QUADRATURE FORMULAE AND ASYMPTOTIC ERROR EXPANSIONS FOR WAVELET APPROXIMATIONS OF SMOOTH FUNCTIONS*

WIM SWELDENS†‡ AND ROBERT PIESSENS†

Abstract. This paper deals with typical problems that arise when using wavelets in numerical analysis applications. The first part involves the construction of quadrature formulae for the calculation of inner products of smooth functions and scaling functions. Several types of quadratures are discussed and compared for different classes of wavelets. Since their construction using monomials is ill-conditioned, also a modified, well-conditioned construction using Chebyshev polynomials is presented. The second part of the paper deals with pointwise asymptotic error expansions of wavelet approximations of smooth functions. They are used to to derive asymptotic interpolating properties of the wavelet approximation and to construct a convergence acceleration algorithm. This is illustrated with numerical examples.

Key words. wavelet, multiresolution analysis, quadrature formula, asymptotic error expansion, convergence acceleration, numerical extrapolation

AMS subject classifications. 65D32, 42C05, 41A30

1. Introduction.

1.1. Multiresolution analysis. We will briefly review wavelets and multiresolution analysis. For detailed treatments, one can consult [4, 8, 16, 18]. A *multiresolution analysis* of $\mathrm{L}^2(\mathbf{R})$ is defined as a set of closed subspaces V_j with $j \in \mathbf{Z}$ that exhibit the following properties:

1. $V_j \subset V_{j+1}$,
2. $v(x) \in V_j \Leftrightarrow v(2x) \in V_{j+1}$ and $v(x) \in V_0 \Leftrightarrow v(x+1) \in V_0$,
3. $\bigcup_{j=-\infty}^{+\infty} V_j$ is dense in $\mathrm{L}^2(\mathbf{R})$ and $\bigcap_{j=-\infty}^{+\infty} V_j = \{\mathbf{0}\}$,
4. A *scaling function* $\varphi(x) \in V_0$ exists such that the set $\{\varphi(x \Leftrightarrow l) \mid l \in \mathbf{Z}\}$ is a Riesz basis of V_0.

Consequently, a sequence $(h_k) \in \ell^2(\mathbf{Z})$ exists such that the scaling function satisfies a *refinement equation*

$$\varphi(x) = 2 \sum_k h_k \, \varphi(2x \Leftrightarrow k). \tag{1.1}$$

The set of functions $\{\varphi_{j,l}(x) \mid l \in \mathbf{Z}\}$ with $\varphi_{j,l}(x) = 2^{j/2} \, \varphi(2^j x \Leftrightarrow l)$, is a Riesz basis of V_j. A complement space of V_j in V_{j+1} is denoted by W_j, so $V_{j+1} = V_j \oplus W_j$, and consequently

$$\bigoplus_{j=-\infty}^{+\infty} W_j = \mathrm{L}^2(\mathbf{R}).$$

* Manuscript recieved by the editors ??, accepted for publication ??

† Departement Computerwetenschappen, Katholieke Universiteit Leuven, Celestijnenlaan 200 A, B 3001 Leuven, Belgium

‡ Department of Mathematics, University of South Carolina, Columbia SC 29208. This author is Research Assistant of the National Fund of Scientific Research Belgium, and was partially supported by ONR Grant N00014-90-J-1343.

The complementary spaces are chosen such that

$$v(x) \in W_j \Leftrightarrow v(2x) \in W_{j+1} \quad \text{and} \quad v(x) \in W_0 \Leftrightarrow v(x+1) \in W_0.$$

A function $\psi(x)$ is a *mother wavelet* if the set of functions $\{\psi(x \Leftrightarrow l) \mid l \in \mathbf{Z}\}$ is a Riesz basis of W_0. Since the mother wavelet is also an element of V_1, a sequence $(g_k) \in \ell^2(\mathbf{Z})$ exists such that

$$\psi(x) = 2 \sum_k g_k \, \varphi(2x \Leftrightarrow k). \tag{1.2}$$

The set of wavelet functions $\{\psi_{j,l}(x) \mid l, j \in \mathbf{Z}\}$, with $\psi_{j,l}(x) = 2^{j/2} \, \psi(2^j x \Leftrightarrow l)$, is now a Riesz basis of $\mathrm{L}^2(\mathbf{R})$.

In the case of an *orthogonal* multiresolution analysis, the set of functions $\{\varphi_{j,l}(x) \mid l \in \mathbf{Z}\}$ is an orthonormal basis of V_j and the set $\{\psi_{j,l}(x) \mid j, l \in \mathbf{Z}\}$ is an orthonormal basis of $\mathrm{L}^2(\mathbf{R})$. The orthogonal projection operators onto V_j and W_j, denoted as $\mathcal{P}_j$ and $\mathcal{Q}_j$ respectively, can then be written as

$$\mathcal{P}_j f(x) = \sum_l \langle f, \varphi_{j,l} \rangle \, \varphi_{j,l}(x) \quad \text{and} \quad \mathcal{Q}_j f(x) = \sum_l \langle f, \psi_{j,l} \rangle \, \psi_{j,l}(x).$$

In the *biorthogonal* case, *dual* functions $\tilde{\varphi}_{j,l}(x) = 2^{j/2} \tilde{\varphi}(2^j x \Leftrightarrow l)$ and $\tilde{\psi}_{j,l}(x) = 2^{j/2} \tilde{\psi}(2^j x \Leftrightarrow l)$ exist such that

$$\langle \varphi_{j,l}, \tilde{\varphi}_{j,l'} \rangle = \delta_{l-l'} \quad \text{and} \quad \langle \psi_{j,l}, \tilde{\psi}_{j',l'} \rangle = \delta_{j-j'} \delta_{l-l'} \quad \text{for} \quad j, j', l, l' \in \mathbf{Z}.$$

The projection operators now can be written as

$$\mathcal{P}_j f(x) = \sum_l \langle f, \tilde{\varphi}_{j,l} \rangle \, \varphi_{j,l}(x) \quad \text{and} \quad \mathcal{Q}_j f(x) = \sum_l \langle f, \tilde{\psi}_{j,l} \rangle \, \psi_{j,l}(x).$$

They are not orthogonal in general. The dual scaling function and wavelet satisfy

$$\tilde{\varphi}(x) = 2 \sum_k \tilde{h}_k \, \tilde{\varphi}(2x \Leftrightarrow k), \qquad \tilde{\psi}(x) = 2 \sum_k \tilde{g}_k \, \tilde{\varphi}(2x \Leftrightarrow k), \tag{1.3}$$

and

$$\tilde{\varphi}(2x \Leftrightarrow k) = \sum_l h_{k-2l} \, \tilde{\varphi}(x \Leftrightarrow l) + \sum_l g_{k-2l} \, \tilde{\psi}(x \Leftrightarrow l). \tag{1.4}$$

Define now the Fourier transform of a function $f(x)$ as

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(x) \, e^{-i\omega x} \, dx.$$

Equations (1.1) and (1.2) can then be written as

$$\hat{\varphi}(\omega) = H(w/2) \, \hat{\varphi}(\omega/2) \quad \text{and} \quad \hat{\psi}(\omega) = G(w/2) \, \hat{\varphi}(\omega/2), \tag{1.5}$$

where H and G are 2π-periodic functions given by

$$H(\omega) = \sum_k h_k \, e^{-ik\omega} \quad \text{and} \quad G(\omega) = \sum_k g_k \, e^{-ik\omega}.$$

Similar equations hold for the dual functions. A necessary condition for biorthogonality is then

$$\forall \omega \in \mathbf{R} \; : \; \begin{cases} \overline{H(\omega)}\,\tilde{H}(\omega) & + & \overline{G(\omega)}\,\tilde{G}(\omega) & = & 1 \\ \overline{H(\omega)}\,\tilde{H}(\omega+\pi) & + & \overline{G(\omega)}\,\tilde{G}(\omega+\pi) & = & 0. \end{cases} \tag{1.6}$$

Given the coefficients $\nu_{j,l} = \langle f, \tilde{\varphi}_{j,l} \rangle$ of a function in V_j, one can find its coefficients in the bases of the spaces V_{j-1} and W_{j-1} with decomposition formulae that can be derived using (1.3),

$$\nu_{j-1,l} = \sqrt{2} \sum_k \tilde{h}_{k-2l}\, \nu_{j,k}, \quad \text{and} \quad \mu_{j-1,l} = \langle f, \tilde{\psi}_{j-1,l} \rangle = \sqrt{2} \sum_k \tilde{g}_{k-2l}\, \nu_{j,k}.$$

The inverse step involves a reconstruction formula that can be derived from (1.4),

$$\nu_{j,k} = \sqrt{2} \sum_l h_{k-2l}\, \nu_{j-1,l} + \sqrt{2} \sum_l g_{k-2l}\, \mu_{j-1,l}.$$

When applied recursively, these formulae define a transformation, the *fast wavelet transform* [16, 17].

Examples:. Several families of orthogonal wavelets exist and are described in [2, 15, 18]. A well known example of orthogonal wavelets with compact support was constructed by Ingrid Daubechies in [7]. Compactly supported biorthogonal wavelets are described in [5]. The wavelets constructed in [4, 22] are *semi-orthogonal* wavelets, which means that the wavelets that belong to one subspace W_j are not orthogonal but the subspaces W_j are still mutually orthogonal. The scaling functions and wavelets are compactly supported splines, but the dual functions do not have compact support.

1.2. Wavelets and Polynomials. The moments of the scaling function and wavelet are defined as,

$$\mathcal{M}_p = \int_{-\infty}^{+\infty} x^p\, \varphi(x)\, dx \quad \text{and} \quad \mathcal{N}_p = \int_{-\infty}^{+\infty} x^p\, \psi(x)\, dx \quad \text{with} \quad p \geq 0.$$

The scaling function is usually normalized with $\mathcal{M}_0 = 1$. Similar definitions hold for the dual functions. The moments of a scaling function can be calculated from the coefficients h_k with a p–term recursion relation,

$$\mathcal{M}_p = \frac{1}{2^p \Leftrightarrow 1} \sum_{i=1}^{p} \binom{p}{i} m_i\, \mathcal{M}_{p-i}, \tag{1.7}$$

where the m_i are the discrete moments of the sequence (h_k),

$$m_i = \sum_k h_k\, k^i.$$

The number of vanishing dual wavelet moments is denoted by N where N is at least 1,

$$\tilde{\mathcal{N}}_p = 0 \quad \text{for} \quad 0 \leq p < N, \quad \text{and} \quad \tilde{\mathcal{N}}_N \neq 0.$$

This is equivalent with

$$\hat{\tilde{\psi}}^{(p)}(0) = 0 \quad \text{for} \quad 0 \leq p < N,$$

and, since $\hat{\tilde{\varphi}}(0) = \tilde{\mathcal{M}}_0 \neq 0$, also with

$$\tilde{G}^{(p)}(0) = 0 \quad \text{for} \quad 0 \leq p < N.$$

The sequence $(\tilde{g}_k)$ thus has N vanishing discrete moments. Using equation (1.6) we see that this is also equivalent with

$$H^{(p)}(\pi) = 0 \quad \text{for} \quad 0 \leq p < N.$$

Consequently,

$$i^p \hat{\varphi}^{(p)}(2k\pi) = \delta_k \, \mathcal{M}_p \quad \text{for} \quad 0 \leq p < N, \tag{1.8}$$

and, by the Poisson summation formula,

$$\sum_l (x \Leftrightarrow l)^p \, \varphi(x \Leftrightarrow l) = \mathcal{M}_p \quad \text{for} \quad 0 \leq p < N. \tag{1.9}$$

By rearranging the last expression we see that any polynomial with degree smaller than N can be written as a linear combination of the functions $\varphi(x \Leftrightarrow l)$ with $l \in \mathbf{Z}$. The coefficients in the linear combination themselves are polynomials in l. More precisely, if Π^p denotes the set of polynomials of degree p,

$$\forall A \in \Pi^{N-1}, \exists B \in \Pi^{N-1} : A(x) = \sum_l B(l) \, \varphi(x \Leftrightarrow l) = \sum_l B(x \Leftrightarrow l) \, \varphi(l). \tag{1.10}$$

This last equation holds because left and right hand side are polynomials that match at every integer.

The number of vanishing dual moments determines the order of convergence of the wavelet approximation of smooth functions. More precisely, if $f(x) \in \mathcal{C}^N$, then [11, 20, 21],

$$\| f(x) \Leftrightarrow \mathcal{P}_n f(x) \| = \mathcal{O}(h^N) \quad \text{with} \quad h = 2^{-n}. \tag{1.11}$$

The conditions (1.8) are usually referred to as the Strang–Fix conditions.

1.3. Contents of the paper. This paper deals with typical problems that arise when using wavelets in numerical analysis applications. Section 2 adresses constructions of quadrature formulae to approximate the inner products $\langle f(x), \tilde{\varphi}(2^j x \Leftrightarrow l) \rangle$. We discuss and compare the accuracy of different quadrature formulae for several types of wavelets.

In section 3 we develop pointwise asymptotic expansions for the error $f(x) \Leftrightarrow \mathcal{P}_n f(x)$ in powers of h, with $h = 2^{-n}$. We show how the expansion can be used to derive asymptotic interpolation properties and to accelerate the convergence. Also we point out a simple trick to improve the accuracy of the first term of the expansion.

2. Quadrature formulae.

2.1. General idea. The idea of a quadrature formula is to find *weights* w_k and *abscissae* x_k such that

$$\int_{-\infty}^{+\infty} f(x) \varphi(x) \, dx \;\approx\; Q[f(x)] = \sum_{k=0}^{r} w_k \, f(x_k). \tag{2.1}$$

DEFINITION 2.1. *The* degree of accuracy *of a quadrature formula is q if it yields the exact result for every polynomial of degree less than or equal to q.*

The degree of accuracy determines the convergence order as follows: if $f(x)$ belongs to $\mathcal{C}^{q+1}$, then

$$\frac{\nu_{n,l} \Leftrightarrow 2^{-n/2}\, Q[f(2^{-n}x)]}{\nu_{n,l}} = \mathcal{O}(h^{q+1}) \quad \text{with} \quad h = 2^{-n}. \tag{2.2}$$

This can easily be seen using the Taylor expansion. It also follows from Peano's theorem in [10]. Note that we do not impose any regularity conditions on $\varphi(x)$. The number of abscissae r determines the efficiency of a quadrature formula since the number of function evaluations and algebraic operations is proportional to r. The quadrature formula is usually constructed by demanding that

$$Q[x^i] = \mathcal{M}_i \quad \text{for} \quad 0 \le i \le q,$$

which leads to an algebraic system. In case the abscissae x_k are fixed, this system is linear in the unknowns w_k. More efficient quadrature formula can be constructed by also treating the abscissae as unknowns, cf. Gauss quadrature formulae.

In connection with a multiresolution analysis, the quadrature formula will be used at a fixed, finest level n to approximate the $\nu_{n,l}$. The coefficients on the coarser levels $\nu_{j,l}$ and $\mu_{j,l}$ with $j < n$, can then be calculated using the fast wavelet transform. Note that this implies that the error will be $\mathcal{O}(2^{-n(q+1)})$ for all coefficients, independent of their level j.

The abscissae should be chosen equidistant because in some applications the function $f(x)$ is only known at equidistant abscissae and because then quadrature formulae for neighboring coefficients can share common points. Therefore we let the abscissae be of the form $k2^{-s} + \tau$, where τ is an unknown.

Comparing equations (1.11) and (2.2), we see that the degree of accuracy should be at least equal to $N \Leftrightarrow 1$, otherwise the quadrature formula will ruin the convergence order of the wavelet approximation.

2.2. Trapezoidal rule. A simple quadrature formula is the *trapezoidal rule*, where

$$Q[f(x)] = \sum_k \varphi(k)\, f(k). \tag{2.3}$$

In general the application of this rule is limited because it only has a degree of accuracy equal to 1. Here however, the following preposition holds:

PREPOSITION 2.2. *If the scaling function satisfies the Strang-Fix condition (1.8), the degree of accuracy of the trapezoidal rule (2.3) is equal to $N \Leftrightarrow 1$.*

This easily seen from equation (1.9) for $x = 0$. The trapezoidal rule can be used in a multiresolution analysis. However in general it is not very efficient. In case $\varphi(x)$ is not compactly supported, the sum in (2.3) has to be broken of which usually leads to a large number of abscissae. But also when φ is compactly supported, this formula is not really efficient: the Daubechies' orthogonal scaling functions have a support length of $2N \Leftrightarrow 1$, such that $r = 2N \Leftrightarrow 2 = 2q$ while even with fixed abscissae we can achieve $r = q + 1$. Only in the case of cardinal B-splines, the trapezoidal rule is useful because here $r = N \Leftrightarrow 1 = q$.

2.3. One point formulae. Since the integral of the scaling function is 1, we can write a one point formula as $Q[f(x)] = f(x_1)$. Evidently, if $x_1 = \mathcal{M}_1$, the degree of accuracy is equal to 1. In the case of orthogonal wavelets, the following theorem holds:

THEOREM 2.3. *If $\varphi(x)$ is an orthogonal scaling function with $N > 1$, then $\mathcal{M}_2 = \mathcal{M}_1^2$.*

Proof. Define

$$\kappa_m = \langle\, x, \varphi(x)\, \varphi(x \Leftrightarrow m) \,\rangle\,.$$

Because of the orthogonality holds that

$$\kappa_{-m} = \langle\, x \Leftrightarrow m, \varphi(x \Leftrightarrow m)\, \varphi(x) \,\rangle = \kappa_m.$$

Consequently

$$0 = \sum_m m\, \kappa_m = \langle\, x, \varphi(x) \sum_m m\, \varphi(x \Leftrightarrow m) \,\rangle\,,$$

and, since $N > 1$,

$$\sum_m m\, \varphi(x \Leftrightarrow m) = x \Leftrightarrow \mathcal{M}_1.$$

Combining the last two equations yields $\mathcal{M}_2 \Leftrightarrow \mathcal{M}_1^2 = 0$. □

Note: This theorem was proven independently in case of Daubechies' scaling functions in [14].

This means that for orthogonal scaling functions the degree of accuracy of the one point quadrature is 2. Consequently it can be used in case $N \leq 3$.

In [8, 9] Ingrid Daubechies constructed orthogonal scaling functions with compact support that have $N \Leftrightarrow 1$ vanishing moments,

$$\mathcal{M}_p = 0 \quad \text{for} \quad 1 \leq p < N, \tag{2.4}$$

where again N is the number of vanishing wavelet moments. These wavelets were called *coiflets* after Ronald Coifman who asked for their construction. We see from (1.9) that they also satisfy

$$\sum_k k^p \varphi(k) = \delta_p \quad \text{for} \quad 0 \leq p < N. \tag{2.5}$$

In this case the one point quadrature formula with $x_1 = 0$ immediately has a degree of accuracy of $N \Leftrightarrow 1$. This formula was used in numerical analysis applications in [3]. The coiflets have the disadvantage that their support width is $3N \Leftrightarrow 1$. This results in a 50% increase in computational cost for the fast wavelet transform in comparison with the original Daubechies' wavelets with the same number of vanishing moments whose support width is only $2N \Leftrightarrow 1$.

2.4. Practical aspects. In applications such as signal and image processing, usually discrete samples a_l are given. Then there are several ways to start the multiresolution analysis. First one can construct a function $a(x) \in V_n$ [17],

$$a(x) = \sqrt{h} \sum_l a_l\, \varphi_{n,l}(x) \quad \text{with} \quad h = 2^{-n}.$$

We can see that the continuous function $a(x)$ will in a way "follow" the discrete samples a_l. The quadrature formula can help us to find a relationship between the function $a(x)$ and the discrete samples a_i. Indeed, using the biorthogonal notation,

$$\sqrt{h}\, a_l = \langle\, a, \tilde{\varphi}_{n,l}\,\rangle \quad \text{and} \quad \langle\, a, \tilde{\varphi}_{n,l}\,\rangle = \sqrt{h}\,[a(h\,(\mathcal{M}_1 + l)) + \mathcal{O}(h^t)],$$

so

$$a_l = a(h\,(\mathcal{M}_1 + l)) + \mathcal{O}(h^t).$$

This means that $a(x)$ satisfies a quasi-interpolating property. Here $t = 2$ in general, $t = 3$ for orthogonal wavelets and $t = N$ for coiflets.

Secondly, one can consider the samples a_l as function evaluations, $a_l = f(hl)$. This corresponds to the one point quadrature formula with $x_1 = 0$. Then the following theorem is important:

THEOREM 2.4. *If $f(x) \in \mathcal{C}^N$ with $f^{(i)}(x)$ bounded for $i \le N$, then ($h = 2^{-n}$)*

$$\sum_l f(hl)\,\varphi(2^n x \Leftrightarrow l) = \sum_l \varphi(l)\, f(x \Leftrightarrow hl) + \mathcal{O}(h^N).$$

Proof.

$$\begin{aligned}
\sum_l f(hl)\,\varphi(2^n x \Leftrightarrow l) &= \sum_{i=0}^{N-1} f^{(i)}(x)\frac{(\Leftrightarrow h)^i}{i!}\sum_l (2^n x \Leftrightarrow l)^i\,\varphi(2^n x \Leftrightarrow l) + \mathcal{O}(h^N) \\
&= \sum_{i=0}^{N-1} f^{(i)}(x)\frac{(\Leftrightarrow h)^i}{i!}\sum_l l^i \varphi(l) + \mathcal{O}(h^N) \\
&= \sum_l \varphi(l)\, f(x \Leftrightarrow hl) + \mathcal{O}(h^N)
\end{aligned}$$

□

This theorem states that taking function evaluations as coefficients results in approximating a different function $\tilde{f}_n(x) = \sum_l \varphi(l)\, f(x \Leftrightarrow hl)$ with an error of $\mathcal{O}(h^N)$. This function can be seen as a "blurred" version of $f(x)$ as $\varphi(l)$ is a low pass filter. Now, $\tilde{f}_n(x)$ will converge to $f(x)$ for $n \to \infty$ since $\sum_k \varphi(k) = 1$. However, in general this convergence is only $\mathcal{O}(h)$. In the case of the coiflets we see from (2.5) that $\tilde{f}_n(x) = f(x) + \mathcal{O}(h^N)$.

Finally, one can consider the samples a_l as inner products,

$$a_l = 2^n \,\langle\, f(x), \varphi^{(0)}(2^n x \Leftrightarrow l)\,\rangle\,,$$

where $\varphi^{(0)}(x)$ is the box function $\chi_{[0,1]}$. One can then construct a recursive multiresolution scheme as follows:

$$\nu_{n,l}^{(0)} = 2^{-n/2}\, a_l \quad \text{and} \quad \nu_{j-1,l}^{(n-j+1)} = \sqrt{2}\sum_k h_{k-2l}\,\nu_{j,k}^{(n-j)}.$$

Here is

$$\nu_{j,l}^{(m)} = 2^{j/2}\,\langle\, f(x), \varphi^{(m)}(2^j x \Leftrightarrow l)\,\rangle\,, \quad \text{with} \quad \varphi^{(m)}(x) = T^m \varphi^{(0)}(x),$$

and T an operator defined as

$$Tg(x) = 2 \sum_k h_k \, g(2x \Leftrightarrow k).$$

It is proven in [7] that $\lim_{m \to \infty} \varphi^{(m)}(x) = \varphi(x)$. This scheme is applicable for a wide range of functions $\varphi^{(0)}(x)$ with integral 1.

2.5. Multiple point formula.

2.5.1. Construction scheme. Since the degree of accuracy of a one point formulae is limited, we construct multiple point integration formulae. In this section we assume that $\varphi(x)$ has compact support $[0, L]$, and satisfies a refinement equation (1.1) with $L+1$ non-zero coefficients h_k. Although the construction is general, we focus on scaling functions with compact support since in this case we have the extra limitation that the abscissae should fall inside the integration interval. We construct an r point quadrature formula with $x_k = d_k \Leftrightarrow \tau$, $d_k = (k \Leftrightarrow 1)2^s$ and $(r \Leftrightarrow 1)2^s \Leftrightarrow L \leq \tau \leq 0$. The range of the shift τ is determined by the requirement that no abscissae should fall outside the integration interval. In order to have a non-zero range for the shift τ, the parameters r and s should be chosen such that $(r \Leftrightarrow 1)2^s < L$. This technique to construct quadrature formulae is also used in [3] but there the shift τ is given a fixed value.

Since there are $r+1$ unknowns $\{\tau, w_1, \ldots, w_r\}$, one can try to achieve a degree of accuracy r. This results in the following system which is nonlinear in the unknown τ,

$$\sum_{k=1}^{r} w_k \, [d_k \Leftrightarrow \tau]^i = \mathcal{M}_i \qquad 0 \leq i \leq r. \tag{2.6}$$

The value of the shift τ can be determined using the product polynomial $\Pi(x)$. This polynomial is defined as

$$\Pi(x) = \prod_{k=1}^{r} (x \Leftrightarrow x_k) = \prod_{k=1}^{r} (x + \tau \Leftrightarrow d_k) = \sum_{i=0}^{r} p_i(\tau) x^i,$$

where $p_i(\tau)$ is a polynomial of degree $r \Leftrightarrow i$. Since the degree of accuracy is r, the quadrature formula gives the exact result for the product polynomial $\Pi(x)$ so

$$0 = Q_r[\Pi(x)] = \sum_{i=0}^{r} p_i(\tau) \, \mathcal{M}_i = , \, (\tau).$$

The latter expression is a polynomial of degree r in τ. For the quadrature formula to exist, , (τ) must have a root in the interval $[(r \Leftrightarrow 1)2^s \Leftrightarrow L, 0]$. However, the existence of such a root is not theoretically guaranteed. If there is no root in this interval, an arbitrary value for τ must be chosen and one degree of accuracy is lost. Once τ is determined, the weights are the solution of the linear system formed by r equations of (2.6). In order to construct , (τ) we write

$$p_i(\tau) = \sum_{j=0}^{r-i} p_{i,j} \, \tau^j \quad \text{and} \quad , \, (\tau) = \sum_{j=0}^{r} \left(\sum_{i=0}^{r-j} \mathcal{M}_i \, p_{i,j} \right) \tau^j.$$

The coefficients $p_{i,j}$ are symmetric ($p_{i,j} = p_{j,i}$) since the product polynomial is symmetric in τ and x and can be found as $p_{i,j}^{(r)}$ where

$$\Pi^{(m)}(x) = \prod_{k=1}^{m} (x + \tau \Leftrightarrow d_k) = \sum_{i=0}^{m} \sum_{j=0}^{m-i} p_{i,j}^{(m)} \tau^j x^i.$$

An algorithm to calculate the $p_{i,j}$ can be derived by writing

$$\Pi^{(m)}(x) = (x + \tau \Leftrightarrow d_m)\, \Pi^{(m-1)}(x),$$

and identifying the coefficients of the powers of x and τ. A disadvantage of this construction is that the system of equations (2.6) is ill-conditioned if r is large. In the construction of Q_{13} for the Daubechies scaling function with $N = 7$, the condition number of the linear system for the weights is 5.10^{16}!

2.5.2. Modified construction. The ill-conditioning problem of the previous section can be overcome if we use the basis of Chebyshev polynomials. This technique is also used successfully in [12, 19]. The Chebyshev polynomial $T_n(x)$ of degree n is defined by $T_0(x) = 1$, $T_1(x) = x$ and $T_n(x) = 2\,x\,T_{n-1}(x) \Leftrightarrow T_{n-2}(x)$ for $n > 1$ [1]. Since the interesting properties of these polynomials only hold in the interval $[\Leftrightarrow 1, 1]$, we transform the scaling function $\varphi(x)$ to this interval yielding a function $\varphi^*(y)$. We will use the notation y to indicate an independent variable that varies between $\Leftrightarrow 1$ and 1,

$$2\varphi^*(y) = L\varphi(x) \quad \text{with} \quad 2x = L(y+1).$$

The refinement equation (1.1) becomes

$$\varphi^*(y) = 2 \sum_k h_k\, \varphi^*(2y \Leftrightarrow 2k/L + 1).$$

We construct a quadrature formula,

$$\int_{-1}^{1} \varphi^*(y) f\left(\frac{L(y+1)}{2}\right) dy = \int_{-1}^{1} \varphi^*(y) f^*(y) dy \;\approx\; \sum_{k=1}^{r} w_k f^*(y_k) = Q_r[f(x)],$$

with $y_k = d_k^* \Leftrightarrow \tau^*$, $d_k^* = 2d_k/L \Leftrightarrow 1$, and $\tau^* = 2\tau/L$. Let $\mathcal{M}_p^*$ denote the modified moments,

$$\mathcal{M}_p^* = \int_0^1 T_p(x)\, \varphi^*(x)\, dx.$$

The new system can be written as

$$\sum_{k=1}^{r} w_k\, T_i(d_k^* \Leftrightarrow \tau^*) = \mathcal{M}_i^* \qquad 0 \le i \le r. \tag{2.7}$$

The solution procedure is similar to the one in the previous section. We construct a polynomial , $^*(\tau^*)$, written as a linear combination of Chebyshev polynomials, and try to find a root in the appropriate interval. In order to construct , $^*(\tau^*)$ we write

$$\Pi(y) = 2^{-(r-1)} \sum_{i=0}^{r} \sum_{j=0}^{r-i} q_{i,j}\, T_j(\tau^*)\, T_i(y)$$

and

$$, ^*(\tau^*) = 2^{-(r-1)} \sum_{j=0}^{r} \sum_{i=0}^{r-j} q_{i,j}\, \mathcal{M}_i^*\, T_j(\tau^*).$$

Now let

$$2^{(m-1)}\, \Pi^{(m)}(y) = 2^{(m-1)} \prod_{k=1}^{m} (y + \tau^* \Leftrightarrow d_k^*) = \sum_{i=0}^{m} \sum_{j=0}^{m-i} q_{i,j}^{(m)}\, T_j(\tau^*)\, T_i(y) \tag{2.8}$$

and

$$\begin{aligned} 2^{(m-1)}\, \Pi^{(m)}(y) &= 2 \sum_{i=0}^{m-1} \sum_{j=0}^{m-i-1} q_{i,j}^{(m-1)}\, (y + \tau^* \Leftrightarrow d_m)\, T_j(\tau^*)\, T_i(y) \\ &= \sum_{i=1}^{m} \sum_{j=0}^{m-i} q_{i-1,j}^{(m-1)}\, T_j(\tau^*)\, T_i(y) + \sum_{i=-1}^{m-2} \sum_{j=0}^{m-2-i} q_{i+1,j}^{(m-1)}\, T_j(\tau^*)\, T_{|i|}(y) \\ &\quad + \sum_{i=0}^{m-1} \sum_{j=1}^{m-i} q_{i,j-1}^{(m-1)}\, T_j(\tau^*)\, T_i(y) + \sum_{i=0}^{m-1} \sum_{j=-1}^{m-2-i} q_{i,j+1}^{(m-1)}\, T_{|j|}(\tau^*)\, T_i(y) \\ &\Leftrightarrow 2\, d_m \sum_{i=0}^{m-1} \sum_{j=0}^{m-1-i} q_{i,j}^{(m-1)}\, T_j(\tau^*)\, T_i(y). \end{aligned} \tag{2.9}$$

An algorithm for the calculation of the $q_{i,j} = q_{i,j}^{(r)}$ can be found by identifying the coefficients of the Chebyshev polynomials of equal degree in x and in τ^* in (2.8) and (2.9). It is given in appendix A.

The condition number of the system for the construction of the same Q_{13} formula as in the previous section is now 1.10^3 ! The roots of the polynomial $, (\tau^*)$ can be found as the eigenvalues of its Chebyshev companion matrix. The effects of an orthogonal basis on the condition of the roots of a polynomial is discussed in [13]. It is stated there that the interval of orthogonality should contain the roots of interest. This condition is satisfied in most cases here.

2.5.3. Calculation of the modified moments.. It is possible to calculate the modified moment as a linear combination of the monomial moments using the coefficients of the Chebyshev polynomials. However, a considerable loss of significant digits will occur since these coefficients tend to be large and different in sign. The condition would essentially be as bad as in the construction of the previous section. We need a formula to calculate the modified moments directly. We know that

$$\mathcal{M}_p^* = \sum_k h_k \int_{-1}^{1} T_p\left(\frac{u \Leftrightarrow 1 + 2k/L}{2}\right) \varphi^*(u)\, du.$$

In order to find a recursion formula, we write this last, shifted and dilated Chebyshev polynomial as a sum of Chebyshev polynomials of degree less than or equal to p,

$$T_p\left(\frac{y \Leftrightarrow 1 + 2k/L}{2}\right) = 2^{-p} \sum_{i=0}^{p} w_i^{(p)}(k)\, T_i(y),$$

TABLE 2.1
Errors of the integration rules.

n	5.2^n	Trapezoidal rule	One point formula	Q_5^*	Q_5	Q_{10}
0	5	7.08e-04	1.17e-02	6.13e-04	2.15e-03	-
1	10	4.17e-03	1.43e-03	9.78e-05	4.40e-05	1.03e-08
2	20	7.96e-04	1.76e-04	4.30e-06	6.51e-07	1.11e-12
3	40	1.15e-04	2.19e-05	1.52e-07	9.38e-09	4.21e-15
4	80	1.53e-05	2.74e-06	5.03e-09	1.38e-10	9.99e-16
5	160	1.98e-06	3.43e-07	1.61e-10	2.09e-12	-
6	320	2.50e-07	4.28e-08	5.10e-12	3.19e-14	-
7	640	3.15e-08	5.35e-09	1.60e-13	1.11e-16	-
8	1280	3.96e-09	6.69e-10	4.66e-15	-	-
9	2560	4.96e-10	8.37e-11	2.22e-16	-	-
10	5120	6.20e-11	1.04e-11	-	-	-

such that

$$\mathcal{M}_p^* = \frac{1}{2^p \Leftrightarrow 1} \sum_{i=0}^{p-1} \left(\sum_{k=0}^{L} h_k w_i^{(p)}(k) \right) \mathcal{M}_i^*.$$

Numerical tests show this to be a stable recursion formula. The $w_i^{(p)}(k)$ can be calculated recursively. We will use the notation $w_i^{(p)} = w_i^{(p)}(k)$ and $\lambda = 2k/L \Leftrightarrow 1$ for simplicity here. Now

$$T_{p+1}\left(\frac{y+\lambda}{2}\right) = 2^{-(p+1)} \sum_{i=0}^{p+1} w_i^{(p+1)}\, T_i(y) \tag{2.10}$$

and

$$\begin{aligned} T_{p+1}\left(\frac{y+\lambda}{2}\right) &= (y+\lambda)\, T_p\left(\frac{y+\lambda}{2}\right) \Leftrightarrow T_{p-1}\left(\frac{y+\lambda}{2}\right) \\ &= 2^{-(p+1)} \left(\sum_{i=1}^{p+1} w_{i-1}^{(p)}\, T_i(y) + \sum_{i=-1}^{p-1} w_{i+1}^{(p)}\, T_{|i|}(y) \right. \\ &\qquad \left. + 2\lambda \sum_{i=0}^{p} w_i^{(p)}\, T_i(y) \Leftrightarrow 4 \sum_{i=0}^{p-1} w_i^{(p-1)}\, T_i(y) \right) \end{aligned} \tag{2.11}$$

The algorithm can be found by identifying the coefficients of the Chebyshev polynomials of equal degree in (2.10) and (2.11). It is given in appendix A.

2.6. Numerical results. When calculating the coefficients at a level n with a quadrature formula, one usually wants to avoid evaluating (or "sampling") the function $f(x)$ at abscissae with spacing smaller then 2^{-n}. This means that $s \geq 0$. For a certain r, one wants the maximal s so the abscissae spread out over the whole integration interval. The maximal s, within the requirement that $(r \Leftrightarrow 1)2^s < L$, however corresponds to the smallest admittance interval for τ. As mentioned above, there is no theoretical certainty that a quadrature formula with degree of accuracy r exists. If for a formula with $s > 0$ no τ can be found, one can always try to find a formula with spacing 2^{s-1}.

We constructed quadrature formulae for the Daubechies' orthogonal scaling functions and verified that formulae Q_r with $2 \leq r \leq 2N \Leftrightarrow 1$ exist for $2 \leq N \leq 10$. For $r = 2$, one of the weights is always zero and we return to the one point formula. In 9 of these 90 formulae, no τ could be found for the maximal value of s and s was taken one smaller. This was in most cases for $r = N$. The weights of these formulae can vary in sign, but their absolute value does not grow too large when the number of points increases.

We compare now different quadrature formulae in a practical example. We construct several multiresolution trees, each with coarsest level 0 and finest level n, and this for several n. We compare each time $\nu_{0,0}$. Notice however that the error is of the same order in the whole multiresolution tree.

As an example we take for $\varphi(x)$ the Daubechies orthogonal scaling function with $N = 3$, $f(x) = \sin(x)$ and

$$\nu_{0,0} = \int_0^5 \varphi(x) \sin(x) dx \approx 0.741104421925905 \tag{2.12}$$

We compare the one point formula, Q_5, Q_{10} (with $s = \Leftrightarrow 1$ applied at level $n \Leftrightarrow 1$), and the trapezoidal rule. The total number of evaluations is then respectively $5.2^n \Leftrightarrow 4$, 5.2^n, 5.2^n and $5.2^n \Leftrightarrow 1$. We also use a formula Q_5^* were τ was given a fixed value equal to $\Leftrightarrow 1/2$. The results are given in table 2.1. They show that for sufficiently differentiable functions $f(x)$, it is useful to search for the optimal value of the shift τ.

3. Asymptotic error expansions.

3.1. General idea. In the second part of the paper we study pointwise asymptotic error expansions. Define therefore the error of the wavelet approximation as

$$\mathcal{E}_n f(x) = f(x) \Leftrightarrow \mathcal{P}_n f(x),$$

and note that

$$\mathcal{E}_n f(x) = \sum_{j=0}^{\infty} \mathcal{Q}_{n+i} f(x). \tag{3.1}$$

The general idea is to first construct an expansion for $\mathcal{Q}_n$ and then apply this formula to find the expansion for $\mathcal{E}_n$. For simplicity we assume that the wavelets are orthogonal and compactly supported.

3.2. Construction of error expansions. We know

$$\mathcal{Q}_n f(x) = \sum_l \mu_{n,l}\, \psi(2^n x \Leftrightarrow l), \quad \text{with} \quad \mu_{n,l} = 2^n \langle f(y), \psi(2^n y \Leftrightarrow l) \rangle \tag{3.2}$$

We suppose that $f(x) \in \mathcal{C}^{N+1}$ with $f^{(l)}(x)$ bounded for $l \leq N + 1$. In the sequel we will always have $h = 2^{-n}$. We also suppose that the support of $\psi(x)$ is $[A, B]$ with $A, B \in \mathbf{Z}$. In formula (3.2) both x and y belong to $I_{n,l} = [h(A + l), h(B + l)]$ and consequently $|x \Leftrightarrow y| \leq Lh$ with $L = B \Leftrightarrow A$. We also suppose that $AB \leq 0$ such that $hl \in I_{n,l}$. For n fixed, at most $L + 1$ functions $\psi_{n,l}(x)$ are non-zero at x, namely the ones with $2^n x \Leftrightarrow B \leq l \leq 2^n x \Leftrightarrow A$. We can follow two strategies. First, using the Taylor formula around hl, we can write

$$\mu_{n,l} \quad = \quad \langle f(hz + hl), \psi(z) \rangle$$

$$
\begin{aligned}
&= \langle \sum_{p=0}^{M+N} h^p f^{(p)}(hl)\,\frac{z^p}{p!} + h^{N+M+1} f^{(N+M+1)}(\xi)\,\frac{z^{N+M+1}}{(N+M+1)!}, \psi(z)\rangle \\
&\quad \text{with } \xi \text{ between } hl \text{ and } hl+hz \\
&= \sum_{p=N}^{N+M} f^{(p)}(hl)\frac{\mathcal{N}_p}{p!} + \rho_{n,l}\, h^{N+M+1},
\end{aligned}
$$

with

$$
|\rho_{n,l}| \;\le\; \max_{\xi\in I_{n,l}} |f^{(N+M+1)}(\xi)|\,\Psi_{N+M+1}.
$$

Here Ψ_j is defined as

$$
\Psi_j = \frac{1}{j!}\int_A^B |z|^j\,|\psi(z)|\,dz.
$$

Now

$$
\mathcal{Q}_n f(x) = \sum_{p=N}^{N+M} \frac{h^p\,\mathcal{N}_p}{p!} \sum_l f^{(p)}(hl)\,\psi(2^n x \Leftrightarrow l) + K_n\, h^{N+M+1}, \tag{3.3}
$$

with

$$
|K_n| \;\le\; (L+1) \max_{|\xi-x|\le L/2^n} |f^{(N+M+1)}(\xi)|\,\Psi_{N+M+1}\,\Psi_0. \tag{3.4}
$$

Consequently

$$
\mathcal{Q}_{n+i} f(x) = \sum_{p=N}^{N+M} \frac{h^p\,\mathcal{N}_p}{p!} \sum_l \frac{f^{(p)}(hl/2^i)}{2^{ip}}\,\psi(2^{n+i}x \Leftrightarrow l) + \frac{K_{n+i}}{2^{i(N+M+1)}} h^{N+M+1}.
$$

Since the upper bound (3.4) of $|K_{n+i}|$ cannot grow as i increases, we still have a $\mathcal{O}(h^{N+M+1})$ term if we sum the $\mathcal{Q}_{n+i}f(x)$ and thus,

$$
\boxed{\mathcal{E}_n f(x) = \sum_{p=N}^{N+M} \frac{h^p\,\mathcal{N}_p}{p!} \sum_{i=0}^{\infty}\sum_l \frac{f^{(p)}(hl/2^i)}{2^{ip}}\,\psi(2^{n+i}x \Leftrightarrow l) + \mathcal{O}(h^{N+M+1}).} \tag{3.5}
$$

The advantage of this formula is that the contributions of each subspace W_{n+i} can be distinguished. The disadvantage is that, because of the double summation, it is not very practical to work with. Therefore we now derive a second formula using the Taylor formula around $y = x$ in (3.2),

$$
\begin{aligned}
\mu_{n,l} &= 2^n \langle \sum_{p=N}^{N+M} f^{(p)}(x)\,\frac{(y\Leftrightarrow x)^p}{p!} + f^{(N+M+1)}(\xi)\frac{(y\Leftrightarrow x)^{N+M+1}}{(N+M+1)!}, \psi(2^n y \Leftrightarrow l)\rangle \\
&\quad \text{with } \xi \text{ between } x \text{ and } y \\
&= \sum_{p=N}^{N+M} \frac{2^n f^{(p)}(x)}{p!}\,\langle (y\Leftrightarrow x)^p, \psi(2^n y\Leftrightarrow l)\rangle + \rho_{n,l}\, h^{N+M+1},
\end{aligned}
$$

with

$$|\rho_{n,l}| \;\leq\; \frac{L^{N+M+1}}{(N+M+1)!} \max_{\xi \in I_{n,l}} |f^{(N+M+1)}(\xi)|\, \Psi_0.$$

Now

$$\begin{aligned} 2^n \left\langle (y \Leftrightarrow x)^p, \psi(2^n y \Leftrightarrow l) \right\rangle &= \left\langle (hz + hl \Leftrightarrow x)^p, \psi(z) \right\rangle \\ &= h^p \sum_{j=0}^{p-N} \binom{p}{j} \mathcal{N}_{p-j}\, (l \Leftrightarrow 2^n x)^j. \end{aligned}$$

Thus:

$$\mu_{n,l} = \sum_{p=N}^{N+M} \frac{h^p\, f^{(p)}(x)}{p!} \sum_{j=0}^{p-N} \binom{p}{j} \mathcal{N}_{p-j}\, (l \Leftrightarrow 2^n x)^j + \rho_{n,l}\, h^{N+M+1},$$

and

$$\mathcal{Q}_n f(x) = \sum_{p=N}^{N+M} h^p\, f^{(p)}(x) \sum_{j=0}^{p-N} \frac{\mathcal{N}_{p-j}}{(p \Leftrightarrow j)!\, j!} (\Leftrightarrow)^j\, \sigma_j(2^n x) + K_n\, h^{N+M+1},$$

with $\sigma_j(x) \in \mathrm{L}^2([0,1])$ defined as

$$\sigma_j(x) = \sum_l (x \Leftrightarrow l)^j\, \psi(x \Leftrightarrow l),$$

and

$$|K_n| \;\leq\; (L+1)\, \frac{L^{N+M+1}}{(N+M+1)!} \max_{|\xi - x| \leq L/2^n} |f^{(N+M+1)}(\xi)|\, \Psi_0^2. \tag{3.6}$$

We can write this as

$$\mathcal{Q}_n f(x) = \sum_{q=0}^{M} h^{N+q}\, f^{(N+q)}(x)\, \sigma_q^*(2^n x) + K_n\, h^{N+M+1}, \tag{3.7}$$

with

$$\sigma_q^*(x) = \sum_{j=0}^{q} \frac{\mathcal{N}_{N+q-j}}{(N+q \Leftrightarrow j)!\, j!} (\Leftrightarrow)^j\, \sigma_j(x). \tag{3.8}$$

Again we can sum the contributions all up such that,

$$\boxed{\mathcal{E}_n f(x) = \sum_{q=0}^{M} h^{N+q}\, f^{(N+q)}(x)\, \tau_q(2^n x) + \mathcal{O}(h^{N+M+1}),} \tag{3.9}$$

with $\tau_q(x)$ defined as the limit function of a uniformly convergent series,

$$\tau_q(x) = \sum_{i=0}^{\infty} \frac{\sigma_q^*(2^i x)}{2^{i(N+q)}}. \tag{3.10}$$

This formula is more practical to work with but has the disadvantage that one cannot distinguish the contributions of the different wavelet subspaces any more. The projection $\mathcal{Q}_n f(x)$ belongs by definition to W_n. However, in formula (3.7) the first term doesn't belong completely to W_n. We can understand that for sufficiently smooth $f(x)$ and large n it has a big component in this space. As we will see, also the $\mathcal{O}(h^{N+1})$ term of (3.7) has a component in W_n which can make the first term a bad estimate of $\mathcal{Q}_n f(x)$ and consequently the first term of (3.9) a bad first estimate of $\mathcal{E}_n f(x)$. We will come back to this later.

3.3. Interpolation. It is easy to see from their definition that $\sigma_p(x)$, $\sigma_p^*(x)$, and $\tau(x)$ are one periodic functions. So the general term consists of a power of h, the same order of derivative of $f(x)$ and a highly oscillating factor. In this section we take a closer look at the first term of the expansion. Therefore we first derive some properties of $\sigma_0(x)$ and $\tau_0(x)$. We use the notation $\sigma(x) = \sigma_0(x)$ and $\tau(x) = \tau_0(x)\, N!/\mathcal{N}_N$. Using equation (1.2) we can obtain the following relations:

$$\sigma(x) = \sum_l (\Leftrightarrow 1)^l \varphi(2x \Leftrightarrow l) = 2 \sum_l \varphi(2x \Leftrightarrow 2l) \Leftrightarrow 1,$$

$$\sigma(x + 1/2) = \Leftrightarrow \sigma(x), \quad \text{and} \quad \tau(x + 1/2) = \tau(x) \Leftrightarrow 2\,\sigma(x). \tag{3.11}$$

Using Poisson's summation formula the Fourier series of $\sigma(x)$ can be written as

$$\sigma(x) = \sum_{k=-\infty}^{+\infty} \hat{\psi}(2k\pi)\, \exp(ik2\pi x).$$

In the Haar case, when $\varphi(x) = \chi_{[0,1]}(x)$, $\sigma(x)\, \chi_{[0,1]}(x)$ is the Haar wavelet $\varphi(2x) \Leftrightarrow \varphi(2x \Leftrightarrow 1)$ and $\tau(x)$ is a sawtooth function with $\tau(x) = 2 \Leftrightarrow 4x$ for $0 \leq x < 1$. The $\sigma(2^j x)$ functions with $j \geq 0$ are then called the Rademacker functions which are well known in probability theory. Figures 3.1 and 3.2 show $\tau(x)$ and $\sigma(x)$ corresponding to the Daubechies wavelets for different values of N. For $N > 5$, the plots of $\sigma(x)$ and $\tau(x)$ almost coincide visually. For $N = 10$, they look like a shifted sine function. This makes sense because the smoother $\psi(x)$, the faster the decay of $\hat{\psi}(\xi)$ and the more $\sigma(x)$ will look like its fundamental frequency component.

THEOREM 3.1. *If $\varphi(x)$ is continuous, then $\sigma(x)$ has at least two zeros in [0,1).*

Proof. If $\varphi(x)$ is continuous, so is $\sigma(x)$. The proof then follows directly from formula (3.11). Moreover, if x_0 is a zero in $[0,1)$ so is $(x_0 + 1/2) \bmod 1$. □

THEOREM 3.2. *If $\varphi(x)$ is continuous and $N > 1$, then $\tau(x)$ has at least two zeros in [0,1).*

Proof. If $\varphi(x)$ is continuous, so is $\tau(x)$ since the series (3.10) converges uniformly. The proof follows from

$$\tau(0) = \tau(1) = \frac{2^N}{2^N \Leftrightarrow 1}\sigma(0) \quad \text{and} \quad \tau(1/2) = \Leftrightarrow \frac{2^N \Leftrightarrow 2}{2^N \Leftrightarrow 1}\sigma(0),$$

which means that $\tau(x)$ has at least two changes of sign in $[0,1]$. □

In the cases we considered, it turns out that $\sigma(x)$ and $\tau(x)$ have exactly two zeros as can be seen in the figures. We will denote the zeros of $\tau(x)$ in $[0,1)$ by x_1 and x_2 with $x_1 < x_2$. In the sequel we will suppose that the conditions of these theorems are satisfied. Note also that

$$\mathcal{N}_N\, \tau(x) = x^N \Leftrightarrow \mathcal{P}_0\, x^N.$$

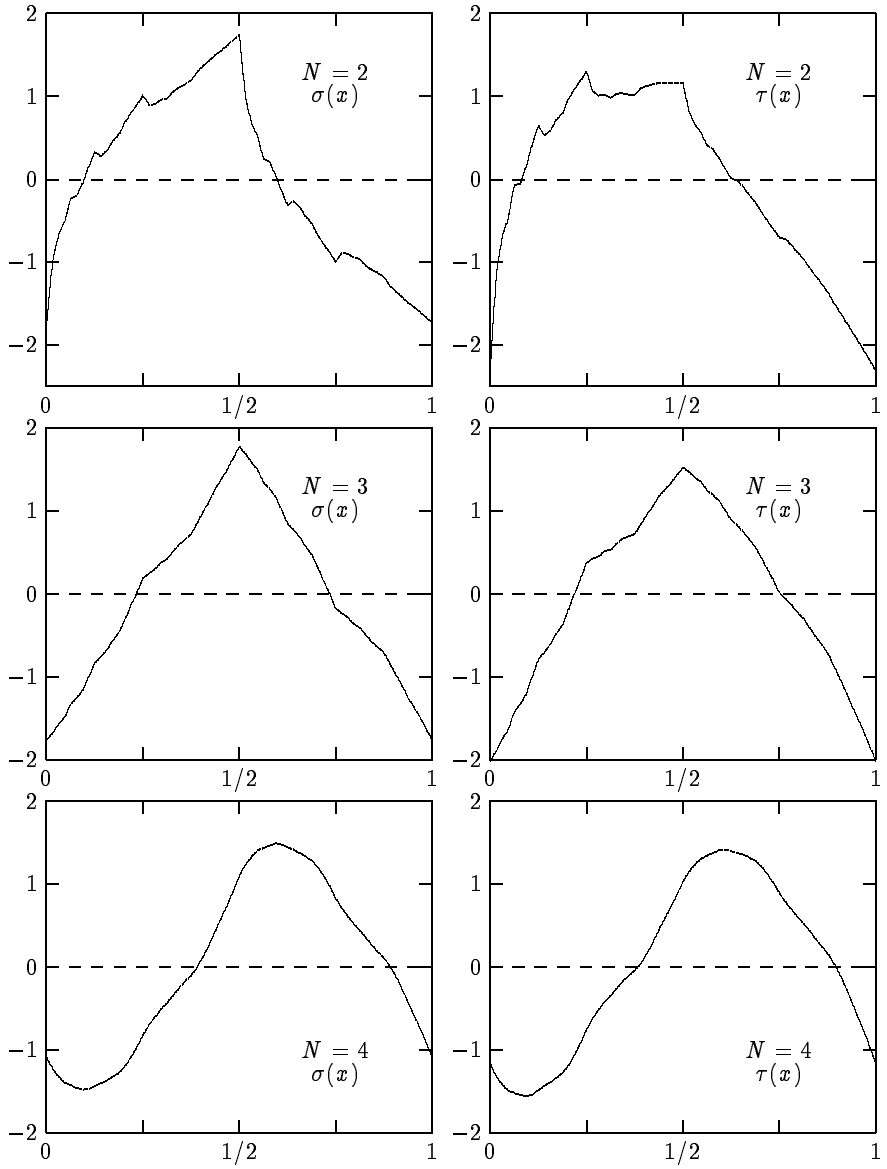

FIG. 3.1. *$\sigma(x)$ and $\tau(x)$ for Daubechies' wavelets with $N = 2, 3, 4$.*

This means that in case the scaling functions are spline functions, $\tau(x)$ is a mono-spline. Therefore, we might want to call $\tau(x)$ in general a *mono-wavelet*.

The envelopes of the first term are given by

$$\frac{h^N \mathcal{N}_N}{N!} f^{(N)}(x)\, \tau_{max} \quad \text{and} \quad \frac{h^N \mathcal{N}_N}{N!} f^{(N)}(x)\, \tau_{min}, \tag{3.12}$$

with

$$\tau_{max} = \max_{x \in [0,1]} \tau(x) \quad \text{and} \quad \tau_{min} = \min_{x \in [0,1]} \tau(x).$$

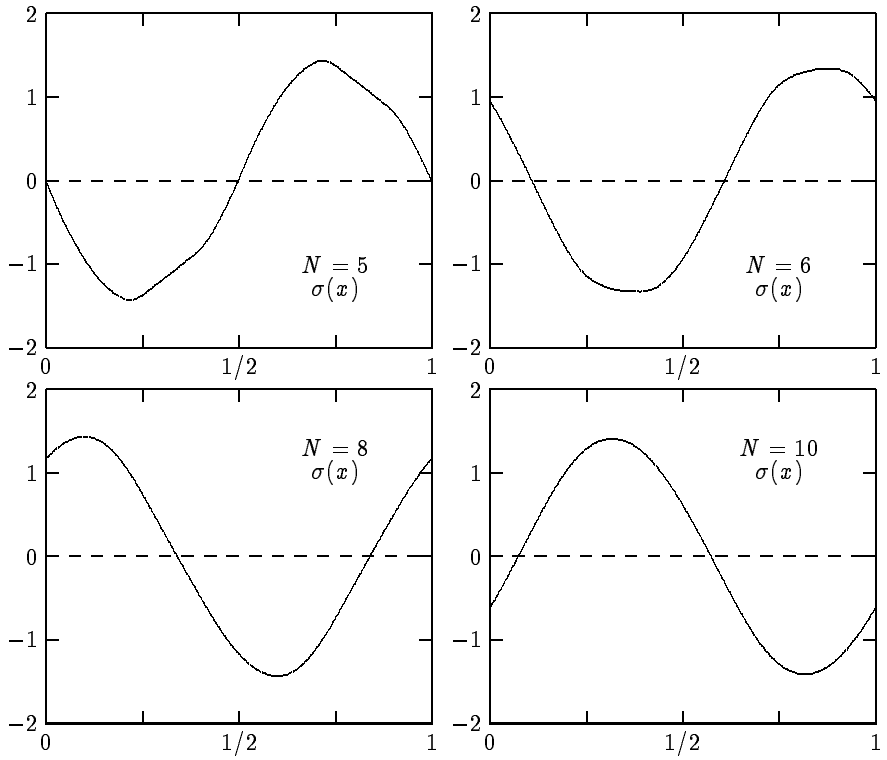

FIG. 3.2. *$\sigma(x)$ for Daubechies' wavelets with $N = 5, 6, 8, 10$.*

Note that $|\tau_{max}| \neq |\tau_{min}|$ so the oscillation is not necessarily balanced around the axis. This difference however becomes smaller for increasing N.

Theorem 3.2 now implies that the first term of formula (3.9) will have at least 2^{n+1} zeros per unit length. For sufficiently small h, the approximation $P_n f(x)$ will thus interpolate the function $f(x)$ also in roughly 2^{n+1} points per unit length. Note that this is about twice the number of basis functions. The interpolation points z_k with $P_n f(z_k) = f(z_k)$ satisfy asymptotically

$$z_{2k} = (x_1 + k)\, h + \mathcal{O}(h^2) \quad \text{and} \quad z_{2k+1} = (x_2 + k)\, h + \mathcal{O}(h^2).$$

3.4. Numerical extrapolation. We can use the the error expansion to accelerate the accuracy of a wavelet series with extrapolation techniques. Indeed, the multiresolution scheme consists of a number of approximations at different levels which can be used to estimate the components of the error expansion and eliminate them. If m is the coarsest level of the multiresolution scheme and $2^m x \in \mathbf{Z}$, the asymptotic error expansion at x, consists of powers of h whose coefficients are independent of h due to the periodicity of $\tau_j(x)$. This means that classical extrapolation techniques such as Richardson extrapolation become applicable.

Table 3.1 shows the results of a numerical experiment where $f(x) = \exp(\Leftrightarrow 20\,(x \Leftrightarrow 0.5)^2)$ and the wavelet used is the Daubechies wavelet with $N = 2$ vanishing moments. The first column shows the the relative error of the approximation at $x = 1/4$ at levels n from 2 to 9. The other columns are the relative errors of the values obtained with

TABLE 3.1
The Richardson extrapolation table for x = 1/4.

n	$\lvert\frac{\mathcal{E}_n f(1/4)}{f(1/4)}\rvert$							
2	1.2e-01	-	-	-	-	-	-	-
3	5.6e-02	3.3e-02	-	-	-	-	-	-
4	1.7e-02	3.5e-03	6.8e-04	-	-	-	-	-
5	4.3e-03	1.9e-04	2.8e-04	2.5e-04	-	-	-	-
6	1.1e-03	2.5e-06	3.0e-05	1.3e-05	5.5e-06	-	-	-
7	2.7e-04	2.3e-06	2.3e-06	4.3e-07	1.4e-08	7.3e-08	-	-
8	6.6e-05	4.2e-07	1.5e-07	1.2e-08	1.3e-09	1.6e-09	9.9e-10	-
9	1.6e-05	6.2e-08	9.9e-09	3.4e-10	3.6e-11	1.6e-11	3.6e-12	3.1e-13

the following Richardson extrapolation scheme:

$$t_{n,1} = P_n f(1/4) \quad \text{for} \quad 2 \le n \le 9,$$

and

$$t_{n,j} = \frac{2^{N+j-2}\, t_{n,j-1} \Leftrightarrow t_{n-1,j-1}}{2^{N+j-2} \Leftrightarrow 1} \quad \text{for} \quad 3 \le n \le 9 \quad \text{and} \quad n \le j \le 9.$$

The first column shows a convergence of h^2, i.e. on each level the error is roughly divided by 4. Every next column corresponds to eliminating one term of the expansion. We see that in this case simple linear combinations result in an increase from 5 to almost 13 accurate digits. This however only works for the points with $2^m x \in \mathbf{Z}$.

In a more general setting, it is also possible to consider f as a distribution. Note that a distribution f is characterized by inner products $\langle f, \eta \rangle$ where η belongs to the class of test functions. If f is a distribution we can approximate the inner products $\langle f, \eta \rangle$ by $\langle P_n f, \eta \rangle$. We can use the asymptotic error expansion of the wavelet approximation to get an asymptotic error expansion for these inner products. For a certain class of test functions this error expansion will consist of powers of h whose coefficients are independent of h. This again will make Richardson extrapolation applicable.

3.5. A more accurate first term. In this section we try to slightly modify the first term of the error expansion (3.9) so that it becomes a more accurate estimate of $\mathcal{E}_n f(x)$. Therefore we first consider the case of approximating $f(x) = x^{N+1}$ at level 0. Equation (3.7) yields

$$\mathcal{Q}_0\, x^{N+1} = (N+1)\, \mathcal{N}_N\, x\, \sigma(x) + \mathcal{N}_{N+1}\, \sigma(x) \Leftrightarrow (N+1)\, \mathcal{N}_N\, \sigma_1(x).$$

As mentioned above, this formula has the disadvantage that the components of each subspace cannot be distinguished. Therefore we will try to rewrite this formula and isolate its component in W_0. First we group the first two terms,

$$\mathcal{Q}_0\, x^{N+1} = (N+1)\, \mathcal{N}_N\, (x+\alpha)\, \sigma(x) \Leftrightarrow (N+1)\, \mathcal{N}_N\, \sigma_1(x), \quad \text{with} \quad \alpha = \frac{\mathcal{N}_{N+1}}{(N+1)\, \mathcal{N}_N}.$$

Secondly, we isolate the component of $\sigma_1(x)$ in W_0, by letting $\mathcal{Q}_0\, \sigma_1(x) = \beta\, \sigma(x)$, $\beta = \langle \sigma_1(y), \psi(y) \rangle$, and $\bar{\sigma}_1(x) = \beta\, \sigma(x) \Leftrightarrow \sigma_1(x)$, such that $\sigma(x)$ and $\bar{\sigma}_1(x)$ are orthogonal. Then,

$$\mathcal{Q}_0\, x^{N+1} = (N+1)\, \mathcal{N}_N\, (x \Leftrightarrow \gamma)\, \sigma(x) + (N+1)\, \mathcal{N}_N\, \bar{\sigma}_1(x), \quad \text{with} \quad \gamma = \beta \Leftrightarrow \alpha.$$

In appendix B an algorithm to calculate β is described. In this formula the second term has no component in W_0. We see that in order to isolate the component in W_0, the modulating function (in this case x) has to be shifted over a distance γ. We now need the following theorem.

THEOREM 3.3. *If $\zeta(x)$ is a periodic function with period one and r is the number of vanishing moments of $\zeta(x)\,\psi(x)$ and $g(x) \in \mathcal{C}^r$, then $\mathcal{Q}_n\,[g(x)\,\zeta(2^n x)] = \mathcal{O}(h^r)$ where $h = 2^{-n}$.*

Proof. We know that

$$\mathcal{Q}_n[g(x)\,\zeta(2^n x)] = \sum_l \mu_{n,l}\,\psi(2^n x \Leftrightarrow l),$$

with

$$\mu_{n,l} = 2^n\ \langle\, g(y), \zeta(2^n y)\,\psi(2^n y \Leftrightarrow l)\,\rangle\ =\ \langle\, g(h\,(y+l)), \zeta(y)\,\psi(y)\,\rangle\,.$$

The proof follows from the Taylor formula. □

We assume that $f(x) \in \mathcal{C}^{N+2}$. Then from (3.7)

$$\mathcal{Q}_n f(x) = q_0(x)\,h^N + q_1(x)\,h^{N+1} + \mathcal{O}(h^{N+2}),$$

where $q_0(x) = f^{(N)}(x)\,\sigma_0^*(2^n x)$, $q_1(x) = f^{(N+1)}(x)\,\sigma_1^*(2^n x)$, and, from (3.8),

$$\begin{aligned}
\sigma_0^*(2^n x) &= \frac{\mathcal{N}_N}{N!}\sigma(2^n x) \\
\sigma_1^*(2^n x) &= \frac{\mathcal{N}_{N+1}}{N+1!}\sigma(2^n x) \Leftrightarrow \frac{\mathcal{N}_N}{N!}\sigma_1(2^n x) = \frac{\mathcal{N}_N}{N!}\left[\Leftrightarrow\gamma\,\sigma(2^n x) + \bar\sigma_1(2^n x)\right].
\end{aligned}$$

The problem here is that both $q_0(x)$ and $q_1(x)$ have a component in W_n. Indeed, since $\sigma(x)\,\psi(x)$ has a non vanishing integral, theorem 3.3 says that $\mathcal{Q}_n q_0(x) = \mathcal{O}(h^0)$ and $\mathcal{Q}_n q_1(x) = \mathcal{O}(h^0)$. We can solve this problem by grouping the components in W_n,

$$\begin{aligned}
\mathcal{Q}_n f(x) &= \frac{h^N \mathcal{N}_N}{N!}\left[\left(f^{(N)}(x) \Leftrightarrow h\gamma f^{(N+1)}(x)\right)\sigma(2^n x) + h f^{(N+1)}(x)\bar\sigma_1(2^n x) + \mathcal{O}(h^2)\right] \\
&= r_{0,0}(x)\,h^N + r_{0,1}(x)\,h^{N+1} + \mathcal{O}(h^{N+2}),
\end{aligned}$$

where

$$r_{0,0}(x) = \frac{\mathcal{N}_N}{N!}f^{(N)}(x \Leftrightarrow h\gamma)\,\sigma(2^n x) \quad \text{and} \quad r_{0,1}(x) = \frac{\mathcal{N}_N}{N!}f^{(N+1)}(x)\,\bar\sigma_1(2^n x).$$

In figure 3.3, $\sigma_1(x)$ and $\bar\sigma_1(x)$ are shown for the Daubechies wavelets. As a consequence of the orthogonalization, $\bar\sigma_1(x)$ is smaller then $\sigma_1(x)$ and new first term is more accurate. From theorem 3.3 with $\sigma(x)$ and $\bar\sigma_1(x)$ as $\zeta(x)$, respectively, we see that $\mathcal{Q}_n r_{0,0}(x) = \mathcal{O}(h^0)$ and $\mathcal{Q}_n r_{0,1}(x) = \mathcal{O}(h^1)$. This means the component in W_n of the $\mathcal{O}(h^{N+1})$ term will tend to zero if $h \to 0$. Now, for $i > 0$,

$$\begin{aligned}
\mathcal{Q}_{n+i} f(x) &= \frac{h^N\,\mathcal{N}_N}{N!\,2^{iN}}\Big[f^{(N)}(x)\,\sigma(2^{n+i}x) + \\
&\qquad\qquad h\frac{f^{(N+1)}(x)}{2^i}\left(\Leftrightarrow\gamma\,\sigma(2^{n+i}x) + \bar\sigma_1(2^{n+i}x)\right) + \mathcal{O}(h^2)\Big] \\
&= r_{i,0}(x)\,h^N + r_{i,1}(x)\,h^{N+1} + \mathcal{O}(h^{N+2}),
\end{aligned}$$

TABLE 3.2
The shift for the original Daubechies wavelets.

N	γ	N	γ	N	γ	N	γ
1	0	6	-1.9942	11	-3.9906	16	-6.0034
2	-0.4301	7	-2.3909	12	-4.3924	17	-6.4068
3	-0.8187	8	-2.7892	13	-4.7947	18	-6.8104
4	-1.2077	9	-3.1888	14	-5.1973	19	-7.2142
5	-1.5996	10	-3.5893	15	-5.6003	20	-7.6179

TABLE 3.3
The shift for the "most symmetric" Daubechies wavelets.

N	γ	N	γ
2	-0.4301	6	-0.0189
3	-0.8187	7	0.0683
4	0.0860	8	-0.0741
5	0.6132	9	0.3431

with

$$r_{i,0}(x) = \frac{\mathcal{N}_N}{N!\,2^{iN}} f^{(N)}(x \Leftrightarrow \gamma h)\, \sigma(2^{n+i}x)$$

and

$$r_{i,1}(x) = \frac{\mathcal{N}_N}{N!\,2^{i(N+1)}} f^{(N+1)}(x) \left((2^i \Leftrightarrow 1)\gamma\, \sigma(2^{n+i}x) + \bar{\sigma}_1(2^{n+i}x)\right).$$

Lemma C.3 states that $\sigma_j(2^i x)\,\psi(x)$ has $2N$ vanishing moments if $0 \le j < N$. Thus, using theorem 3.3 with $\sigma(2^i x)$ and $\sigma_1(2^i x)$ as $\zeta(x)$, yields respectively $\mathcal{Q}_n r_{i,0}(x) = \mathcal{O}(h^{2N})$ and $\mathcal{Q}_n r_{i,1}(x) = \mathcal{O}(h^{2N})$. So the only term with a component in W_n that is independent of h is $r_{0,0}(x)$. Consequently,

$$\boxed{\mathcal{E}_n f(x) = \frac{h^N\, \mathcal{N}_N}{N!} f^{(N)}(x \Leftrightarrow h\gamma)\, \tau(2^n x) + \mathcal{O}(h^{N+1}),} \tag{3.13}$$

where shifting the modulating function, yields an $\mathcal{O}(h^{N+1})$ term with a component in W_n that is $\mathcal{O}(h)$ and thus a more accurate first term. This is illustrated with two examples in the next section.

It is easy to see that the shift γ is zero for (anti-)symmetric wavelets. This shift can be seen as a measure for the symmetry of the wavelet. Ingrid Daubechies showed that, except for the Haar case, no symmetric compactly supported orthogonal wavelets exist [7]. In [9] she constructed so called "most symmetric" wavelets, who have the closest to linear phase of all wavelets with support length $2N \Leftrightarrow 1$ and N vanishing moments.

In table 3.2 and 3.3 we give numerical values for γ in function of the number of vanishing moments both for the original Daubechies wavelets and for the "most symmetric" ones. For the original Daubechies the absolute value of the shift seems to be increasing linearly with N. As could be expected, the shift is smaller for the "most symmetric" ones.

3.6. Numerical Example. In this section we consider an example with

$$f(x) = \exp(\Leftrightarrow 20\,(x \Leftrightarrow 0.5)^2),$$

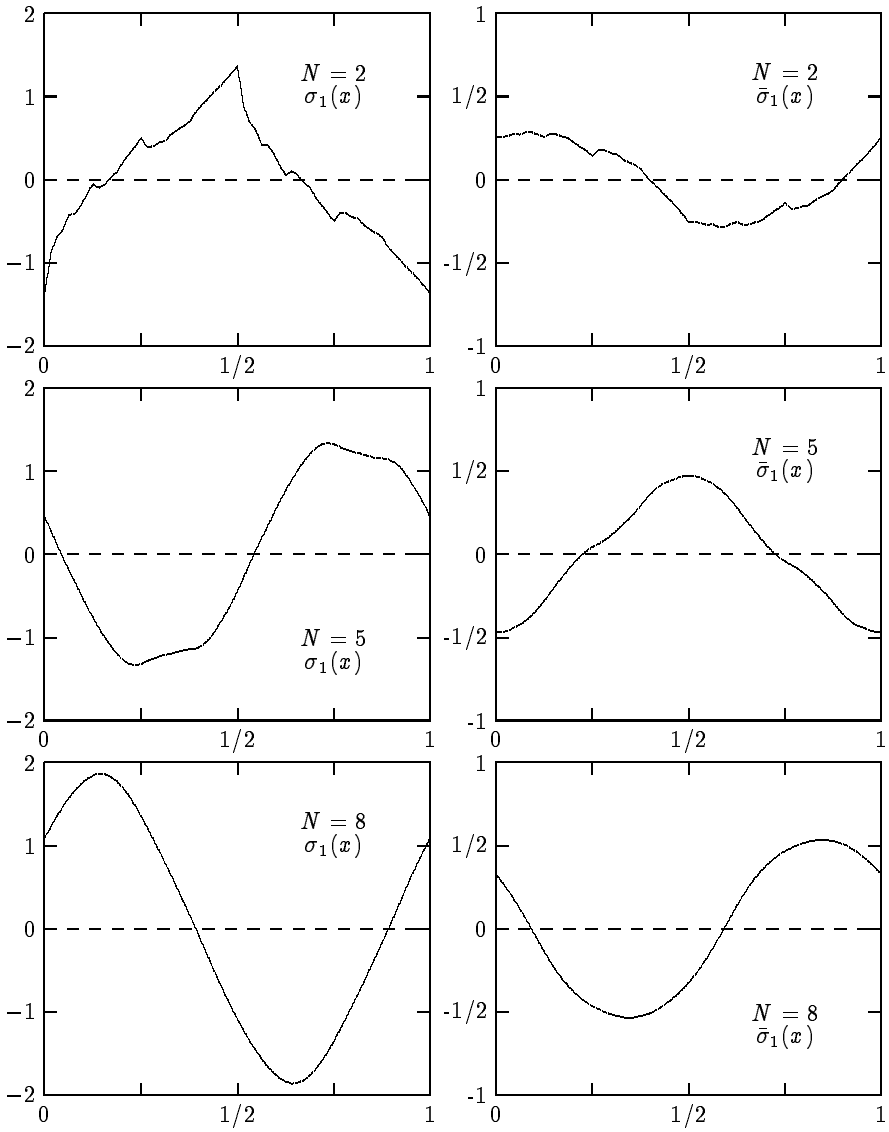

FIG. 3.3. $\sigma_1(x)$ *and* $\bar{\sigma}_1(x)$ *for Daubechies' wavelets with* $N = 2, 5, 8$

where we look at the error in the interval $[0,1]$. Figure 3.4 shows the error $\mathcal{E}_6 f(x)$ obtained with a computer program calculating $\mathcal{P}_6 f(x) \Leftrightarrow f(x)$. The wavelet used is the Daubechies wavelet with 3 vanishing moments. One can check that the interpolating properties are satisfied. On the same figure, the envelopes (3.12) of the first term of the expansion (3.9) are drawn dashed and the envelopes of the first term of (3.13) are drawn solid. Figure 3.5 shows the error $\mathcal{E}_6 f(x)$ using the Daubechies wavelet with 9 vanishing moments and the same envelopes. We see that the first term of the error expansion already gives a reasonable approximation of the actual error and secondly that shifting the modulating function indeed yields more accurate results. For both these examples the inner products $\langle f(x), \varphi_{n,l}(x) \rangle$ were calculated using a quadrature

formula with an error of $\mathcal{O}(h^{2N})$ so this influence can be neglected.

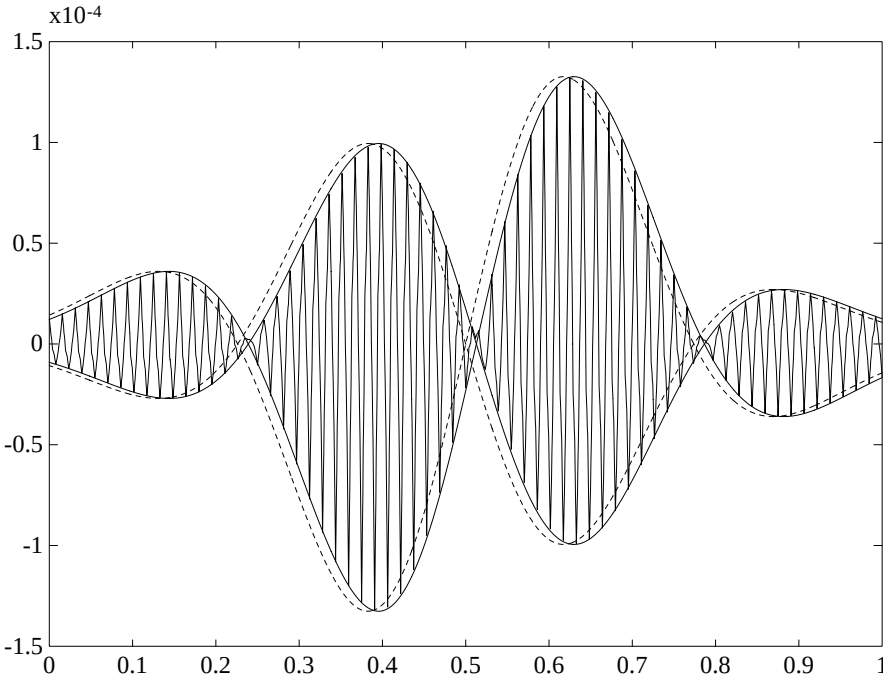

FIG. 3.4. *The error and envelopes for* $N = 3$ *and* $n = 6$.

3.7. Remarks. Using a generalization of Bernoulli polynomials it is possible to derive an asymptotic error expansion for the quadrature formula [6, 23]. As a result it is possible to use extrapolation techniques similar to Romberg integration.

Future research includes the careful study of how the use of a quadrature formula will influence the error expansion. First experiment show that the use of a quadrature formula with degree of accuracy $q = N \Leftrightarrow 1$ can ruin the interpolating properties of the wavelet expansion. Therefore one might want to use quadrature formulae with $q > N \Leftrightarrow 1$.

Acknowledgment. The authors are indebted to Pierre Verlinden for many fruitful discussions. He also pointed out a shorter way for some of the proofs.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of mathematical functions*, Dover Publications, New York, 1965.
[2] G. BATTLE, *A block spin construction of ondelettes*, Comm. Math. Phys., 110 (1987), pp. 601–615.
[3] G. BEYLKIN, R. COIFMAN, AND V. ROKHLIN, *Fast wavelet transforms and numerical algorithms I*, Comm. Pure and Appl. Math., 44 (1991), pp. 141–183.
[4] C. K. CHUI, *An Introduction to Wavelets*, Academic Press, San Diego, 1992.
[5] A. COHEN, I. DAUBECHIES, AND J. FEAUVEAU, *Bi-orthogonal bases of compactly supported wavelets*, Comm. Pure and Appl. Math., 45 (1992), pp. 485–560.
[6] W. DAHMEN AND C. A. MICCHELLI, *Using the refinement equation for evaluating integrals of wavelets*, SIAM J. Num. Anal., 30 (1993), pp. 507–537.

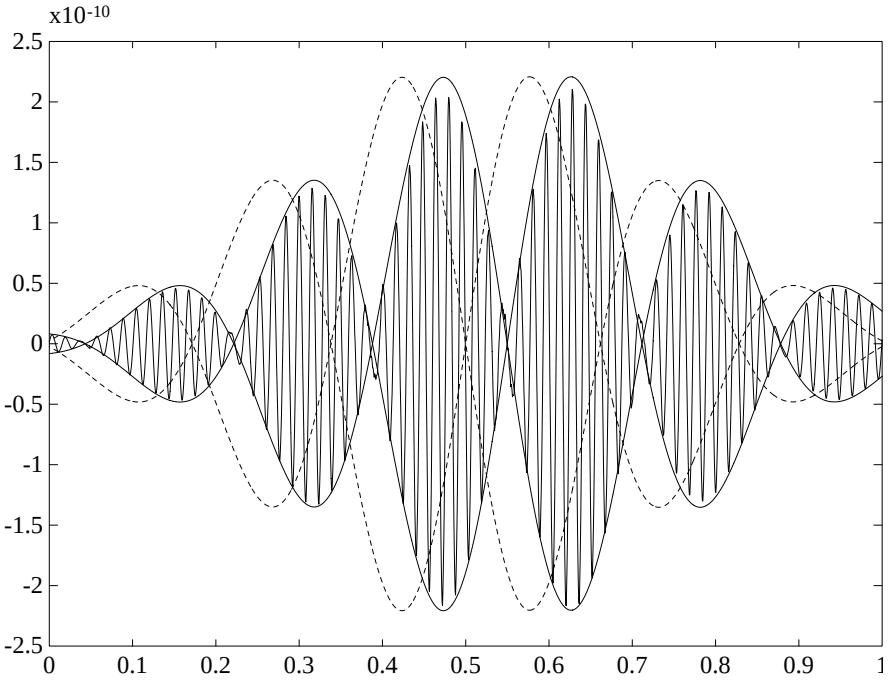

FIG. 3.5. *The error and envelopes for* $N = 9$ *and* $n = 6$.

[7] I. DAUBECHIES, *Orthonormal bases of compactly supported wavelets*, Comm. Pure and Appl. Math., 41 (1988), pp. 909–996.

[8] ———, *Ten Lectures on Wavelets*, no. 61 in CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[9] ———, *Orthonormal bases of compactly supported wavelets II. Variations on a theme*, SIAM J. Math. Anal., 24 (1993), pp. 499–519.

[10] P. DAVIS AND P. RABINOWITZ, *Methods of numerical integration*, Academic Press, London, 1984.

[11] G. FIX AND G. STRANG, *Fourier analysis of the finite element method in Ritz-Galerkin theory*, Stud. Appl. Math, 48 (1969), pp. 265–273.

[12] W. GAUTSCHI, *On the construction of Gaussian quadrature rules from modified moments*, Math. Comp., 24 (1970), pp. 245–260.

[13] ———, *Numerical condition related to polynomials*, in Recent Advances in Numerical Analysis, C. de Boor and G. H. Golub, eds., Mathematics Research Center, The University of Wisconsin, Academic Press, 1978, pp. 45–72.

[14] R. A. GOPINATH AND C. S. BURRUS, *On the moments of the scaling function*, in Proceedings of the ISCAS-92, San Diego, 1992.

[15] P.-G. LEMARIÉ, *Ondelettes a localisation exponentielle*, J. de Math. Pures et Appl., 67 (1988), pp. 227–236.

[16] S. G. MALLAT, *Multifrequency channel decompositions of images and wavelet models*, IEEE Trans. on Acoust. Signal Speech Process., 37 (1989), pp. 2091–2110.

[17] ———, *Multiresolution approximations and wavelet orthonormal bases of* $L^2(\mathbf{R})$, Trans. Amer. Math. Soc., 315 (1989), pp. 69–87.

[18] Y. MEYER, *Ondelettes et Opérateurs, I: Ondelettes, II: Opérateurs de Calderón-Zygmund, III:* (with R. Coifman), *Opérateurs multilinéaires*, Hermann, Paris, 1990.

[19] R. PIESSENS AND M. BRANDERS, *The evaluation and application of some modified moments*, BIT, 13 (1973), pp. 443–450.

[20] G. STRANG, *Wavelets and dilation equations: A brief introduction*, SIAM Review, 31 (1989), pp. 614–627.

[21] G. STRANG AND G. FIX, *A Fourier analysis of the finite element variational method*, in

Constructive aspects of Functional Analysis, Rome, 1973, Edizione Cremonese.

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

[23] P. Verlinden and A. Haegemans, *An asymptotic expansion in wavelet analysis and its application to accurate numerical wavelet decomposition*, Numer. Algor., 2 (1992), pp. 287–298.

A. Algorithms for the modified construction. Algorithm for the calculation of the coefficients $q_{i,j}$ in section 2.5.2:

$q_{0,0}^{(0)} \leftarrow 1/2$

for $m \leftarrow 1\,(1)\,r$

  for $i \leftarrow 0\,(1)\,m$

    for $j \leftarrow 0\,(1)\,m \Leftrightarrow i$

      $q_{i,j}^{(m)} \leftarrow q_{i-1,j}^{(m-1)} + q_{i,j-1}^{(m-1)} + q_{i+1,j}^{(m-1)} + q_{i,j+1}^{(m-1)} \Leftrightarrow 2\, d_m\, q_{i,j}^{(m-1)}$

      if i = 1 **then** $q_{1,j}^{(m)} \leftarrow q_{1,j}^{(m)} + q_{0,j}^{(m-1)}$

      if j = 1 **then** $q_{i,1}^{(m)} \leftarrow q_{i,1}^{(m)} + q_{i,0}^{(m-1)}$

    end for

  end for

end for

Algorithm for the calculation of the coefficients $w_i^{(p)}$ in section 2.5.3:

$w_0^{(0)} \leftarrow 1,\ w_0^{(1)} \leftarrow \lambda,\ w_1^{(1)} \leftarrow 1$

$w_0^{(2)} \leftarrow 2\lambda^2 \Leftrightarrow 3,\ w_1^{(2)} \leftarrow 4\lambda,\ w_2^{(2)} \leftarrow 1$

for $p \leftarrow 2\,(1)\,\ldots$

  $w_0^{(p+1)} \leftarrow w_1^{(p)} + 2\lambda w_0^{(p)} \Leftrightarrow 4 w_0^{(p-1)}$

  $w_1^{(p+1)} \leftarrow 2w_0^{(p)} + w_2^{(p)} + 2\lambda w_1^{(p)} \Leftrightarrow 4 w_1^{(p-1)}$

  for $i \leftarrow 2\,(1)\,p \Leftrightarrow 1$

    $w_i^{(p+1)} \leftarrow w_{i-1}^{(p)} + w_{i+1}^{(p)} + 2\lambda w_i^{(p)} \Leftrightarrow 4 w_i^{(p-1)}$

  end for

  $w_p^{(p+1)} \leftarrow w_{p-1}^{(p)} + 2\lambda w_p^{(p)}$

  $w_{p+1}^{(p+1)} \leftarrow w_p^{(p)} = 1$

end for

B. Calculation of β. In this section an algorithm to calculate β accurately is described. We can write

$$\begin{aligned}
\beta &= \langle x, \sigma(x)\,\psi(x)\rangle \\
&= 2\sum_k g_k \langle x, \sigma(x)\,\varphi(2x \Leftrightarrow k)\rangle \\
&= 1/2\sum_k g_k \langle x, \sigma((x+k)/2)\,\varphi(x)\rangle + 1/2\sum_k g_k\, k \langle \sigma((x+k)/2), \varphi(x)\rangle \\
&= 1/2\,(p_1 + p_2),
\end{aligned}$$

with

$$p_2 = \sum_k g_k\, k \langle \sigma((x+k)/2), \varphi(x)\rangle = \sum_k h_{1-k}\, k \langle \sigma(x/2), \varphi(x)\rangle = 1 \Leftrightarrow \mathcal{M}_1,$$

and

$$p_1 = \sum_k g_k \langle x, \sigma((x+k)/2)\,\varphi(x)\rangle = \langle x, \sigma(x/2)\,\varphi(x)\rangle = \sum_l (\Leftrightarrow 1)^l \kappa_l,$$

where

$$\kappa_l = \langle x, \varphi(x \Leftrightarrow l)\,\varphi(x)\rangle.$$

The κ_l can be found by solving the linear system

$$\kappa_m = \sum_l a_{l-2m}\,\kappa_l + b_{2m},$$

with

$$a_i = \sum_k h_k\, h_{k+i} \quad \text{and} \quad b_i = \sum_k k\, h_k\, h_{k+i}.$$

C. Proof of some Lemma's.

LEMMA C.1. *The function $\sigma_j(x)\,\varphi(x)$ has N vanishing moments if $0 \le j < N$ and no vanishing moments if $N \le j$.*

Proof.

$$\langle x^p, \sigma_j(x)\,\varphi(x)\rangle = \sum_l \langle (x+l)^p, x^j\,\psi(x)\,\varphi(x+l)\rangle = \mathcal{M}_p\,\mathcal{N}_j \qquad \text{if } 0 \le p < N$$

This is zero if $0 \le j < N$ and non-zero if $p = 0$ and $j = N$. □

LEMMA C.2. *The following functions also have N vanishing moments if $0 \le j < N$ and no vanishing moments if $N \le j$: $\sigma_j(x)\,\varphi(x \Leftrightarrow l)$ with $l \in \mathbf{Z}$, $\sigma_j(2^i x)\,\varphi(x)$ with $i \in \mathbf{N}$, $\sigma_j^*(x)\,\varphi(x)$, and $\tau_j(x)\,\varphi(x)$.*

LEMMA C.3. *The function $\sigma_j(2^i x)\,\psi(x)$ with $i > 0$ has $2N$ vanishing moments if $0 \le j < N$ and N vanishing moments if $N \le j$.*

Proof. For $i = 1$

$$\begin{aligned}
\langle x^p, \psi(x)\,\sigma_j(2x)\rangle &= 2\sum_l g_l \langle x^p, \varphi(2x \Leftrightarrow l)\,\sigma_j(2x)\rangle \\
&= 2^{-p}\sum_l g_l \langle x^p, \varphi(x \Leftrightarrow l)\,\sigma_j(x)\rangle \\
&= 2^{-p}\sum_l g_l \langle (x+l)^p, \varphi(x)\,\sigma_j(x)\rangle
\end{aligned}$$

If $0 \leq p < N$ this is zero for all j because of lemma C.1. If $N \leq p < 2N$ the inner product in the summation is a polynomial of degree $p \Leftrightarrow N < N$. To prove the lemma for $i > 1$, we can follow the same reasoning as above and use lemma C.2. □