Wavelet Calculus and Finite Difference Operators*

Kent McCormick and Raymond O. Wells, Jr.

March 17, 1993

Abstract

This paper shows that the naturally induced discrete differentiation operators induced from a wavelet-Galerkin finite-dimensional approximation to a standard function space approximates differentiation with an error of order $O(h^{2d+2})$, where d is the degree of the wavelet system. The degree of a wavelet system is defined as one less than the degree of the lowest order non-vanishing moment of the fundamental wavelet. We consider in this paper compactly supported wavelets of the type introduced by Daubechies in 1988. The induced differentiation operators are described in terms of connection coefficients which are intrinsically defined functional invariants of the wavelet system (defined as L^2 inner products of derivatives of wavelet basis functions with the basis functions themselves). These connection coefficients can be explicitly computed without quadrature and they themselves have key moment vanishing properties proved in this paper which are dependent upon the degree of the wavelet system. This is the basis for the proof of the principal results concerning the degree of approximation of the differenitaion operator by the wavelet-Galerkin discrete differentiation operator.

*AMS Classification Numbers: Primary 65D25, Secondary 65L12, 42C99; Keywords: wavelets, finite-difference operators, Galerkin approximation, connection coefficients

Contents

1 Introduction

The recently discovered compactly supported wavelets of Daubechies [4] have proven to be a useful tool in the numerical solutions of partial differential equations (see [7,12,17,19,20,16] for recent numerical solutions of partial differential equations using wavelet-Galerkin techniques. In each of the papers above the notion of connection coefficients has played a key role in the approximations of partial derivatives as well as nonlinear terms in the wavelet discretization of the differential equations. Connection coefficients and some of their properties and algorithms for computing them are described in [11] and [1].

The purpose of this paper is to show that there is a relationship between discrete differentiation using connection coefficients and discrete differentiation using finite-difference operators.

A certain class of finite-difference operators have the property that operating on the discretization of a polynomial of degree d is equivalent to differentiating the polynomials and then discretizing. This implies that the finite-difference operator approximates the derivative up to order d, and conversely. Let's explain what this means. Consider a lattice

$$\Lambda = \{x_0 + ih\}, \quad i \in \mathbf{Z},$$

in $\mathbf{R}$, for some fixed $h \in \mathbf{R}$. Let $x_i = x_0 + ih$ denote a generic point of Λ, and let $C(\Lambda)$ denote the continuum (= arbitrary) real-valued functions on Λ. Then a finite-difference operator ∇ is a mapping of the form

$$\begin{array}{ccc} C(\Lambda) & \stackrel{\nabla}{\to} & C(\Lambda) \\ & & \\ f(x_i) & \mapsto & g(x_i) \end{array} \tag{1.1}$$

where

$$g(x_i) = \nabla f(x_i) = \sum_{k=-N_1}^{N_1} \frac{w_i^k}{h} f(x_i + kh) \tag{1.2}$$

where the w_i^k are the *weights* of the finite-difference operator. For instance, there is a classical 5-point symmetric difference operator of the form

$$\begin{aligned}\nabla_4(f)(x_i) &= \frac{1}{12h}\left[f(x_i-2h)-8f(x_i-h)+8f(x_i+h)-f(x_i+2h)\right] \\ &= \frac{1}{12h}\left[f(x_{i-2})-8f(x_{i-1})+8f(x_{i+1})-f(x_{i+2})\right],\end{aligned} \tag{1.3}$$

and the first form of the operator indicates the origin of the name finite-difference operator. This finite-difference operator has the simpler form (defined at any point $x \in \mathbf{R}$)

$$\nabla f(x) = \sum_{k=-N}^{N} \frac{w^k}{h} f(x+kh), \tag{1.4}$$

which is *independent* of the point x where it is evaluated. This is the only kind of finite-difference operators we shall consider in this paper.

By using wavelet interpolation to discretize continuous functions (see [19]), and by using the connection coefficients determined by the wavelet system to obtain a discrete approximation one obtains a wavelet discrete differentiation operator of the form

$$D_J(f)(x_0) = \sum_{k=-2g+2}^{2g-2} 2^J \Gamma_0^k f(x_0+kh). \tag{1.5}$$

Here $h = 2^{-J}$ is the mesh size for some fixed scaling J of the wavelet system, and

$$\Gamma_0^k = \int \varphi'(x)\varphi(x-k)\,dx$$

are connection coefficients as defined in Section 3 below, where $\varphi(x)$ is the scaling function of the given wavelet system.

The main point of this paper is that there is a strong relation between the classical finite-difference operators and the induced discrete wavelet differentiation operators using connection coefficients. As an example, for Daubechies wavelets of genus 2 (four coefficients), the corresponding operator D_J (given in Section 5) coincides precisely with the classical operator ∇_4 given in (1.3) above, where $h = 2^{-J}$ (as was observed independently by Beylkin [1]). More generally, we will see that the wavelet differntiation operator D for a given wavelet system with L vanishing moments of the wavelet function, i.e.,

$$\int_{\mathbf{R}} x^l \psi(x)\,dx = 0, \quad i = 0, \ldots, L, \tag{1.6}$$

has the property that D commutes with differentiation up to degree $2L+2$. We recall that the Daubechies wavelets are characterized by having $g-1$ moments vanishing for wavelet systems defined by $2g$ coefficients (these are wavelet systems of rank 2 and genus g as in [10]). These wavelet systems have the property that they *represent* polynomials up to degree $g-1$ exactly, but the wavelet differentiation operator for these wavelets commutes with differentiation of polynomials up to degree $2g$, slightly more than twice the degree one might expect from the exact representation.

In Section 2 we find a characterization of finite-difference operators which approximate the first derivative to a certain order in terms of the action on polynomials up to the same degree. This will be used to show that the wavelet discrete differentiation has certain approximation properties. In Section 3 we give a description of the wavelet systems, wavelet interpolation, and wavelet differentiation. In Section 4 we prove a fundamental result about the vanishing of discrete moments for the connection coefficients. From this we derive the polynomial mapping property of wavelet differentiation. In Section 6 we find a comparison of the classical Stirling series of finite difference operators and the induced wavelet discrete differentiation operators.

In an earlier paper, similar results to the approximation properties of wavelet operators described in this paper were found by Gopinath, Lawton and Burrus [9]. The connection coefficients used in our paper (and for which there are algorithms for their computation in [1,11])) are sets of specific rational numbers which provide concrete realizations of the wavelet-Galerkin discretization and approximations of linear operators described in [9]. The authors of [9] express their results in terms of convolution operators and approximations of linear translation invariant operators. They have analogues of the vanishing moments results that we have here, but the proofs and the formulation in the two papers are independent. We formulated our results in the language applicable to the papers in partial differential equations at the beginning of this introduction, where the connection coefficients play a key computational role in the numerical solutions. It's clear that the results of the present paper will have generalization to higher order derivatives, etc., but we don't carry out that program at this time.

The authors would like to thank Ramesh Gopinath and Xiaodong Zhou for useful conversations during the preparation of this paper.

2 Finite difference approximations to differentiation

Let $\delta : C^k(\mathbf{R}) \to C(\Lambda)$ be a discretization operator mapping functions on $\mathbf{R}$ to Λ of the form: $\delta(f)(x_i) = cf(x_i)$, for $f \in C^k(\mathbf{R})$ and for some constant c (a multiple of the restriction of f from $\mathbf{R}$ to Λ). Now consider the diagram

$$\begin{array}{ccc} C^k(\mathbf{R}) & \stackrel{D}{\to} & C^{k-1}(\mathbf{R}) \\ \downarrow \delta & & \downarrow \delta \\ C(\Lambda) & \stackrel{\nabla}{\to} & C(\Lambda), \end{array} \tag{2.7}$$

for any finite-difference operator of the form (1.2) and where $D = \frac{d}{dx}$ denotes the differentiation operator. We will say that ∇ *commutes with D up to degree d* if the restricted diagram

$$\begin{array}{ccc} P_d(\mathbf{R}) & \stackrel{\frac{d}{dx}}{\to} & P_{d-1}(\mathbf{R}) \\ \downarrow \delta & & \downarrow \delta \\ C(\Lambda) & \stackrel{\nabla}{\to} & C(\Lambda) \end{array} \tag{2.8}$$

commutes where $P_d(\mathbf{R})$ is the vector space of polynomials of degree $\leq d$. The "diagram (2.8) commutes" means that

$$\delta \circ D = \nabla \circ \delta,$$

i.e., the linear mappings in the diagram are compatible. This is clearly a strong restriction on a finite-difference operator for large d. We can now formulate the following classical result.

Theorem 2.1 *A finite-difference operator ∇ of the form* (1.4) *commutes with D up to degree d if and only if*

$$f'(x_i) - \nabla f(x_i) = \mathrm{O}(h^d), \tag{2.9}$$

for $f \in C^{d+1}(\mathbf{R})$.

The proof is well known and follows from the Peano representation of the functional $D - \nabla$ (see, e.g. [6]).

The operator ∇_4 given in (1.3) satisfies the estimate (2.9) for $d = 4$, and hence ∇_4 commutes with differentiation up to degree 4. This is a special case of a general class of symmetric finete-difference operators ∇_{2n} studied by Stirling in the 18th century (see [15,3]), where ∇_{2n} estimates f to order $O(h^{2n})$ as in equation (2.9). These points are discussed in further detail in Section 6. We shall say that a finite-difference operator ∇ is an *approximate derivative of order* d if ∇ satisfies (2.9).

By virtue of Theorem 2.1, to verify that a finite-difference operator is an approximate derivative of order d, it suffices to show that it commutes with differentiation of polynomials up to degree d, and we'll use this later in this paper.

3 Wavelet Interpolation

In this section we review some fundamental facts concerning wavelet systems. See [2] and [5] for more details on the general theory of wavelets. Let $\{a_k\}, k = 0, \ldots, 2g-1, g \geq 1$, be a vecto which satisfies: the *quadratic* and *linear* conditions:

$$\sum a_k a_{k+2l} = 2\delta_{0,l} \qquad \sum a_k = 2. \tag{3.10}$$

Let φ and ψ be compactly supported L^2 functions which satisfy the *scaling equation*:

$$\varphi(x) = \sum_{k=0}^{2g-1} a_k \varphi(2x-k), \tag{3.11}$$

and the *wavelet equation*

$$\psi(x) = \sum_{k=0}^{2g-1} b_k \varphi(2x-k), \tag{3.12}$$

respectively, where the coefficients a_k satisfy (3.10) above, and where $b_k = (-1)^k a_{2g-1-k}$. The support of φ and ψ is the compact interval $[0, 2g-1]$, and these will generate (by translations and dilations) a *wavelet system* for representing functions in $L^2(\mathbf{R})$ by means of a series like the orthonomal series above.

More specifically we define the *wavelet system* associated with the scaling coefficients a_k by

$$\varphi_k(x) := \varphi(x-k), \quad \psi_{jk}(x) := 2^{j/2}\psi(2^j x - k), \; j \geq 0$$

and the *wavelet expansion*

$$f(x) = \sum_k f_k \varphi_k(x) + \sum_{jk} f_{jk}\psi_{jk}(x)$$
$$f_k = \int f(x)\varphi_k(x)dx, \quad f_{jk} = \int f(x)\psi_{jk}(x)dx.$$

This is the wavelet series expansion with expansion (wavelet) coefficients f_k and f_{jk}. The vector $(a_0, \ldots, a_{2g-1})$ is called the *scaling vector* ("vector of scaling coefficients for the scaling difference equation"), and the vector $(b_0, \ldots, b_{2g-1})$ is called the *wavelet vector*. The parameter g is called the genus of the wavelet vector, and the number 2 is called the multiplier or rank of the wavelet system (there are wavelet systems of higher rank, see, e.g., [10], [8]). Examples of wavelet vectors are given in [5] as well as many other places. The Daubechies wavelets are characterized by having maximal vanishing moments of the wavelet function, i.e.,

$$\int x^l \psi(x)\, dx = 0, \quad l = 0, \ldots, g-1,$$

which determines φ and ψ up to a finite set of choices of the a_k, and we call the *Daubechies wavelets* of genus g the specific choice having maximal vanishing moments given in her paper [4].

Assume that a scaling vector a_k is given so that the associated wavelet system is an orthonormal basis for $L^2(\mathbf{R})$ (almost all scaling vectors have this property and, in particular the Daubechies wavelets have this property, see [13]). Define

$$\varphi_{jk}(x) = 2^{j/2}\varphi(2^j x - k)$$

the rescaling and translates of the scaling function. Let, for $j \in Z$,

$$V_j := \text{closure}(\text{span}\{\varphi_{jk} : k \in \mathbf{Z}\}).$$

Then $L^2(\mathbf{R}) = \bigcup_j V_j$, that is, for any function $f \in L^2(\mathbf{R})$, if we let P_j denote the orthogonal projection $L^2(\mathbf{R}) \to V^j$, then $P_j f$ converges to f in the L^2 norm. The coefficients of $P_j f$ are given by the classical orthonormal expansion type formulas (of the type discussed earlier)

$$P_j f(x) = \sum_{k \in \mathbf{Z}} f^0_{jk}\varphi_{jk}(x),$$

where

$$f_{jk}^0 = \int_{\mathbf{R}} f(x)\varphi_{jk}(x)\,dx,$$

since the φ_{jk}'s are an orthonormal basis for V_j. We will denote the expansion coefficients of the scaled and translated scaling functions generically by $f_{*,*}^0$ and of the scaled and translated wavelet functions by $f_{*,*}^1$, letting the superscripts $\{0\}$ and $\{1\}$ distinguish between the two cases.

If we let W_j denote the orthogonal complement of V_j in V_{j+1}, then, for a fixed $J \in \mathbf{Z}^+$ (where $\mathbf{Z}^+$ denotes the nonnegative integers),

$$V_0 \oplus W_0 \oplus W_1 \oplus \ldots \oplus W_{J-1} = V_J.$$

Moreover, each of the spaces W_j is spanned by the L^2 wavelet functions ψ_{jk}, for j fixed, and for $k \in \mathbf{Z}$.

If we let

$$f(x) = \sum_{k\in\mathbf{Z}} f_{0k}^0 \varphi_{0k}(x) + \sum_{j\in\mathbf{Z}^+} \sum_{k\in\mathbf{Z}} f_{jk}^1 \psi_{jk}(x),$$

for $f \in L^2(\mathbf{R})$, then we find that,

$$\sum_{k\in\mathbf{Z}} f_{Jk}^0 \varphi_{Jk}(x) = \sum_{k\in\mathbf{Z}} f_{0k}^0 \varphi_{0k}(x) + \sum_{j=0}^{J-1} \sum_{k\in\mathbf{Z}} f_{jk}^1 \psi_{jk}(x). \tag{3.13}$$

The left hand side represents expansion information at a *fixed* scale, while the right hand side has the same information represented in terms of a sequence of scales (the multiscale representation). In both cases the expansion coefficients are given formally by the usual orthonormal expansion coefficient formulas

$$\begin{aligned} f_{jk}^0 &= \int_{\mathbf{R}} f(x)\varphi_{jk}(x)\,dx, \quad j \in \mathbf{Z}^+, \\ f_{jk}^1 &= \int_{\mathbf{R}} f(x)\psi_{jk}(x)\,dx, \quad j \in \mathbf{Z}^+. \end{aligned}$$

Moreover, it follows from the 2-scale difference (3.11) equation that one can determine the coefficients on the right hand side in terms of the coefficients on the left hand side, and conversely. This is the Mallat algorithm, and the formulas are easy to derive ([14], see also [19]). We won't need them explicitly in this paper.

Moreover, we want to remark that if we consider *any* expansion of the form (the scaling function expansion at a fixed scale or level)

$$\tilde{f}_J(x) := \sum_{k\in\mathbf{Z}} f^0_{Jk}\varphi^J_k(x),$$

where the coefficients f^0_{Jk} are prescribed in some fashion (perhaps relating to some given L^2 function f or not), then there is a corresponding expansion of the form (the wavelet or multiscale expansion)

$$\tilde{f}_J(x) = \sum_{k\in\mathbf{Z}} f^0_{0k}\varphi_{0k}(x) + \sum_{j=0}^{J-1}\sum_{k\in\mathbf{Z}} f^1_{jk}\psi(x)_{jk}(x),$$

and the coefficients in the wavelet expansion are determined in terms of the coefficients in scaling function expansion and the converse is true. This is the *Mallat algorithm* in general. It doesn't have anything to do with the coefficients being generalized Fourier coefficients of some given function.

Let f be a smooth L^2 function and represent f as a wavelet series:

$$f(x) = \sum f_k\varphi_k(x) + \sum f_{jk}\psi_{jk}(x),$$

and differentiate

$$f'(x) = \sum f_k\varphi'_k(x) + \sum f_{jk}\psi'_{jk}(x)$$

Expand the derivatives of the scaling and wavelet functions in a wavelet series (assuming that the derivatives are computed either classically or as distributions):

$$\begin{aligned}
\varphi'_k(x) &= \sum_l \Gamma^l_k\varphi_l + \sum_{il}\Gamma^{il}_k\psi_{il}\\
\psi'_{jk}(x) &= \sum_l \Gamma^l_{jk}\varphi_l + \sum_{il}\Gamma^{il}_{jk}\psi_{il}
\end{aligned}$$

The $\Gamma^l_k, \Gamma^{il}_k$, etc., are called *connection coefficients*, and have been explicitly calculated [1,11] for specific families of wavelets including the Daubechies wavelets which have a maximal moment vanishing property. The formulas are formally given by:

$$\Gamma^l_k = \int \varphi'(x-k)\varphi(x-l)dx, \qquad \Gamma^{il}_k = \int \varphi'(x-k)\psi_{il}(x)dx, \quad (3.14)$$

with the Γ_{jk}^{l} and Γ_{jk}^{il} being given by similar formulas. We see that the formulas for the connection coefficients are given as integrals. If one had to approximate these integrals numerically in order use the connection coefficients then one would have accuracy problems in general. As it turns out (and is shown in [1,11]), the values of the connection coefficients can be calcuated as the solution of specific linear equations, and in fact the values turn out to be rational numbers in general. Programs have been written which generate these connection coefficient efficiently in various cases, and for a given computation the connection coefficients can be stored for use in a given problem. For the Daubechies wavelet systems for genus 2 (four coefficient Daubechies wavelets) these numbers can be calculated, and one obtains for instance:

$$\Gamma_0^l = \left(\frac{1}{12}, -\frac{8}{12}, 0, \frac{8}{12}, -\frac{1}{12}\right), \quad l = -2, -1, 0, 1, 2. \tag{3.15}$$

Other examples of connection coefficients are computed in reference tables in [1,11].

Let us consider the projection of a given function $f \in L^2(\mathbf{R})$ into V_J for some fixed scale J (fixed resolution). There are two natural projections:

- L^2 *Orthogonal Projection:*

$$\begin{aligned} f &\mapsto P_j f \\ P_J &= \sum f_{Jk}^{P} \varphi_k(x) \\ f_{Jk}^{P} &:= \int f(x)\varphi_{Jk}(x)dx \end{aligned}$$

- *Interpolation Projection:*

$$\begin{aligned} f &\mapsto I_J f \\ I_J f &= \sum f_{Jk}^{I} \varphi_k(x) \\ f_{Jk}^{I} &:= f\left(\tfrac{k}{2^J}\right) \end{aligned}$$

  In $I_J f$ the evaluation could be at some other point centered near the support of the wavelet.

Since $\int \varphi_{Jk}(x)dx = 2^{-\frac{J}{2}}$, and $\operatorname{supp} \varphi_{Jk} = \left[\frac{k}{2^J}, \frac{k+2g-1}{2^J}\right]$ it follows that

$$2^{-\frac{J}{2}} f\left(\frac{k}{2^J}\right) \approx \int f(x)\varphi_{Jk}(x)dx$$

where we note that $2^{\frac{J}{2}}\varphi_{Jk}$ is an approximation to the Dirac delta function near the point $\frac{k}{2^J}$.

Thus for large J, both $P_J f$ and $I_J f$ are good approximations to each other and to f (see [19] for an error analysis of these projections).

In practice, for a digitized signal, we *identify* for a fixed high resolution scale J the scaling expansion coefficients at that scale (f_{Jk}) with the values of the digitized function values, and then use the Mallat algorithm to obtain the wavelet expansion coefficients for coarser resolution scales $(j < J)$. There is *no integration* in the evaluation of the wavelet expansion coefficients.

If we consider a discrete function (or digitized continous function), which we consider as an approximation of a given funcition then we can use the connection coefficients to calculate derivatives. We will describe this explicitly in Section 5 below, but first we want to derive some important properties of connection coefficients as defined above.

4 Moment Properties of Connection Coefficients

We now consider connection coefficents Γ_0^l and Γ_0^{il} for a given wavelet system. Note that knowing these connection coefficents with subscript 0 determine the full set of connection coefficients Γ_k^l and Γ_k^{il} by translation, and we will concentrate on these specific connection coefficients in our results in this section.

We now have the following fundamental theorem concerning these coefficients.

Theorem 4.1 *If* $\int x^j \psi(x)dx = 0$ *for* $j = 0 \ldots L$*, then:*

(a) $\sum_l l^\alpha \Gamma_0^l = 0$ *for* $0 \leq \alpha \leq 2L+2, \alpha \neq 1$, $\sum_l l\Gamma_0^l = -1$,

(b) $\sum_{jk} k^\alpha \Gamma_0^{jk} = 0$ *for* $0 \leq \alpha \leq L$.

The proof will occupy the remainder of this section and will involve several lemmas.

Lemma 4.2 *If* $\sum_k k^\alpha \Gamma_0^{0,k} = 0$ *for* $0 \leq \alpha \leq n$*, then* $\sum_{jk} k^\alpha \Gamma_0^{jk} = 0$ *for* $0 \leq \alpha \leq n$.

Proof: We have the following relationship of the connection coefficients, which is easily derived from the definition (3.14) and from the wavelet equation (3.12):

$$\Gamma_0^{jk} = \sqrt{2}\sum_{m=0}^{2g-1} a_m \Gamma_0^{j-1,k-2^{j-1}m} \tag{4.16}$$

which implies

$$\sum_k k^\alpha \Gamma_0^{jk} = \sqrt{2}\sum_{m=0}^{2g-1} a_m \sum_k k^\alpha \Gamma_0^{j-1,k-2^{j-1}m} \tag{4.17}$$

since we have assumed a base case, induction proves the formula for all j. Thus we obtain that $\sum_k k^\alpha \Gamma_0^{jk} = 0$ for each j, which implies that $\sum_{j,k} k^\alpha \Gamma_0^{jk} = 0$. □

Now consider the expansion of $\psi'(x)$ (using distribution derivative if necessary, depending on the smoothness of the wavelet system).

$$\psi'(x) = \sum_l \Gamma_{00}^l \varphi(x) + \sum_{jk} \Gamma_{00}^{jk} \psi_{jk}(x)$$

Now we have L vanishing moments, so for $j = 0 \ldots L$, we have

$$\int x^j \psi'(x) = \sum_l \Gamma_{00}^l \int x^j \varphi_l(x)dx + \sum_{jk} \Gamma_{00}^{jk} \int x^j \psi_{jk}(x)dx$$

and

$$\int x^j \psi'(x)dx = -\int j x^{j-1}\psi(x)dx = 0$$

Thus

$$\begin{aligned} 0 &= \sum_l \Gamma_{00}^l \int (x+l)^j \varphi(x)dx \\ 0 &= \sum_l \Gamma_{00}^l \sum_{i=0}^{j} \begin{pmatrix} j \\ i \end{pmatrix} \int x^0 l^{j-i}\varphi(x)dx = \sum_l \Gamma_{00}^l \sum_{i=0}^{j} \begin{pmatrix} j \\ i \end{pmatrix} l^{j-i}\mathrm{Mom}_i(\varphi) \end{aligned}$$

If $j = 0$, then we obtain

$$0 = \left(\sum_l \Gamma_{00}^l\right) \mathrm{Mom}_0\varphi,$$

which imples that $\sum_l \Gamma^l_{00} = 0$, and if $j = 1$, we see that

$$0 = \left(\sum_l l\Gamma^l_{00}\right) \mathrm{Mom}_0\varphi + \left(\sum_l \Gamma^l_{00}\right) \mathrm{Mom}_1\varphi,$$

which implies that

$$\sum_l l\Gamma^l_{00} = 0.$$

Thus, continuing by induction, we find that

$$0 = \sum_l l^j \Gamma^l_{00} \quad \text{for} \quad j = 0, \ldots, L.$$

But

$$\sum_l l^j \Gamma^l_{00} = \sum_l l^j \left(-\Gamma^{00}_l\right) = \sum_l -l^j \Gamma^{0,-l}_0 = 0,$$

which implies that $\sum_l l^j \Gamma^{0,l}_0 = 0$ as required. Thus the hypothesis for Lemma 4.2 is proven, and hence we conclude that part (b) is valid for all L. $\square$

To prove part (a) we proceed by induction on L. We start with the case where $L = 0$. Since $\Gamma^l_0 = -\Gamma^{-l}_0$ (as is easy to verify, see (4.24) below),

$$\sum_l l^\alpha \Gamma^l_0 = 0 \tag{4.18}$$

for all even α. In this case ($L = 0$) we need only consider $0 \leq \alpha \leq 2, \alpha \neq 1$ and this means α is either 0 or 2 which are even, and hence part (a) is verified for $L = 0$. So now we assume (a) is valid for $0 \leq \alpha \leq 2L, \alpha \neq 1$ and we shall use the fact that (b) has already been shown to be true for $0 \leq \alpha \leq L$.

We will need the following lemma relating moments of φ and ψ to the connection coefficients.

Lemma 4.3 *Let p be an arbritrary positive integer, then*

$$\begin{aligned} -p\mathrm{Mom}_{p-1}(\varphi) &= \sum_{i=0}^{p} \binom{p}{i} \sum_l l^i \Gamma^l_0 \mathrm{Mom}_{p-i}(\varphi) \\ &+ 2^{-j(p+\frac{1}{2})} \sum_{i=0}^{p} \binom{p}{i} \sum_{jk} k^i \Gamma^{jk}_0 \mathrm{Mom}_{p-i}(\psi) \end{aligned} \tag{4.19}$$

Proof: Write, for any positive integer p,

$$x^p = \sum_k c_k \varphi_k(x) + \sum_{jk} d_{jk} \psi_{jk}(x), \tag{4.20}$$

where c_k and d_{jk} are the wavelet expansion coefficients. We can calculate the coefficients and obtain:

$$\begin{aligned} c_k &= \int x^p \varphi_k(x)\,dx, \\ &= \int x^p \varphi(x-k)\,dx, \\ &= \int (u+k)^p \varphi(u)\,ddu, \\ &= \sum_{i=0}^{p} \begin{pmatrix} p \\ i \end{pmatrix} k^i u^{p-i} \varphi(u)\,du \\ &= \sum_{i=0}^{p} \begin{pmatrix} p \\ i \end{pmatrix} k^i \mathrm{Mom}_{p-i}(\varphi) \end{aligned} \tag{4.21}$$

By exactly the same type of calculation and using the definition of $\psi_{jk}(x)$ in terms of $\psi(x)$ we obtain

$$d_{jk} = 2^{-j(p+\frac{1}{2})} \sum_{i=0}^{p} \begin{pmatrix} p \\ i \end{pmatrix} k^k \mathrm{Mom}_{p-i}(\psi). \tag{4.22}$$

Now by differentiating (4.20) and multiplying by $\varphi(x)$ and integrating we obtain

$$p \int x^{p-1} \varphi(x)\,dx = \sum_k c_k \int \varphi_k'(x)\varphi(x)\,dx + \sum_{jk} d_{jk} \int \psi_{jk}'(x)\varphi(x)\,dx, \tag{4.23}$$

and we see that (integrating by parts)

$$\Gamma_k^0 := \int \varphi_k'(x)\varphi(x)\,dx = -\int \varphi_k(x)\varphi'(x)\,dx = -\Gamma_0^k, \tag{4.24}$$

and that

$$\Gamma_{jk}^0 = \int \psi_{jk}'(x)\varphi(x)\,dx = -\int \psi_{jk}(x)\varphi'(x)\,dx = -\Gamma_0^{jk}. \tag{4.25}$$

By substituting the expressions in (4.24) and (4.25) into (4.23) and rearranging we obtain the desired relation (4.19). □

Continuing the proof of the inductive step for part (a) we let $p = 2L+1$ and use (4.19) obtaining

$$-(2L+1)\mathrm{Mom}_{2L}(\varphi) = \textstyle\sum_{i=0}^{2L+1} \left(\begin{array}{c} 2L+1 \\ i \end{array} \right) \sum_l l^i \Gamma_0^l \mathrm{Mom}_{2L+1-i}(\varphi)$$

$$+2^{-j(2L+1+\frac{1}{2})} \textstyle\sum_{i=0}^{2L+1} \left(\begin{array}{c} 2L+1 \\ i \end{array} \right) \left(\sum_{jk} k^i \Gamma_0^{jk} \right) \mathrm{Mom}_{2L-i+1}(\psi) \tag{4.26}$$

Now $\mathrm{Mom}_\alpha(\psi) = 0$ for $0 \le \alpha \le L$ by the hypothesis of the theorem, so $2L - i + 1$ must be greater than L for any possibly nonzero term in the first sum of the right hand term of (4.26 which implies that for such terms $i \le L$. But if $i \le L$ then $\sum_{jk} k^i \Gamma_0^{jk} = 0$ by part (b). Thus the right hand term in (4.26) vanishe, and we are left with the identity

$$-(2L+1)\mathrm{Mom}_{2L}(\varphi) = \sum_{i=0}^{2L+1} \left(\begin{array}{c} 2L+1 \\ i \end{array} \right) \left(\sum_l l^i \Gamma_0^l \right) \mathrm{Mom}_{2L+1-i}(\varphi). \tag{4.27}$$

Now by the induction hypothesis for part (a) we have

$$\sum_l l^i \Gamma_0^l = 0 \text{ for } 0 \le i \le 2L, \quad i \ne 1.$$

Moreover, we see from (4.27) for the case where $L = 0$ (using (4.18)),

$$\sum_l l \Gamma_0^l = -1, \tag{4.28}$$

which is one of the required assertions for part (a). In addition we know from the normalization condition for the wavelet system $\sum a_k = 2$ that $\mathrm{Mom}_0(\varphi) = 1$. Substituting these facts into (4.27) we find that

$$\sum_l l^{2L+1} \Gamma_0^l = 0$$

As we remarked above, $\sum_l l^\alpha \Gamma_0^l = 0$ for all even α, so $\sum_l l^{2L+2} \Gamma_0^l = 0$, and hence the induction and the theorem is proven where (4.27) evaluates $\sum l^\alpha \Gamma_0^l$ for the case where $\alpha = 1$. This concludes the proof of Theorem 3.1.

5 Wavelet Differentation

In Section 3 we discussed the Mallat algorithm, and the wavelet interpolation of a given function to a specific scale. One obtains that, for a sufficiently small scale (large J), a function can approximated by the scaling function expansion at that level, using values of the function at the lattice as coefficients for the wavelet expansion. That is,

$$f(x) \approx I_J f(x) = \sum_n 2^{-\frac{J}{2}} f(x_n) \varphi_{Jn}(x)$$

where $\varphi_{Jn}(x) = \varphi(2^J x - n) \cdot 2^{\frac{J}{2}}$, as before, and where $x_n = nh$, for $h = 2^{-J}$. Now the Mallat algorithm allows this sum to be expressed exactly as

$$\sum_n c_n \varphi_n(x) + \sum_{\substack{k \\ 0 \le j \le J-1}} c_{jk} \psi_{jk}(x) = \sum_n 2^{-\frac{J}{2}} f(x_n) \varphi_{Jn}(x)$$

This relationship is exact and the coefficients c_k and c_{jk} are linear combinations of the coefficients $f(x_n)$.

Now by using connection coefficients we may differentiate both sides:

$$\begin{aligned} \left(\sum_n c_n \varphi_n(x) + \sum_{\substack{k \\ 0 \le j \le J-1}} c_{jk} \psi_{jk}(x) \right)' &= \sum_n c_n \left(\sum_l \Gamma^l_n \varphi_l(x) + \sum_{mn_2} \Gamma^{mn_2}_n \psi_{mn_2}(x) \right) \\ &+ \sum_{\substack{k \\ 0 \le j \le J-1}} c_{jk} \left((\sum_l \Gamma^l_{jk} \varphi_l(x) + \sum_{mn_2} \Gamma^{mn_2}_{jk} \psi_{mn_2}(x) \right) \end{aligned}$$

and

$$\left[\sum_n 2^{-\frac{J}{2}} f(x_n) \varphi_{Jn}(x) \right]' = \sum_n 2^{-\frac{J}{2}} f(x_n) \varphi'_{Jn}(x)$$

Now $\{\varphi_{Jn}(x)\}_n$ and $\{\psi_{jk}(x)\}_{j \ge J,k}$ also form a basis for $L^2(\mathbf{R})$, so we may express $\varphi'_{Jn}(x)$ in this basis.

$$\begin{aligned} \langle \varphi'_{Jn}(x) \mid \varphi_{Jj}(x) \rangle &= 2^{2J} \int 2^J \varphi'(2^J x - n) \varphi(2^J x - j) dx \\ &= 2^J \int \varphi'(u - n) \varphi(u - j) du = 2^J \Gamma^j_n, \end{aligned}$$

and

$$\begin{aligned}\langle \varphi'_{Jn}(x) \mid \psi_{jk}(x)\rangle &= \int 2^{J+\frac{J}{2}+\frac{j}{2}}\varphi'(2^J x - n)\psi(2^j x - k)dx \\ &= \int 2^{J+\frac{j}{2}-\frac{J}{2}}\varphi'(u-n)\psi(2^{j-J}u - k)dx \\ &= \int 2^J \varphi'(u-n)\psi_{j-J,k}(u)du = 2^J\Gamma_n^{j-J,k}.\end{aligned}$$

Thus

$$\sum_n 2^{-\frac{J}{2}} f(x_n)\varphi'_{Jn}(x) = \sum_n 2^{-\frac{J}{2}} f(x_n)\left(\sum_l 2^J\Gamma^l_n\varphi_{Jl}(x) + \sum_{rs} 2^J\Gamma_n^{r-J,s}\psi_{rs}(x)\right), \quad r \geq J.$$

Now since the basis elements $\{\psi_{jk}\}_{j\geq J}$ are common to both sides of the equation and they are orthogonal to each of the other terms, their coefficients must be equal. So removing these terms from both sides we obtain

$$\begin{aligned}\sum_n 2^{-\frac{J}{2}} f(x_n)\left(2^J\sum_l \Gamma^l_n\varphi_{Jl}(x)\right) &= \sum_n c_n\left(\sum_l \Gamma^l_n\varphi_l(x) + \sum_{mn_2}\Gamma_n^{mn_2}\psi_{mn_2}(x)\right) \\ &+ \sum_{\substack{k \\ 0\leq j\leq J-1}} c_{jk}\left(\sum_l \Gamma^l_{jk}\varphi_l(x) + \sum_{mn_2}\Gamma_{jk}^{mn_2}\psi_{mn_2}(x)\right).\end{aligned}$$

Thus derivative computations which are performed within a limited number of scales are exactly equivalent to the computation

$$\sum_n 2^{-\frac{J}{2}} f(x_n)(2^J\sum_l \Gamma^l_n\varphi_{Jl}(x))$$

which is performed on the finest scale ($h = 2^{-J}$).

Now

$$\sum_n 2^{-\frac{J}{2}} f(x_n)\left(2^J\sum_l \Gamma^l_n\varphi_{Jl}(x)\right) = \sum_l 2^{-\frac{J}{2}}\varphi_{Jl}(x)\sum_n 2^J f(x_n)\Gamma^l_n$$

So

$$\frac{d}{dx}I_J f(x) = \left[\sum f(x_n)\varphi_{Jn}(x)\right]' = \sum_l 2^{-\frac{J}{2}}\varphi_{Jl}(x)\sum_n f(x_n)2^J\Gamma^l_n.$$

Interpreting this formula as a relationship on *discrete* data points we see that the Γ_n^l's provide a weighting of adjacent points to compute the derivative.

We define the wavelet discrete differentiation operator of level J

$$D_J f(x_l) = \sum_n 2^J \Gamma_n^l f(x_n),$$

where D_J depends on the connection coefficients $\Gamma = \{\Gamma_n^l\}$ (defined in (3.14) of the wavelet system $\{\varphi_k, \psi_{jk}\}$ and on the level J of the sampling of the values of f. We see that the wavelet discrete operator has coefficients which depend only on level zero scaling functions multiplied by the scale factor 2^J.

Now if the coefficients come from a basis with L-vanishing moments, then we have from Theorem 3.1 that

$$\sum_l l^\alpha \Gamma_0^l = 0 \text{ for } 0 \leq \alpha \leq 2L+2, \quad \alpha \neq 1, \ \sum l \Gamma_0^l = -1$$

Consider any polynomial of degree $2L+2$ or less

$$p(x) = \sum_{n=0}^{2L+2} b_n x^n$$

We want to verify that

$$D_J \delta(p(x)) = \delta(D(p(x)),$$

where the discretization δ is given by

$$\delta(f(x) = 2^{-\frac{J}{2}} f(x_k), \quad k \in \mathbf{Z},$$

and where $x_k = kh, h = 2^{-J}$, as before.

By linearity we need only check that

$$D_J \circ \delta(x^n) = \delta \circ D(x^n), \quad \text{for } n \leq 2L+2.$$

We see that

$$\begin{aligned} \delta(x^n)(x_k) &= 2^{-\frac{J}{2}} (x_k)^n, \\ &= 2^{-\frac{J}{2}} h^n k^n, \end{aligned}$$

and similarly,

$$\begin{aligned} \delta \circ D(x^n)(x_k) &= 2^{-\frac{J}{2}} n (x_k)^{n-1} \\ &= 2^{-\frac{J}{2}} n h^{n-1} k^{n-1}. \end{aligned}$$

We calculate, recalling that $2^J = h^{-1}$,

$$\begin{aligned}
D_J \circ \delta(x^n)(x_k) &= 2^J \textstyle\sum_l \Gamma_l^k 2^{-\frac{J}{2}} h^n l^n \\
&= 2^{-\frac{J}{2}} h^{n-1} \textstyle\sum_l \Gamma_l^k l^n \\
&= 2^{-\frac{J}{2}} h^{n-1} \textstyle\sum_l \Gamma_0^{k-l} l^n \\
&= -2^{-\frac{J}{2}} h^{n-1} \textstyle\sum_l \Gamma_0^{l-k} l^n \\
&= -2^{-\frac{J}{2}} h^{n-1} \textstyle\sum_m \Gamma_0^m (m+k)^n \\
&= -2^{-\frac{J}{2}} h^{n-1} \textstyle\sum_m \Gamma_0^m \left(\sum_{l=0}^n \begin{pmatrix} n \\ l \end{pmatrix} m^l k^{n-l} \right) \\
&= -2^{-\frac{J}{2}} h^{n-1} \textstyle\sum_{l=0}^n \begin{pmatrix} n \\ l \end{pmatrix} k^{n-l} \left(\sum_m m^l \Gamma_0^m \right).
\end{aligned}$$

But from Theorem 4.1, we see that $\sum_m m^l \Gamma_0^m = 0$ for $l \neq 1$, and $l \leq 2L+2$, and this same sum is equal to -1 for $l = 1$, hence we find that the sum over l in the above equation collapses to one term leaving

$$D_J \circ \delta(x^n)(x_k) = 2^{-\frac{J}{2}} n h^{n-1} k^{n-1},$$

as desired. □

Thus a wavelet system with L vanishing moments gives the exact derivatives for discrete points of a polynomial of degree $2L+2$. Therefore, by Theorem 2.1, we obtain the following result.

Theorem 5.1 *Let D_J be the level J approximation to the derivative operator D, then if $f \in C^{d+1}$,*

$$Df(x_k) - D_J f(x_k) = O(h^d), \quad \textit{where } h = 2^{-J}.$$

6 Comparison of Wavelet Differentiation with Classical Finite-Difference Operators

Beylkin points out in [1] that the connection coefficients considered for a given wavelet system can be computed as rational functions of the coordinates of the reduced parameter space for all compactly supports wavelet systems (see [18] for a discussion of this parameter space). Those reduced parameter space coordinates which correspond to Daubechies wavelets (maximal vanishing moments) have rational coordinates in the reduced parameter space, and consequently the connection coefficients for Daubechies wavelet systems consists of rational numbers. This is proved independently in [11], which gives a general algorithm for various families of connection coefficients. In particular, this latter paper considers the question of connection coefficients involving integrals of triple products of scaling and wavelet functions and their derivatives, which are critical for nonlinear problems involving quadratic nonlinearities (see for example [17,16]).

The Daubechies wavelet systems of genus 1 and 2 have scaling functions which are not classically differentiable, hence the integral

$$\Gamma_0^k = \int \varphi'(x)\varphi(x-k)\,dx,$$

which is the definition of the connection coefficients, don't exist as classical integral. Nevertheless the algorithms which give values as rational numbers for higher genus ([1,11]) extend as formulas to these cases, and one finds that (see [1]):

$$\begin{aligned} \Gamma_0^k &= \{-\tfrac{1}{2}, 0, \tfrac{1}{2}\}, \quad \text{genus} = 1, \text{ (Haar)} \\ \Gamma_0^k &= \{-\tfrac{1}{12}, \tfrac{2}{3}, 0, -\tfrac{2}{3}, \tfrac{1}{12}\}, \quad \text{genus} = 2, \text{ (Daubechies 4-coefficient)} \end{aligned}$$

There is a classical sequence of finite-difference operators due to Stirling which has the following form (see [3], Formula 7.5.4 in Chapter 7, and see [15] for a more thorough discussion of this sequence and its error analysis). The sequence has the leading terms:

$$hf_0' = \mu\delta f_0 - \frac{1}{6}\mu\delta^3 f_0 + \frac{1}{30}\mu\delta^5 f_0 - \frac{1}{140}\mu\delta^7 f_0 + \ldots, \tag{6.29}$$

where $f_k = f(x_k)$, and where

$$\delta f(x_k) = f\left(x_k + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right), \tag{6.30}$$

$$\mu f(x_k) = \frac{1}{2}\left[f\left(x_k + \frac{h}{2}\right) + f\left(x - \frac{h}{2}\right)\right]. \tag{6.31}$$

$$\tag{6.32}$$

We can easily compute that

$$\mu\delta f_0 = -\frac{1}{2}f(x_0 - h) + \frac{1}{2}f(x_0 + h), \tag{6.33}$$

Daubechies Connection Coefficients, $g = 3$	9-Point Stirling finite-difference
-3.424657e-4	-3.57142857143e-3
-1.4611872e-2	3.80952380952e-2
1.452054e-1	-2e-1
-7.452054e-1	8e-1
0	0
7.452054e-1	-8e-1
-1.452054e-1	2e-1
1.4611872e-2	-3.80952380952e-2
-3.424657e-4	-3.57142857143e-3

Table 6.1: Comparison of Daubechies Connection Coefficients with Stirling's formula for 9 points

$$\mu\delta f_0 - \frac{1}{6}\mu\delta^3 f_0 = \frac{1}{12}f(x_0 - 2h) - \frac{2}{3}f(x_0 - h) \qquad (6.34)$$
$$+\frac{2}{3}f(x_0 + h) - \frac{1}{12}f(x_0 + 2h). \qquad (6.35)$$

Thus we see that the Stirling finite-difference formulas agrees with the Daubechies connection coefficients for genus 1 and 2. For genus 3 however, one finds that the next Stirling formula which can be compared with the Daubechies connection coefficients (the 9-point symmetric difference formula, which includes terms in (6.29) up to $-\frac{1}{140}\mu\delta^7 f_0$) does not agree with the Daubechies connection coefficients for genus = 3. The results are compared in Table 6.

Thus we see that the connection coefficients provide a new series of finite-difference opeartors which are of the symmetric difference type, which agree with the Stirling finite-difference operators initially, but then diverge. This series of operators $D_{J,g}$, where the subscript g denotes the dependence on the genus of the Daubechies wavelet system, have the same approximation properities as the classical Stirling operators, i.e., if we consider the $d+1$-point Stirling operator ∇_{d+1}, then one has that

$$\nabla_{d+1} f(x_k) - Df(x_k) = \mathrm{O}(h^d).$$

By Theorem 5.1 we see that the Daubechies connection coeficients provide automatically in the wavelet context an aproximate derivative

$D_{J,g}$ which satisfies the same estimate as the Stirling operator for the same number of points. In particular, the wavelet approximate derivative at a fixed scale J provides an approximation to the first derivative where the approximation error is of the form:

$$\mathrm{O}(2^{-2gJ}),$$

where g is the genus of the wavelet system. Thus using a wvelet systems to represent functions includes automatically a discrete differentiation with a predetermined rate of accuracy depending on the choice of the system, where the error above is for the special case of Daubechies wavelet systems.

References

[1] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numerical Analysis*, 29(6):1716–1740, December 1992.

[2] Charles Chui. *Wavelet Theory.* Academic Press, Cambridge, MA, 1991.

[3] Germund Dahlquist and Ake Bjorck. *Numerical Methods.* Prentice-Hall, Englewood Cliffs, 1974.

[4] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 41:906–966, 1988.

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

[6] Philip J. Davis. *Interpolation and Approximation.* Blaisdell, New York, 1963.

[7] R. Glowinski, W. Lawton, M. Ravachol, and E. Tenenbaum. Wavelet solution of linear and nonlinear elliptic, parabolic and hyperbolic problems in one dimension. In R. Glowinski and A. Lichnewski, editors, *Proceedings of the Ninth International Conference on Computing Methods in Applied Sciences and Engineering*, pages 55–120, Philadelphia, 1990. SIAM.

[8] R. Gopinath and C. S. Burrus. On the moments of the scaling function ϕ. In *Proceedings of ISCAS '92*, pages 963–966, May 1992.

[9] R. A. Gopinath, W. M. Lawton, and C. S. Burrus. Wavelet-galerkin approximation of linear translation invariant operators. In *Proc. ICASSP-91*, pages 2021–2024. IEEE, 1991.

[10] P. Heller, H. L. Resnikoff, and R. O. Wells, Jr. Wavelet matrices and the representation of discrete functions. In Charles Chui, editor, *Wavelets: A Tutorial*, pages 15–50. Academic Press, Cambridge, MA, 1992.

[11] A. Latto, H. L. Resnikoff, and E. Tenenbaum. The evaluation of connection coefficients of compactly supported wavelets. In Y. Maday, editor, *Proceedings of the French-USA Workshop on Wavelets and Turbulence, June 1991*, New York, 1993. Princeton University, Springer-Verlag. To appear.

[12] W. Lawton, W. Morrell, E. Tenenbaum, and J. Weiss. The wavelet-galerkin method for partial differential equations. Technical Report AD901220, Aware, Inc., 1990.

[13] Wayne M. Lawton. Necessary and sufficient conditions for constructing orthonormal wavelet bases. *J. Math. Phys.*, 32(1):57–61, January 1991.

[14] S. Mallat. Multiresolution approximation and wavelet orthonormal bases of $l^2(\mathbf{r})$. *Trans. Amer. Math. Soc.*, 315:69–87, 1989.

[15] L. M. Milne-Thompson. *The Calculus of Finite Differences*. Macmillan and Co., Ltd., London, 1933.

[16] Sam Qian and John Weiss. Wavelets and the numerical solution of partial differential equations. *J. Comp. Phys.*, 1993. To appear.

[17] J. Weiss. Wavelets and the study of two dimensional turbulence. In Y. Maday, editor, *Proceedings of the French-USA Workshop on Wavelets and Turbulence, June 1991*, New York, 1993. Princeton University, Springer-Verlag. To appear.

[18] R. O. Wells, Jr. Parametrizing smooth compactly supported wavelets. *Trans. Amer. Math. Soc.*, 1992. to appear.

[19] R. O. Wells, Jr. and Xiaodong Zhou. Wavelet interpolation and approximate solutions of elliptic partial differential equations. Technical Report 92-03, Rice University, 1992. Computational Mathematics Laboratory.

[20] R. O. Wells, Jr. and Xiaodong Zhou. Wavelet solutions for the dirichlet problem. Technical Report 92-02, Rice University, 1992. Computational Mathematics Laboratory.