A Remark on Vanishing Moments *

Jun Tian and Raymond O. Wells, Jr.
Computational Mathematics Laboratory, Rice University
Houston, Texas 77005-1892
juntian@rice.edu, wells@rice.edu

Abstract

It is well known that with more vanishing moments on the scaling function and the wavelet function, one may get a better wavelet sampling approximation. In this paper, we will present a quantitative treatment of the above statement. It will be shown that equal vanishing moments distribution between the scaling function and the wavelet function will be an ideal choice for the wavelet sampling approximation. Based on this observation, we introduce biorthogonal Coifman wavelet systems, a family of biorthogonal wavelet systems with the vanishing of moments equally distributed between scaling functions and wavelet functions. These wavelet systems provide good wavelet sampling approximation with exponential decay, and the scaling vectors are all dyadic rational.

1. Introduction

In 1988, I. Daubechies, in her celebrated paper [1], introduced a class of compactly supported orthonormal wavelet systems in general, as well as a family with growing smoothness for large support, the Daubechies wavelet systems. The Daubechies wavelet system put all freedom of parameters into vanishing moments of the wavelet function, or in other words, they are the minimal length wavelet systems with vanishing moments imposed on wavelet functions. The smoothness of the Daubechies wavelet systems increases linearly with the support width. One year later, R. Coifman suggested that it might be worthwhile to construct orthogonal wavelet systems with vanishing moments not only for wavelet functions, but also for scaling functions. Such orthogonal wavelet systems with vanishing moments equally distributed between scaling functions and wavelet functions (which are called *Coiflets* by Daubechies in [2]) are closer to being symmetric than the Daubechies wavelet systems, which suggests that with vanishing moments on scaling functions, we might be able to design more symmetric wavelet systems. This is important since in signal processing, symmetry corresponds to linear phase. A method to construct Coiflets was proposed in [2] but the general existence problem is still open.

In a wavelet sampling approximation, vanishing moments of the scaling function and the wavelet function are actually playing two different roles. Vanishing moments of the wavelet function is related to the wavelet orthogonal projection, while vanishing moments of the scaling function is connected to the wavelet sampling approximation inside the projection subspace. In this paper, we will present a quantitative discussion of these approximation issues. The main result is that for the optimal wavelet sampling approximation, vanishing moments should be equally distributed between the scaling function and the wavelet function. This is true not only for orthogonal wavelet systems, but also for more general wavelet systems (semiorthogonal and biorthogonal).

The paper is organized as follows. We review two basic results of the wavelet orthogonal projection and the wavelet sampling approximation in Section 2 and 3. The wavelet sampling approximation theorem is stated in Section 4. It combines two types of approximation and gives an optimal result on vanishing moments distribution. *Biorthogonal Coifman wavelet systems* is introduced in Section 5 and we outline some properties of this wavelet family. The paper is concluded in Section 6.

2. Wavelet Orthogonal Projection

As usual, in orthonormal wavelet system, we define the *wavelet orthogonal projection* of an $L^2(\mathbf{R})$ function $f(x)$ at the j-th level by

$$P^j(f) \;:=\; \sum_{k\in\mathbf{Z}} \left(\int_{-\infty}^{\infty} f(t)\phi_{j,k}(t)dt \right) \cdot \phi_{j,k}(x)\,,$$

*Copyright 1996 IEEE, in *Proc. of the 30th Asilomar Conference on Signals, Systems, and Computers*, A. Singh, Editor, 983–987, IEEE Computer Society Press, Pacific Grove, CA, 1996. This work was supported in part by Air Force Office of Scientific Research F49620-94-1-0006 and Texas Advanced Technology Program.

where $\phi(x)$ is the scaling function of the orthonormal wavelet system, and $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$, $j, k \in \mathbf{Z}$. The wavelet orthogonal projection is exactly the orthogonal projection of $f(x)$ onto the j-th subspace V_j in the multiresolution analysis of the orthonormal wavelet system [5].

The basic result is that with more vanishing moments on the wavelet function $\psi(x)$ of the orthonormal wavelet system, the difference between the original function $f(x)$ and the orthogonal projection $P^j(f)$ will get smaller and smaller. We denote by $C_0^{N,1}(\mathbf{R})$ the set of compactly supported functions having derivatives of order $\leq N$ and whose N-th derivative is Lipschitz (thus it contains all compactly supported functions having derivatives of order $\leq N+1$).

Theorem 2.1 *In an orthonormal wavelet system with scaling function $\phi(x)$ and wavelet function $\psi(x)$, for $f(x) \in C_0^{N,1}(\mathbf{R})$, define, for $j \in \mathbf{N}$,*

$$P^j(f) \;:=\; \sum_{k\in\mathbf{Z}} \left(\int_{-\infty}^{\infty} f(t)\phi_{j,k}(t)dt \right) \cdot \phi_{j,k}(x)\,,$$

where $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$. If $\psi(x)$ has vanishing moments up to degree N, i.e.,

$$\int_{-\infty}^{\infty} x^n \psi(x)\,dx \;=\; 0\,, \quad \text{for}\, n = 0, 1, \cdots, N\,,$$

then

$$\left|\left| f(x) \;-\; P^j(f) \right|\right|_{L^2} \;\leq C 2^{-j\,(N+1)}\,,$$

where C is a constant, independent of j and N.

This theorem states that the wavelet orthogonal projection converges to the original function with exponential decay as the number of vanishing moments of the wavelet function goes to infinity, or the scale goes to infinity. The proof can be found in [3].

3. Wavelet Sampling Approximation

We define the *wavelet sampling approximation* of an $L^2(\mathbf{R})$ function $f(x)$ at the j-th level by

$$S^j(f) \;:=\; 2^{-j/2} \sum_{k\in\mathbf{Z}} f\left(\frac{k}{2^j}\right) \phi_{j,k}(x)\,.$$

A similar result to Theorem 2.1 above is that imposing vanishing moments on the scaling function $\phi(x)$ will reduce the difference between the wavelet orthogonal projection $P^j(f)$ and the wavelet sampling approximation $S^j(f)$.

Theorem 3.1 *In an orthonormal wavelet system with scaling function $\phi(x)$, for $f(x) \in C_0^{N,1}(\mathbf{R})$, define, for $j \in \mathbf{N}$,*

$$S^j(f) \;:=\; 2^{-j/2} \sum_{k\in\mathbf{Z}} f\left(\frac{k}{2^j}\right) \phi_{j,k}(x)\,,$$

where $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$. If $\phi(x)$ has vanishing moments up to degree N, i.e.,

$$\int_{-\infty}^{\infty} \phi(x)\,dx \;=\; 1\,,$$

$$\int_{-\infty}^{\infty} x^n \phi(x)\,dx \;=\; 0\,, \quad \text{for}\, n = 1, \cdots, N\,,$$

then

$$\left|\left| P^j(f) \;-\; S^j(f) \right|\right|_{L^2} \;\leq C 2^{-j\,(N+1)}\,,$$

where C is a constant, independent of j and N.

This theorem states that the difference between the wavelet orthogonal projection and the wavelet sampling approximation will goes to zero as the number of vanishing moments of the scaling function goes to infinity, or the scale goes to infinity. And the convergence is exponentially fast. The proof can be found in [6,9].

4. A Wavelet Approximation Theorem

The wavelet sampling approximation is what is used in most applications of wavelets, as it is the easiest approximation to compute. (Simply let the sampling values of the given function to be the corresponding wavelet expansion coefficients.) In an orthonormal wavelet system, by definition,

$$\left|\left| f(x) - S^j(f) \right|\right|_{L^2}^2 = \left|\left| f(x) - P^j(f) \right|\right|_{L^2}^2 + \left|\left| P^j(f) - S^j(f) \right|\right|_{L^2}^2 .$$

Thus, to minimize the wavelet sampling approximation error (the left hand side), we need to minimize the two terms of the right hand side to the same degree. So based on Theorem 2.1 and Theorem 3.1, the optimal distribution of vanishing moments between the scaling function and the wavelet function will be the equal distribution, the scaling function and the wavelet function will have the same degrees of vanishing moments. Also it follows that vanishing moments of the wavelet function will reduce the error in the wavelet orthogonal projection, or the distance from the original L^2 function to the projection subspace, while vanishing moments of the scaling function will reduce the difference between the wavelet orthogonal projection and the wavelet sampling approximation, where both of these two are in the projection subspace. The wavelet sampling approximation is illustrated in Figure 1. The illustration shows the nature and relationships of the two types of approximations, the wavelet orthogonal projection P^j and the wavelet sampling approximation S^j. And we formulate the result as the following theorem.

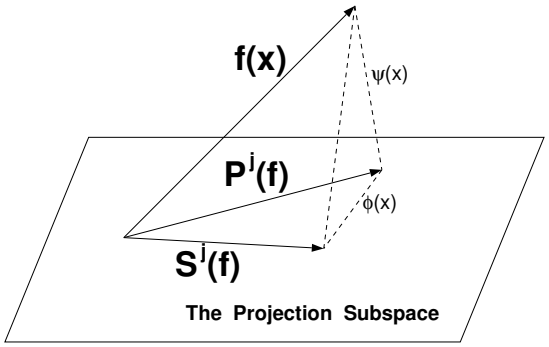

Figure 1. Wavelet Sampling Approximation

Theorem 4.1 *In an orthonormal wavelet system, if both of the scaling function $\phi(x)$ and the wavelet function $\psi(x)$ have vanishing moments up to degree N,*

$$\int_{-\infty}^{\infty} \phi(x)\,dx \;=\; 1\,,$$

$$\int_{-\infty}^{\infty} x^n \phi(x)\,dx \;=\; 0\,, \quad \text{for } n = 1, \cdots, N\,,$$

$$\int_{-\infty}^{\infty} x^n \psi(x)\,dx \;=\; 0\,, \quad \text{for } n = 0, \cdots, N\,,$$

for a function $f(x) \in C_0^{N,1}(\mathbf{R}), j \in \mathbf{N}$, define

$$S^j(f)(x) \;:=\; 2^{-j/2} \sum_{k \in \mathbf{Z}} f\left(\frac{k}{2^j}\right) \phi_{j,k}(x)\,,$$

where $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$. Then

$$||f(x) - S^j(f)(x)||_{L^2} \;\le\; C 2^{-j(N+1)}\,,$$

where C is a constant, independent of j and N.

The proof of Theorem 4.1 is straightforward from Theorem 2.1 and 3.1.

Surprisingly, the orthogonality is not a necessary condition for the above theorem, though it is used in the above argument. All we need is the vanishing moments conditions on the scaling function and the wavelet function. Recall that the vanishing moments conditions of the scaling function and the wavelet function are equivalent to the linear conditions on the scaling vector, as we state in the next lemma [7,4].

Lemma 4.1 *Suppose $\phi(x), \psi(x) \in L^2(\mathbf{R})$ satisfying*

$$\phi(x) \;=\; \sum_{k \in \mathbf{Z}} a_k \phi(2x - k)\,,$$

$$\psi(x) \;=\; \sum_{k \in \mathbf{Z}} (-1)^k a_{-k+1} \phi(2x - k)\,,$$

where $\{a_k\}$ is a finite length sequence, i.e., there exists a positive number K, such that $a_k = 0$ if $|k| > K$. Assume $\phi(x) \in L^1(\mathbf{R})$ and it is normalized

$$\int_{-\infty}^{\infty} \phi(x)\,dx \;=\; 1\,.$$

Then the following two conditions are equivalent,

1. *The vanishing moments of $\phi(x)$ and $\psi(x)$ are both of degree N, i.e.,*

$$\int_{-\infty}^{\infty} x^n \phi(x)\,dx \;=\; 0\,, \quad \text{for } n = 1, \cdots, N\,,$$

$$\int_{-\infty}^{\infty} x^n \psi(x)\,dx \;=\; 0\,, \quad \text{for } n = 0, \cdots, N\,.$$

2. *The sequence $\{a_k\}$ satisfying*

$$\sum_{k \in \mathbf{Z}} (2k)^n a_{2k} \;=\; \sum_{k \in \mathbf{Z}} (2k+1)^n a_{2k+1} \;=\; 0\,,$$

for $n = 1, \cdots, N$, *and*

$$\sum_{k \in \mathbf{Z}} a_{2k} \;=\; \sum_{k \in \mathbf{Z}} a_{2k+1} \;=\; 1\,.$$

Lemma 4.1 converts the vanishing moments conditions on the scaling function and the wavelet function into the linear conditions on the scaling vector. A big advantage of it is that we can design wavelet systems with vanishing moments by solving linear equations of the scaling vector, which is much easier than working directly on the scaling function and the wavelet function. Also by Lemma 4.1 we can present the wavelet approximation theorem with conditions on $\{a_k\}$ only.

Theorem 4.2 (Wavelet Approximation Theorem) *Suppose $\phi(x)$ is an $L^2(\mathbf{R})$ solution of the two-scale difference equation*

$$\phi(x) \;=\; \sum_{k \in \mathbf{Z}} a_k \phi(2x - k)\,,$$

and it is normalized

$$\int_{-\infty}^{\infty} \phi(x)\,dx \;=\; 1\,,$$

where $\{a_k\}$ is a finite length sequence, satisfying

$$\sum_{k \in \mathbf{Z}} (2k)^n a_{2k} \;=\; \sum_{k \in \mathbf{Z}} (2k+1)^n a_{2k+1} \;=\; 0\,,$$

for $n = 1, \cdots, N$, *and*

$$\sum_{k \in \mathbf{Z}} a_{2k} \;=\; \sum_{k \in \mathbf{Z}} a_{2k+1} \;=\; 1 \,.$$

For $f(x) \in C_0^{N,1}(\mathbf{R}), j \in \mathbf{N}$, *define*

$$S^j(f)(x) \;:=\; 2^{-j/2} \sum_{k \in \mathbf{Z}} f\left(\frac{k}{2^j}\right) \phi_{j,k}(x) \;,$$

where $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$. *Then*

$$||f(x) - S^j(f)(x)||_{L^2} \;\leq\; C 2^{-j(N+1)} \,,$$

where C *is a constant, depending only on* $f(x)$ *and* $\{a_k\}$. *If in addition,* $\phi(x) \in C^m(\mathbf{R})$, *where* $m \in \mathbf{Z}, 0 \leq m \leq N$, *then*

$$||f(x) - S^j(f)(x)||_{H^m} \;\leq\; C 2^{-j(N+1-m)} \,,$$

where H^m *is the Sobolev space of* L^2 *functions with weak derivatives of order* $\leq m$, *and* C *is a constant, depending only on* $f(x)$ *and* $\{a_k\}$.

The wavelet approximation theorem doesn't require the orthogonality condition, thus it is true for general wavelet systems (semiorthogonal and biorthogonal). The proof of Theorem 4.2 is based on the following Strang-Fix condition [8].

Lemma 4.2 *Assume the same conditions as in Theorem 4.2, then we have*

$$\sum_{k \in \mathbf{Z}} \phi(x-k) \;=\; 1 \,,$$

and

$$\sum_{k \in \mathbf{Z}} (x-k)^n \phi(x-k) \;=\; 0 \,, \quad \text{for } n = 1, \cdots, N \,.$$

Space limits preclude our reproducing the proof of Theorem 4.2; however it can be found in [6,9].

5. Biorthogonal Coifman Wavelet Systems

For a biorthogonal wavelet system, in the decomposition process, based on Theorem 4.2, we would like to set the analysis scaling vector to have vanishing moments up to some degree, then we can sample on dyadic rationals and take these values as the wavelet expansion coefficients. Especially we can take the sample values as the discrete wavelet transform coefficients at the starting level, apply the Mallat Algorithm [5] on these sample values, and analyze the data (compression, denoising, etc). In the reconstruction process, again, based on Theorem 4.2, if we set the synthesis scaling vector to have vanishing moments up to some degree, we can take the inverse discrete wavelet transform coefficients as the sample values on dyadic rationals and reconstruct the original data. This observation led to the introduction of *biorthogonal Coifman wavelet systems* [9].

Biorthogonal Coifman wavelet systems are compactly supported wavelet systems with the vanishing of moments equally distributed among the analysis wavelet function, the synthesis scaling function, and the synthesis wavelet function. By the perfect reconstruction condition, it can be proved that the analysis scaling function also has vanishing moments up to the same degree of the other three. So a fundamental result for biorthogonal Coifman wavelet systems is that if we use the sample values of a discretized function as scaling function (either analysis or synthesis scaling function) coefficients at a fine scale, then the resulting wavelet series approximates the underlying function with exponentially increasing accuracy.

Using a time domain design method, the closed form solutions of the minimum length biorthogonal Coifman wavelet systems of all degrees are obtained [9].

Theorem 5.1 *The minimum length biorthogonal Coifman wavelet systems of degree* N *have the analysis scaling vector* $\{a_k\}$ *and the synthesis scaling vector* $\{\tilde{a}_k\}$ *of the form*

$$\tilde{a}_0 \;=\; 1 \,, \quad \tilde{a}_{2k} \;=\; 0 \,\text{when}\, k \neq 0 \,,$$

- *if* N *is even,* $N = 2n$,

$$\tilde{a}_{2k+1} \;=\; \frac{(-1)^k(2n+1)}{2^{4n-1}(2k+1)} \left(\begin{array}{c} 2n-1 \\ n-1 \end{array} \right) \left(\begin{array}{c} 2n \\ n+k \end{array} \right)$$

- *if* N *is odd,* $N = 2n-1$,

$$\tilde{a}_{2k+1} \;=\; \frac{(-1)^k(2n-1)}{2^{4n-3}(2k+1)} \left(\begin{array}{c} 2n-2 \\ n-1 \end{array} \right) \left(\begin{array}{c} 2n-1 \\ n+k \end{array} \right)$$

$$a_{2k+1} \;=\; \tilde{a}_{2k+1} \,,$$

$$a_{2k} \;=\; 2\delta_{0,k} \;-\; \sum_{l \in \mathbf{Z}} \tilde{a}_{2l+1} \tilde{a}_{2l+1-2k} \,.$$

The scaling vectors of the minimum length biorthogonal Coifman wavelet systems with degrees $N = 0, 1, 2, 3$ and 4 are listed in Table 1.

As it can be seen in Table 1, the scaling vectors $\{a_k\}$ and $\{\tilde{a}_k\}$ are all dyadic rational, i.e., all the nonzero elements in the scaling vectors are of the form $(2p+1)/2^q$, for some integers p and q. Actually this assertion is true for scaling vectors of the minimum length biorthgonal Coifman wavelet systems of all degrees (see proof in [9]). Thus we have obtained a family of biorthogonal wavelet systems on which we can implement a very fast multiplication-free discrete

Table 1. Biorthogonal Coifman Wavelet Systems

N	a_k	$\tilde{a}_k$
$N=0$	a_0 = 1	$\tilde{a}_0$ = 1
	a_1 = 1	$\tilde{a}_1$ = 1
$N=1$	a_{-2} = -1/4	
	a_{-1} = 1/2	$\tilde{a}_{-1}$ = 1/2
	a_0 = 3/2	$\tilde{a}_0$ = 1
	a_1 = 1/2	$\tilde{a}_1$ = 1/2
	a_2 = -1/4	
$N=2$	a_{-4} = 3/64	
	a_{-3} = 0	
	a_{-2} = -3/16	
	a_{-1} = 3/8	$\tilde{a}_{-1}$ = 3/8
	a_0 = 41/32	$\tilde{a}_0$ = 1
	a_1 = 3/4	$\tilde{a}_1$ = 3/4
	a_2 = -3/16	$\tilde{a}_2$ = 0
	a_3 = -1/8	$\tilde{a}_3$ = -1/8
	a_4 = 3/64	
$N=3$	a_{-6} = -1/256	
	a_{-5} = 0	
	a_{-4} = 9/128	
	a_{-3} = -1/16	$\tilde{a}_{-3}$ = -1/16
	a_{-2} = -63/256	$\tilde{a}_{-2}$ = 0
	a_{-1} = 9/16	$\tilde{a}_{-1}$ = 9/16
	a_0 = 87/64	$\tilde{a}_0$ = 1
	a_1 = 9/16	$\tilde{a}_1$ = 9/16
	a_2 = -63/256	$\tilde{a}_2$ = 0
	a_3 = -1/16	$\tilde{a}_3$ = -1/16
	a_4 = 9/128	
	a_5 = 0	
	a_6 = -1/256	
$N=4$	a_{-8} = 15/16384	
	a_{-7} = 0	
	a_{-6} = -35/2048	
	a_{-5} = 0	
	a_{-4} = 345/4096	
	a_{-3} = -5/128	$\tilde{a}_{-3}$ = -5/128
	a_{-2} = -405/2048	$\tilde{a}_{-2}$ = 0
	a_{-1} = 15/32	$\tilde{a}_{-1}$ = 15/32
	a_0 = 10317/8192	$\tilde{a}_0$ = 1
	a_1 = 45/64	$\tilde{a}_1$ = 45/64
	a_2 = -405/2048	$\tilde{a}_2$ = 0
	a_3 = -5/32	$\tilde{a}_3$ = -5/32
	a_4 = 345/4096	$\tilde{a}_4$ = 0
	a_5 = 3/128	$\tilde{a}_5$ = 3/128
	a_6 = -35/2048	
	a_7 = 0	
	a_8 = 15/16384	

wavelet transform, which consists of only addition and shift operations, on digital computers.

In addition, biorthogonal Coifman wavelet systems converge to the *Sinc wavelet system*, which has the sinc function as its scaling function. Half of the members (more precisely, those with odd degrees) of the family are symmetric. For more details of biorthogonal Coifman wavelet systems, we refer to [9].

6. Conclusions

In this paper we have shown that vanishing moments not only give the smoothness for a wavelet system, but also provide a good wavelet sampling approximation with exponential decay. In addition it may provide us a symmetric wavelet system (in a biorthogonal setting, for example), which will be more important that the sampling approximation in certain applications. And vanishing moments of the scaling function and vanishing moments of the wavelet function are controlling two different approximation errors in the wavelet sampling approximation. Biorthogonal Coifman wavelet systems, which are optimal in the sense of vanishing moments distribution, have lots of interesting properties, which will be quite useful in practice.

References

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

[2] I. Daubechies. Orthonormal bases of compactly supported wavelets II. Variations on a theme. *SIAM J. Math. Anal.*, 24(2):499–519, 1993.

[3] 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*, Philadelphia, 1990. SIAM.

[4] P. N. Heller. Rank m wavelet matrices with n vanishing moments. *SIAM J. Matr. Anal.*, 16:502–518, 1995.

[5] S. G. Mallat. Multiresolution approximation and wavelet orthonormal bases of $L^2(\mathbf{R})$. *Trans. AMS*, 315(1):69–87, Sept. 1989.

[6] H. L. Resnikoff and R. O. Wells, Jr. *Wavelet Analysis and the Scalable Structure of Information*. Springer-Verlag, New York, 1996. To appear.

[7] P. Steffen, P. N. Heller, R. A. Gopinath, and C. S. Burrus. Theory of regular m-band wavelet bases. *IEEE Trans. on Signal Processing*, 41(12), Dec. 1993.

[8] G. Strang. Wavelets and dilation equations: A brief introduction. *SIAM Review*, 31:614–627, 1989.

[9] J. Tian. *The Mathematical Theory and Applications of Biorthogonal Coifman Wavelet Systems*. PhD thesis, Rice University, Houston, Texas, Feb. 1996.