rho_p < 2$, repeated application of lemma 4.1 shows that the infinite series for Q converges, so Q is well-defined, and

$$\|f_n - Q\| \leq \|M_p^n f_0\| + \sum_{j=n}^{\infty} \|M_p^j q\| \to 0 \quad \text{as } n \to \infty.$$

(b) Choose any $\epsilon > 0$. Take norms in (4.3) and apply lemma 4.1:

$$\|f_n\| \le c(\rho_p/2+\epsilon)^n + c\sum_{j=0}^{n-1}(\rho_p/2+\epsilon)^j \le c(\rho_p/2+\epsilon)^n.$$

□

5. Asymptotic Norm Estimates

In this section, we establish bounds on the growth of the 2-norm and average 2-norm of W_L, P_L as $L \to \infty$. We state the main results first, then prove them one by one. Throughout this section, $\|.\|_2$ stands for the upper bound on the 2-norm introduced in section 3.1 (independent of N), and $\|.\|_{2,\text{avg}}$ stands for the upper bounds (3.5), (3.7) on the average 2-norm for large enough N.

Recall the definition of $h(\theta)$, $g(\theta)$ in (2.3). We calculate the trigonometric polynomials

$$|h(\theta)|^2 = \sum_{k=-d_1}^{d_1} \alpha_k e^{ik\theta},$$

$$|h(\theta)|^2 + |g(\theta)|^2 = \sum_{k=-d_2}^{d_2} \beta_k e^{ik\theta},$$

where

$$\alpha_k = \frac{1}{2}\sum_j h_{j+k}h_j,$$

$$\beta_k = \frac{1}{2}\sum_j \left[h_{j+k}h_j + g_{j+k}g_j\right].$$

Let ρ_W be the spectral radius of the matrix A given by

$$A_{ij} = 2\alpha_{2i-j}, \qquad i,j = -d_1\ldots d_1, \tag{5.1}$$

and let ρ_P be the spectral radius of the matrix B given by

$$B_{ij} = 2\beta_{2i-j}, \qquad i,j = -d_2\ldots d_2. \tag{5.2}$$

Theorem 5.1. *(a) If $\rho_W < 1$, then*

$$\|W_L\|_2 \le \textit{const} \qquad \textit{for all } L.$$

If $\rho_W \ge 1$, then

$$\limsup_{L\to\infty} \|W_L\|_2^{1/L} \le \sqrt{\rho_W}.$$

(b) If $\rho_P < 1$, then

$$\lim_{L\to\infty} \|P_L\|_2 = 0.$$

If $\rho_P \ge 1$, then

$$\limsup_{L\to\infty} \|P_L\|_2^{1/L} \le \sqrt{\rho_P}.$$

(c) If $\rho_W < 2$, then

$$\lim_{L\to\infty} \|W_L\|_{2,avg} = \textit{const}.$$

If $\rho_W \geq 2$, then

$$\limsup_{L\to\infty} \|W_L\|_{2,avg}^{1/L} \leq \sqrt{\rho_W/2}.$$

(d) If $\rho_P < 2$, then

$$\lim_{L\to\infty} \|P_L\|_{2,avg} = 0.$$

If $\rho_P \geq 2$, then

$$\lim_{L\to\infty} \|P_L\|_{2,avg}^{1/L} = \sqrt{\rho_P/2}.$$

Proof. We will prove these statements in increasing order of difficulty.

(d) From section 3.2 we recall that

$$\|P_L\|_{2,\text{avg}}^2 = \text{constant term in } p_L,$$

where p_L is defined by (3.8). Equivalently,

$$\|P_L\|_{2,\text{avg}}^2 = \frac{1}{2\pi}\int_0^{2\pi} p_L(\theta)\,d\theta.$$

The result then follows directly from lemma 4.1. The stronger result of Eirola applies, since $|h(\theta)|^2 + |g(\theta)|^2 > 0$.

(c) With the same reasoning,

$$\|W_L\|_{2,\text{avg}}^2 = \frac{1}{2\pi}\int_0^{2\pi} f_L(\theta)\,d\theta,$$

where f_L is defined by (3.6). The result follows from lemma 4.2.

(a) Corresponding to (2.8), we define bounded operators H, G on $L^2[0,2\pi]$ (the space of 2π-periodic, square integrable functions on $\mathbb{R}$ with norm (2.6)) by

$$\begin{aligned} Hf(\theta) &= h(\frac{\theta}{2})f(\frac{\theta}{2}) + h(\frac{\theta}{2}+\pi)f(\frac{\theta}{2}+\pi),\\ Gf(\theta) &= g(\frac{\theta}{2})f(\frac{\theta}{2}) + g(\frac{\theta}{2}+\pi)f(\frac{\theta}{2}+\pi). \end{aligned}$$

By (2.12) and (2.7),

$$\begin{aligned} \|\hat{W}_L\hat{\mathbf{s}}^0\|_2^2 &= \|\hat{\mathbf{s}}^L\|_2^2 + \|\hat{\mathbf{d}}^L\|_2^2 + \cdots + \|\hat{\mathbf{d}}^1\|_2^2\\ &= \frac{N}{2^L}\|\hat{s}^L\|_2^2 + \frac{N}{2^L}\|\hat{d}^L\|_2^2 + \cdots + \frac{N}{2}\|\hat{d}^1\|_2^2\\ &= N\left[\|(\frac{H}{\sqrt{2}})^L\hat{s}^0\|_2^2 + \|(\frac{G}{\sqrt{2}})(\frac{H}{\sqrt{2}})^{L-1}\hat{s}^0\|_2^2 + \cdots + \|(\frac{G}{\sqrt{2}})\hat{s}^0\|_2^2\right]\\ &\leq N\|\hat{s}^0\|_2^2\left[\|(\frac{H}{\sqrt{2}})^L\|_2^2 + \sum_{j=0}^{L-1}\|(\frac{G}{\sqrt{2}})(\frac{H}{\sqrt{2}})^j\|_2^2\right]\\ &= \|\hat{\mathbf{s}}^0\|_2^2\left[\|(\frac{H^*}{\sqrt{2}})^L\|_2^2 + \sum_{j=0}^{L-1}\|(\frac{H^*}{\sqrt{2}})^j(\frac{G^*}{\sqrt{2}})\|_2^2\right]. \end{aligned} \tag{5.3}$$

Here H^* is the adjoint of H, which is easily calculated to be

$$H^*f(\theta) = 2\overline{h(\theta)}f(2\theta),$$

and likewise for G^*. The Cauchy-Schwarz inequality shows

$$\|(H^*)^L\|_2 \leq 2^L \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |h(\theta)|^2 |h(2\theta)|^2 \ldots |h(2^{L-1}\theta)|^2 \, d\theta} \tag{5.4}$$

and similarly for G^* and the mixed terms. Substituting this into (5.3) leads back to the same calculation as in (c), with $|h(\theta)|^2$, $|g(\theta)|^2$ in (3.6) replaced by $2|h(\theta)|^2$, $2|g(\theta)|^2$.

(b) With the same setup as in (a),

$$\begin{aligned} \|P_2 \hat{\mathbf{s}}^0\|_2^2 &= N \left[\|(\frac{H^*}{\sqrt{2}})^2 \hat{s}^0\|_2^2 + \|(\frac{H^*}{\sqrt{2}})(\frac{G^*}{\sqrt{2}})\hat{s}^0\|_2^2 + \|(\frac{G^*}{\sqrt{2}})(\frac{H^*}{\sqrt{2}})\hat{s}^0\|_2^2 + \|(\frac{G^*}{\sqrt{2}})^2 \hat{s}^0\|_2^2 \right] \\ &\leq \|\hat{\mathbf{s}}^0\|_2^2 \left[\|(\frac{H^*}{\sqrt{2}})^2\|_2^2 + \|(\frac{H^*}{\sqrt{2}})(\frac{G^*}{\sqrt{2}})\|_2^2 + \|(\frac{G^*}{\sqrt{2}})(\frac{H^*}{\sqrt{2}})\|_2^2 + \|(\frac{G^*}{\sqrt{2}})^2\|_2^2 \right], \end{aligned} \tag{5.5}$$

and similarly for higher L. Substituting (5.4) into this leads back to the calculation in (d), with a factor of 2 added. □

Remark: The estimates in (c), (d) are exact. (We suspect that the lim sup in (c) could be shown to be a limit, at the cost of further machinery). In contrast, (a), (b) may be strict inequalities resulting from strict inequalities in (5.3), (5.5). In fact, the actual growth rate of the norms in some of our numerical examples appears to be slower than the estimate. However, the upper bounds are almost achieved in other cases, so we don't believe these estimates can be improved without imposing further conditions on the wavelet coefficients. □

As stated at the beginning of this paper, theorem 5.1 imposes no restrictions on the recursion coefficients other than the perfect reconstruction condition (2.4). However, most wavelets used in practice (including all our examples) satisfy the minimal additional conditions

$$h(0) = g(\pi) = 1, \qquad h(\pi) = g(0) = 0, \tag{5.6}$$

or equivalently

$$\sum h_{2k} = \sum h_{2k+1} = \sum (-1)^k g_{2k} = \sum (-1)^{k+1} g_{2k+1} = \sqrt{2}/2.$$

With these additional assumptions, the conclusions of theorem 5.1 can be sharpened.

Theorem 5.2. *If conditions* (5.6) *are satisfied, then*

(a) $\rho_W \geq 1$, *so*

$$\limsup_{L\to\infty} \|W_L\|_2^{1/L} \leq \sqrt{\rho_W}.$$

If $\rho_W < 2$, *then*

$$\lim_{L\to\infty} \|W_L\|_{2,avg} = const.$$

If $\rho_W \geq 2$, *then*

$$\limsup_{L\to\infty} \|W_L\|_{2,avg}^{1/L} \leq \sqrt{\rho_W/2}.$$

(b) $\rho_P > 2$, *so*

$$\limsup_{L\to\infty} \|P_L\|_2^{1/L} \leq \sqrt{\rho_P}$$

$$\lim_{L\to\infty} \|P_L\|_{2,avg}^{1/L} = \sqrt{\rho_P/2}.$$

Remark: The conclusions of theorem 5.2 for the standard L^2 norm are already contained in [1] and [2]: The norms for the wavelet decomposition may remain bounded independent of L; this happens if $\rho_W = 1$, which actually occurs in many cases. The norms for the wave packet decomposition always grow without bound.

We observe that there may be wavelets which are stable "on average", but not in the strict sense. Such wavelets do exist (see table 6 for examples). The wave packet decomposition is still always unstable, even in the average 2-norm. □

Proof. (a) The additional conditions (5.6) imply that each column of the matrix A sums to 1, so $(1, 1, \ldots, 1)$ is a left eigenvector with eigenvalue 1. This implies $\rho_W \geq 1$. The rest is simply theorem 5.1 (a), with one impossible case deleted.

(b) For

$$p(\theta) = |h(\theta)|^2 + |g(\theta)|^2,$$

theorem 6.4 in [2] states (in our notation) that

$$\|M_p^n \mathbf{1}\| \geq c \cdot \lambda^n$$

for some $c > 0$, $\lambda > 1$. Using the relationship between M_p^n and T_p^n as in the proof of lemma 1, this implies $\rho_p > 2$. □

Remark: It was pointed out by one of the referees that if we assume $h(\pi) \neq 0$ instead of (5.6), then $\rho_W > |h(0)|^2$. For the usual normalization $h(0) = 1$ this implies instability.

The argument is the following:

Let $p(\theta) = |h(\theta)|^2$, and define T_p as in section 4. By the theory of positive operators, T_p has a positive largest eigenvalue ρ with non-negative eigenfunction $F(\theta)$. By iterating T_p and looking at $\theta = 0$, we find

$$\rho^n F(0) = p(0)^n F(0) + \sum_{m=1}^{2^n-1} \prod_{j=1}^{n} p(2^{-j} \cdot 2m\pi) F(2^{-n} \cdot 2m\pi).$$

The second term on the right is strictly positive, so $F(0) > 0$, and

$$\rho^n F(0) > p(0)^n F(0),$$

from which the statement follows. □

Table 3 shows $\|W_L\|_2$ and $\|W_L\|_{2,\text{avg}}$ for $L = 1, \ldots, 7$ for the dual binomial wavelet with two vanishing moments, for which $\rho_W \approx 0.790569$, $\rho_P \approx 1.118034$. Both $\|W_L\|/\|W_{L-1}\|$ and $\|W_L\|^{1/L}$ can be seen to converge towards the limits predicted by theorem 5.1. Convergence in this example is actually slower than in many other cases (see tables 4, 5).

Table 3 goes near here

Table 4 shows ρ_W and ρ_P and the predicted growth rates of norms as $L \to \infty$ for a number of different wavelets. On the second line for each wavelet, we show the ratio of actual norms for $L = 7$ and $L = 6$. In many cases, there is already excellent agreement with the predicted limit. Any discrepancies in the average 2-norms are due to slow convergence (as in table 3). Discrepancies in the 2-norms could be due to slow convergence or to strict inequalities in theorem 5.1. The latter is especially noticeable in the Cohen and Barlaud wavelets, for the wave packet case. We note that the stability of the Barlaud wavelet and all Cohen wavelets with minimal regularity is already known.

Table 5 shows the same numbers for the dual wavelets.

Tables 4 and 5 go near here

For the cases where the norms approach a finite limit as $L \to \infty$, the values for $L = 7$ are shown in table 6. This occured for several of the wavelet transforms, but not for any of the wave packet transforms.

Table 6 goes near here

6. Interpretation and Numerical Examples

What are the practical implications of the estimates derived in this paper?

The numbers in tables 4 through 6 suggest that many biorthogonal wavelet decompositions are in fact numerically stable; among the types we tested, only the dual binomial wavelets for larger M have stability problems. The wave packet decomposition has mild instabilities for larger L, but even there the average error growth is usually very slow.

A detailed interpretation depends on the application being considered. We discuss three examples in this section. Numerical experiments are illustrated with four different wavelets (results for other wavelets are similar):

- The Daubechies wavelet with two vanishing moments, as a reference case (orthogonal, completely stable);
- the Cohen wavelet with $(N, \tilde{N}) = (2, 2)$ (stable);
- the dual binomial wavelet with four vanishing moments (mildly unstable);
- and the dual binomial wavelet with six vanishing moments (unstable).

Each case is tested on four different unit vectors of length 16,384 each, with $L = 10$ levels of decomposition:

(a) a random vector (normally distributed with mean 0; the same random vector is used for all examples);
(b) a smooth vector (discretization of $\cos t \cos 2t \cos 3t \cos 5t$ on $[0, 2\pi]$);
(c) a music sample (starting segment of `spacemusic.au`, distributed by Sun Microsystems with their Sparc 1+ workstation); the music was digitized at 8,000 samples per second, which corresponds roughly to telephone line quality;
(d) a digitized speech sample (starting segment of `sample.au`, from the same source as (c)).

We make the simplifying assumption that only errors in the original data and its transform affect the result, ignoring errors introduced during intermediate steps.

All calculations were done in double precision on a DECstation 5000/125 using `matlab`, an interactive matrix calculator.

Example 1: Decomposition. We perform a wavelet transform on the data and estimate the error in the transform. This comes up in signal processing applications where one tries to get information from the transform coefficients.

First we look at the growth of the vectors themselves. Columns 1 and 2 of table 7 show the actual and expected length of the wavelet transform of the four test vectors. The expected lengths are

(a) For a random vector, we would expect the length to grow by $\|W_{10}\|_{2,\mathrm{avg}}$. We estimated W_{10} by extrapolating from the values in table 4.
(b) If the signal $\mathbf{s}^0$ is very smooth and the wavelet has several vanishing moments, the norm of the signal remains approximately the same during transformation ($\mathbf{s}^L$ is a

compressed and scaled version of $\mathbf{s}^0$, with approximately the same norm, and the $\mathbf{d}^n$ are very small). This is indicated by $O(1)$ in the table.

(c), (d) An actual signal would fall somewhere between these extremes, depending on its smoothness. The music and speech signal we tested are rapidly varying, and closer to the random case.

The observed growth of the vectors matches the prediction closely.

For roundoff error, our results imply that if the roundoff error in $\mathbf{s}^0$ is random, we would expect it to grow like signal (a). If the error is more systematic, it might behave more like (c) or (d), i.e. slightly better than random, but not by much.

In practice, this means that for rapidly varying signals the coefficients themselves and the error are of similar magnitude in ℓ_2. Still, individual coefficients may carry a large relative error, so unstable wavelets (like the dual binomial wavelets for higher M) would not be suitable for this application. Using unstable wavelets on a smooth signal would lead to very large errors in the coefficients.

Example 2: Decomposition/Reconstruction. Decomposition followed by reconstruction, without any kind of processing in between, is not something that usually comes up in applications, but it serves to illustrate error growth.

A key observation is that only errors introduced during the reconstruction affect the result. Any errors introduced during the decomposition will disappear again during the reconstruction. (One way to visualize that is to decompose the error vector into singular vectors: During decomposition, the jth error component gets magnified by the corresponding singular value σ_j, but during reconstruction it gets multiplied by $1/\sigma_j$).

We assume that the decomposed vector $W_L\mathbf{s}^0$ acquires a roundoff error of size $\|W_L\mathbf{s}^0\|\epsilon$, where ϵ is the machine epsilon. Assuming this error is randomly distributed, we expect an error of size $\|\tilde{W}_L^T\|_{2,\mathrm{avg}}\|W_L\mathbf{s}^0\|\epsilon$ after reconstruction. Columns 3 and 4 of table 7 show predicted and actual values for

$$\frac{\text{error after reconstruction}}{\text{original error}} \approx \frac{\text{error after reconstruction}}{\epsilon}$$

For the machine epsilon, we used `matlab`'s built-in constant $eps \approx 2.22 \cdot 10^{-16}$. This choice produces correct results for the calibration case of Daubechies wavelets and seems suitable.

Some of the actual results are off by an order of magnitude, but there is a definite correlation between predicted and actual results. We attribute the differences to the fact that the errors in the transform are not truly random.

Example 3: Signal Compression. A common application of the wavelet transform lies in signal compression. When a smooth vector is decomposed using wavelets with several vanishing moments, many coefficients in the decomposed vector are very small and can be set to zero.

As an illustration, we zeroed all coefficients below 10^{-3}, which produced an absolute error of order 10^{-3} in $W_{10}\mathbf{s}^0$. The truncated transform was reconstructed, and we then calculated the ratio

$$\frac{\text{error in reconstructed signal}}{\text{error due to truncation in transformed signal}}.$$

Assuming truncation produces a random error, we would expect this growth factor to be $\|\tilde{W}_{10}\|_{2,\mathrm{avg}}$. The actual results are surprisingly close to the estimates, considering that the assumption of random error is not really satisfied. We attribute the quality of our predictions to the fact that even the worst case error $\|W_{10}\|_2$ is very small in all cases (see table 6). In fact, it would be hard to distinguish average and worst case errors by numerical experiments.

We note that even the most unstable of the wavelets we tested performs quite well in this particular application. This fact was also observed in [8]. Roughly stated, the instability is concentrated in the decomposition, where it does not do much harm, but the reconstruction is stable.

Wavelets with stability problems may actually be preferred if they give better compression.

7. Acknowledgements

I would like to thank Don Ramirez and Charles Dunkl for providing me with a preprint of their paper [6] and a copy of a program to calculate the integrals in section 3.2 numerically. I am also indebted to the referees for valuable suggestions.

References

1. A. Cohen and I. Daubechies, *A stability criterion for biorthogonal wavelet bases and their related subband coding scheme*, Duke Math. J., 68 (1992), pp. 313–335.
2. ———, *On the instability of arbitrary biorthogonal wavelet packets*, SIAM J. Math. Anal., 24 (1993), pp. 1340–1354.
3. A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560.
4. R. R. Coifman, Y. Meyer, S. Quake, and M. V. Wickerhauser, *Signal processing and compression with wave packets*, in Proceedings of the Marseilles Wavelet Meeting 1989, 1991.
5. I. Daubechies, *Ten Lectures on Wavelets*, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.
6. C. F. Dunkl and D. E. Ramirez, *Computing hyperelliptic integrals for surface measure of ellipsoids.* to appear in ACM Trans. Math. Software.
7. T. Eirola, *Sobolev characterization of solutions of dilation equations*, SIAM J. Math. Anal., 23 (1992), pp. 1015–1030.
8. F. Keinert, *Biorthogonal wavelets for fast matrix computations*, Appl. Comput. Harmonic Anal., 1 (1994), pp. 147–156.

Table 1. Coefficients of biorthogonal wavelets used in numerical examples. Dual coefficients are derived from (2.4). See text for explanation and references.

	h_{-4} g_{-4}	h_{-3} g_{-3}	h_{-2} g_{-2}	h_{-1} g_{-1}	h_0 g_0	h_1 g_1	h_2 g_2	h_3 g_3	h_4 g_4	h_5 g_5	h_6 g_6	factor
dual binomial												
$M = 2$					3	2	-1					$\sqrt{2}/4$
					1	-2	1					$\sqrt{2}/4$
3					-1	3	3	-1				$\sqrt{2}/4$
					1	-3	3	-1				$\sqrt{2}/8$
4					-5	20	10	-12	3			$\sqrt{2}/16$
					1	-4	6	-4	1			$\sqrt{2}/16$
5					3	-15	20	20	-15	3		$\sqrt{2}/16$
					1	-5	10	-10	5	-1		$\sqrt{2}/32$
6					7	-42	77	28	-63	30	-5	$\sqrt{2}/32$
					1	-6	15	-20	15	-6	1	$\sqrt{2}/64$
complementary												
$M = 2$	same as dual binomial, $M = 2$											
3			1	0	10	8	-3					$\sqrt{2}/16$
			-1	0	6	-8	3					$\sqrt{2}/16$
4			1	0	23	16	-9	0	1			$\sqrt{2}/32$
			-1	0	9	-16	9	0	-1			$\sqrt{2}/32$
5	-3	0	20	0	166	128	-60	0	5			$\sqrt{2}/256$
	3	0	-20	0	90	-128	60	0	-5			$\sqrt{2}/256$
6	-3	0	25	0	362	256	-150	0	25	0	-3	$\sqrt{2}/512$
	3	0	-25	0	150	-256	150	0	-25	0	3	$\sqrt{2}/512$
Cohen												
$(N, \tilde{N}) = (1, 3)$					1	1						$\sqrt{2}/2$
			-1	-1	8	-8	1	1				$\sqrt{2}/16$
$(1, 5)$					1	1						$\sqrt{2}/2$
	3	3	-22	-22	128	-128	22	22	-3	-3		$\sqrt{2}/256$
$(2, 2)$				1	2	1						$\sqrt{2}/4$
				1	2	-6	2	1				$\sqrt{2}/8$
$(2, 4)$				1	2	1						$\sqrt{2}/4$
		-3	-6	16	38	-90	38	16	-6	-3		$\sqrt{2}/128$
$(3, 1)$				1	3	3	1					$\sqrt{2}/8$
				1	3	-3	-1					$\sqrt{2}/4$
$(3, 3)$				1	3	3	1					$\sqrt{2}/8$
		-3	-9	7	45	-45	-7	9	3			$\sqrt{2}/64$
Barlaud			-1	5	12	5	-1					$\sqrt{2}/20$
			3	-15	-73	170	-73	-15	3			$\sqrt{2}/280$

TABLE 2. Comparison of actual norms and upper bounds on the norms for the dual binomial wavelet with two vanishing moments, $N = 64$.

levels L	$\|W_L\|_2$ $N=64$ actual	$\|W_L\|_2$ any N bound	$\|W_L\|_{2,\mathrm{avg}}$ $N=64$ actual	$\|W_L\|_{2,\mathrm{avg}}$ $N=64$ bound	$\|W_L\|_{2,\mathrm{avg}}$ large N bound
1	1.618034	1.618034	1.116336	1.118034	1.118034
2	2.187316	2.187316	1.207889	1.211920	1.211920
3	2.814412	2.814412	1.272744	1.279343	1.279343
4	3.500128	3.500128	1.316455	1.325641	1.325641
5	4.255668	4.255668	1.344838	1.356656	1.356656
6	4.255668	5.089263	1.344838	1.356656	1.377104
7	4.255668	6.010410	1.338852	1.350885	1.390434

TABLE 3. Comparison of actual and predicted growth of norms as $L \to \infty$, for the dual binomial wavelet with two vanishing moments.

L	$\|W_L\|_2$	$\frac{\|W_L\|_2}{\|W_{L-1}\|_2}$	$\|W_L\|_2^{1/L}$	$\|W_L\|_{2,\mathrm{avg}}$	$\frac{\|W_L\|_{2,\mathrm{avg}}}{\|W_{L-1}\|_{2,\mathrm{avg}}}$	$\|W_L\|_{2,\mathrm{avg}}^{1/L}$
1	1.618034		1.618034	1.118034		1.118034
2	2.187316	1.351836	1.478958	1.211920	1.083974	1.100872
3	2.814412	1.286696	1.411874	1.279343	1.055633	1.085581
4	3.500128	1.243645	1.367795	1.325641	1.036189	1.073017
5	4.255668	1.215861	1.335960	1.356656	1.023396	1.062904
6	5.089263	1.195878	1.311523	1.377104	1.015073	1.054778
7	6.010410	1.180998	1.292028	1.390434	1.009680	1.048214
predicted limit		1.118034	1.118034		1.000000	1.000000

Table 4. Predicted and observed growth rates for 2-norms and average 2-norms as $L \to \infty$, for various biorthogonal wavelets. The first row contains theoretically predicted values, the second row contains actually observed estimates, calculated from the ratio of norms for 7 and 6 levels.

	ρ_W	$\|W_L\|_{2,\text{avg}}$	$\|W_L\|_2$	ρ_P	$\|P_L\|_{2,\text{avg}}$	$\|P_L\|_2$
dual binomial, $M = 2$	1.250000	1.000000	1.118034	2.500000	1.118034	1.581139
		1.009680	1.180998		1.118034	1.296023
3	2.171165	1.041913	1.473487	3.164338	1.257843	1.778859
		1.095362	1.480446		1.257844	1.494621
4	5.948812	1.724647	2.439019	6.566593	1.811987	2.562536
		1.724958	2.444670		1.812040	2.450430
5	12.551057	2.505101	3.542747	12.951300	2.544730	3.598791
		2.505009	3.548769		2.544661	3.549542
6	36.290310	4.259713	6.024144	36.534659	4.274030	6.044391
		4.259428	6.024257		4.273790	6.026930
complementary, $M = 3$	1.000000	1.000000	1.000000	2.234411	1.056980	1.494795
		1.002701	1.104922		1.056979	1.200633
4	1.259807	1.000000	1.122411	2.546766	1.128443	1.595859
		1.010732	1.185437		1.128442	1.318424
5	1.100252	1.000000	1.048929	2.374919	1.089706	1.541077
		1.005450	1.139716		1.089704	1.283048
6	1.271085	1.000000	1.127424	2.583461	1.136543	1.607315
		1.011470	1.190090		1.136541	1.349833
Cohen, $(N, \tilde{N}) = (1, 3)$	1.000000	1.000000	1.000000	2.033044	1.008227	1.425848
		1.000118	1.015260		1.008227	1.099671
$(1, 5)$	1.000000	1.000000	1.000000	2.068254	1.016920	1.438143
		1.000222	1.017930		1.016918	1.155532
$(2, 2)$	1.000000	1.000000	1.000000	2.196252	1.047915	1.481976
		1.000285	1.000000		1.047914	1.259969
$(2, 4)$	1.000000	1.000000	1.000000	2.176069	1.043089	1.475151
		1.000145	1.000000		1.043086	1.262258
$(2, 6)$	1.000000	1.000000	1.000000	2.188206	1.045994	1.479259
		1.000118	1.000000		1.045988	1.274642
$(3, 1)$	1.000000	1.000000	1.000000	3.086562	1.242289	1.756861
		1.001334	1.000000		1.242289	1.632173
$(3, 3)$	1.000000	1.000000	1.000000	2.755849	1.173850	1.660075
		1.000739	1.000000		1.173844	1.554780
$(3, 5)$	1.000000	1.000000	1.000000	2.669468	1.155307	1.633851
		1.000554	1.000000		1.155279	1.535135
$(3, 7)$	1.000000	1.000000	1.000000	2.643471	1.149668	1.625875
		1.000475	1.000000		1.149805	1.524651
Barlaud	1.000000	1.000000	1.000000	2.001102	1.000276	1.414603
		1.000000	1.000092		1.000276	1.024219

Table 5. Predicted and observed growth rates for 2-norms and average 2-norms as $L \to \infty$, for the duals of wavelets in table 4. The first row contains theoretically predicted values, the second row contains actually observed estimates, calculated from the ratio of norms for 7 and 6 levels.

	$\rho_{\tilde{W}}$	$\|\tilde{W}_L\|_{2,\text{avg}}$	$\|\tilde{W}_L\|_2$	$\rho_{\tilde{P}}$	$\|\tilde{P}_L\|_{2,\text{avg}}$	$\|\tilde{P}_L\|_2$
dual binomial,						
$M = 2$	1.000000	1.000000	1.000000	2.500000	1.118034	1.581139
		1.000914	1.002117		1.118034	1.402935
3	1.000000	1.000000	1.000000	3.086562	1.242289	1.756861
		1.001334	1.000000		1.242289	1.632173
4	1.000000	1.000000	1.000000	5.191908	1.611196	2.278576
		1.002089	1.000386		1.611077	2.210747
5	1.000000	1.000000	1.000000	8.012187	2.001523	2.830581
		1.002370	1.000905		1.997884	2.764346
6	1.000000	1.000000	1.000000	15.753056	2.806515	3.969012
		1.002630	1.002111		2.776267	3.871454
complementary,						
$M = 3$	1.000000	1.000000	1.000000	2.209409	1.051049	1.486408
		1.000689	1.005504		1.051049	1.218328
4	1.000000	1.000000	1.000000	2.489314	1.115642	1.577756
		1.001228	1.005480		1.115641	1.371614
5	1.000000	1.000000	1.000000	2.312643	1.075324	1.520738
		1.000974	1.006225		1.075322	1.271940
6	1.000000	1.000000	1.000000	2.497135	1.117393	1.580233
		1.001327	1.006023		1.117391	1.363768
Cohen,						
$(N, \tilde{N}) = (1, 3)$	1.000000	1.000000	1.000000	2.029663	1.007388	1.424662
		1.000105	1.006665		1.007388	1.037886
$(1, 5)$	1.000000	1.000000	1.000000	2.054355	1.013498	1.433302
		1.000189	1.007054		1.013497	1.086007
$(2, 2)$	1.000000	1.000000	1.000000	2.173152	1.042390	1.474162
		1.000640	1.031490		1.042389	1.146879
$(2, 4)$	1.000000	1.000000	1.000000	2.148085	1.036360	1.465635
		1.000257	1.002241		1.036361	1.216647
$(2, 6)$	1.000000	1.000000	1.000000	2.151717	1.037236	1.466873
		1.000220	1.001012		1.037235	1.224165
$(3, 1)$	2.171165	1.041913	1.473487	3.164338	1.257843	1.778859
		1.095362	1.480446		1.257844	1.494621
$(3, 3)$	1.000000	1.000000	1.000000	2.603259	1.140890	1.613462
		1.003222	1.068150		1.140917	1.427730
$(3, 5)$	1.000000	1.000000	1.000000	2.520752	1.122665	1.587688
		1.001320	1.009170		1.122744	1.533942
$(3, 7)$	1.000000	1.000000	1.000000	2.501492	1.118368	1.581611
		1.001122	1.001533		1.117875	1.545197
Barlaud	1.000000	1.000000	1.000000	2.001042	1.000261	1.414582
		1.000000	1.000156		1.000261	1.006566

TABLE 6. Estimated limit of norms as $L \to \infty$. Shown are the norms for $L = 7$, which are close to the limit values. Blank spaces indicate that the norms tend to ∞.

	$\|W_L\|_{2,\text{avg}}$	$\|W_L\|_2$	$\|\tilde{W}_L\|_{2,\text{avg}}$	$\|\tilde{W}_L\|_2$
dual binomial, $M = 2$	1.39		1.19	1.72
3			1.38	2.00
4			2.02	3.09
5			2.77	4.26
6			4.55	7.20
complementary, $M = 3$	1.15	3.21	1.10	1.59
4	1.42		1.22	1.90
5	1.26		1.15	1.78
6	1.44		1.23	1.97
Cohen, $(N, \tilde{N}) = (1, 3)$	1.02	1.36	1.02	1.29
$(1, 5)$	1.03	1.53	1.03	1.40
$(2, 2)$	1.07	1.41	1.09	2.23
$(2, 4)$	1.05	1.41	1.07	1.75
$(2, 6)$	1.05	1.41	1.06	1.74
$(3, 1)$	1.38	2.00		
$(3, 3)$	1.24	2.00	1.39	4.55
$(3, 5)$	1.20	2.00	1.29	2.95
$(3, 7)$	1.19	2.00	1.26	2.76
Barlaud	1.00	1.18	1.00	1.08

TABLE 7. Numerical experiments with $N = 16384$, $L = 10$. (a) Random vector (b) smooth vector (c) digitized music (d) digitized speech. See section 6 for further explanations.

	$\|W_L\mathbf{s}\|_2$ expected	$\|W_L\mathbf{s}\|_2$ actual	$\|\tilde{W}_L^T W_L\mathbf{s} - \mathbf{s}\|_2$ expected	$\|\tilde{W}_L^T W_L\mathbf{s} - \mathbf{s}\|_2$ actual	growth of truncation error expected	growth of truncation error actual
Daubechies $M = 2$						
(a)	1.0	1.0	1.0	1.1	1.0	1.0
(b)	1.0	1.0	1.0	5.0	1.0	1.0
(c)	1.0	1.0	1.0	1.4	1.0	1.0
(d)	1.0	1.0	1.0	1.7	1.0	1.0
Cohen $(N, \tilde{N}) = (2, 2)$						
(a)	1.1	1.1	1.2	2.1	1.1	1.1
(b)	$O(1)$	0.9	1.0	7.8	1.1	1.8
(c)	≤ 1.1	1.0	1.1	2.9	1.1	1.0
(d)	≤ 1.1	1.0	1.1	3.5	1.1	1.0
Dual binomial $M = 4$						
(a)	238.7	205.2	410.4	93.3	2.0	2.2
(b)	$O(1)$	2.7	5.4	4.8	2.0	1.5
(c)	≤ 238.7	172.0	344.0	101.2	2.0	2.1
(d)	≤ 238.7	117.9	235.8	51.7	2.0	2.1
Dual binomial $M = 6$						
(a)	$1.6 \cdot 10^6$	$1.3 \cdot 10^6$	$6.0 \cdot 10^6$	$7.0 \cdot 10^5$	4.6	5.1
(b)	$O(1)$	7.8	35.9	7.1	4.6	3.7
(c)	$\leq 1.6 \cdot 10^6$	$3.1 \cdot 10^5$	$1.4 \cdot 10^6$	$9.5 \cdot 10^4$	4.6	4.8
(d)	$\leq 1.6 \cdot 10^6$	$4.0 \cdot 10^5$	$1.8 \cdot 10^6$	$2.3 \cdot 10^5$	4.6	4.6