ON THE MOMENTS OF THE SCALING FUNCTION ψ_0

R.A.Gopinath   C.S.Burrus

Department of Electrical and Computer Engineering
Rice University, Houston, TX-77251

ABSTRACT

This paper derives relationships between the moments of the scaling function $\psi_0(t)$ associated with multiplicity M, K-regular, compactly supported, orthonormal wavelet bases [6, 5], that are extensions of the multiplicity 2, K-regular orthonormal wavelet bases constructed by Daubechies [2]. One such relationship is that the square of the first moment of the scaling function $(\psi_0(t))$ is equal to its second moment. This relationship is used to show that uniform sample values of a function provides a third order approximation of its scaling function expansion coefficients. For the special case of $M = 2$, the results in this paper have been reported earlier [3].

1. INTRODUCTION

In this paper we derive relationships between the moments of the scaling function $\psi_0(t)$ associated with the compactly supported, multiplicity M, K regular, orthonormal wavelet bases. In particular, we show that the square of the first moment of ψ_0 is the second moment of ψ_0. Hence samples of a function accurately represent its expansion coefficients in terms of the scaling function. The scaling function $\psi_0(t)$ is determined by a sequence $h_0(k)$ (the *scaling vector*), constructed through a spectral factorization process in the K-regular case. However, it is interesting that the moment relationships only depend on and can be derived from the squared magnitude, $|H_0(\omega)|^2$, of the Fourier Transform of $h_0(k)$.

2. RELATIONSHIP BETWEEN MOMENTS OF h_l AND ψ_l

Multiplicity M, K-regular compactly supported, orthonormal wavelet bases are characterized by a scaling vector $h_0(k)$ and $M-1$ wavelet vectors $h_l(k)$, $l = 1, 2, \ldots, M-1$ all of finite length N, that satisfy the equations,

$$\sum_k h_l(k) h_m(k + Mn) = \delta(l-m)\delta(n) \tag{1}$$

for $l = 1, 2, \ldots, M-1$, and

$$\sum_k h_l(k) = \sqrt{M}\delta(l). \tag{2}$$

In terms of the scaling vector h_0, the scaling function is defined by the scale recursive formula,

$$\psi_0(t) = \sqrt{M} \sum_{k=0}^{N-1} h_0(k) \psi_0(Mt - k). \tag{3}$$

Similarly, from the wavelet vectors, and the scaling function, the wavelets $\psi_l(t)$ are defined as

$$\psi_l(t) = \sqrt{M} \sum_{k=0}^{N-1} h_l(k) \psi_0(Mt - k). \tag{4}$$

If the scaling function is normalized to have unit energy, then it can also be shown that

$$\int_{\mathbb{R}} \psi_0(x)\, dx = 1. \tag{5}$$

In wavelet analysis of signals, it is necessary to take inner products of a given signal $f(t)$ with the scaling function and wavelets appropriately scaled. If $f(t)$ is approximated from its samples (only the samples are assumed to be available) using a local polynomial approximation, then the inner-products can be obtained by a discrete convolution using the moments of the scaling function/wavelets [6]. Thus it is important to compute the moments of the scaling function and the wavelets. Even though it is not possible to obtain the scaling function or wavelets exactly (analytically), it turns out that the moments of $\psi_0(t)$ and $\psi_l(t)$ can be computed exactly. This is because the moments of the scaling function and wavelets are related to the discrete moments of the scaling vectors and wavelet vectors respectively. The latter are easy to compute, since these vectors are known explicitly.

Lemma 1 *Define for all $l = 0, 1, \ldots, M-1$*

$$m_{l,n} = \int_{\mathbb{R}} t^n \psi_l(t)\, dt \tag{6}$$

$$\mu_{l,n} = \sum_{k=0}^{N-1} k^n h_l(k) \tag{7}$$

Then we have the following relationship between these moments

$$m_{l,n} = \frac{1}{M^{n+\frac{1}{2}}} \sum_{i=0}^{n} \binom{n}{i} \mu_{l,i} m_{0,n-i} \tag{8}$$

Proof: Invoking Eqn. 3 for $\psi_0(x)$ in Eqn. 6, we get

$$\begin{aligned} m_{l,n} &= \sqrt{M} \sum_k h_l(k) \int_{\mathbb{R}} x^n \psi_0(Mx - k)\, dx \qquad (9) \\ &= \frac{1}{M^{\frac{1}{2}}} \sum_k h_0(k) \int_{\mathbb{R}} \left(\frac{x+k}{M}\right)^n \psi_0(x)\, dx \end{aligned}$$

$$= \frac{1}{M^{n+\frac{1}{2}}} \sum_k h_l(k) \int_{\mathbb{R}} \sum_{i=0}^{n} \binom{n}{i} x^{n-i} k^i \psi_0(x)\, dx$$
$$= \frac{1}{M^{n+\frac{1}{2}}} \sum_{i=0}^{n} \binom{n}{i} \left[\sum_k h_l(k) k^i \right] m_{0,n-i}$$
$$= \frac{1}{M^{n+\frac{1}{2}}} \sum_{i=0}^{n} \binom{n}{i} \mu_{l,i} m_{0,n-i}$$
□

From Eqn. 8, we have a recursive formula to obtain the moments of the scaling/wavelet function from the moments of the corresponding scaling/wavelet vector and the lower order moments of the scaling function. From now on we consider only the moments of the scaling function and scaling vector. Clearly Eqn. 5 and Eqn. 2 implies that $m_{0,0} = 1$ and $\mu_{0,0} = \sqrt{M}$. Besides, Eqn. 8 implies that $m_{0,1} = \mu_{0,1}/\sqrt{M}$. It is more convenient to scale the discrete moments of the sequences h_l by defining

$$d_{l,n} = \mu_{l,n}/\sqrt{M}. \tag{10}$$

Then Eqn. 8 becomes

$$m_{l,n} = \frac{1}{M^n} \sum_{i=0}^{n} \binom{n}{i} d_{l,i} m_{0,n-i} \tag{11}$$

An important consequence of Eqn. 11 is that if all the moments $m_{0,n}$ are given by $m_{0,n} = (m_{0,1})^n$, then the continuous moments are precisely equal to $d_{0,n}/(M-1)^n$.

Lemma 2 *For a given ψ_0 and integer $k \geq 0$, the following statements are equivalent:*

1. *For all n, $0 \leq n \leq k$, $m_{0,n} = (m_{0,1})^n$*
2. *For all n, $0 \leq n \leq k$, $d_{0,n} = (d_{0,1})^n$*

Moreover if either condition is satisfied, then

$$d_{0,n} = (d_{0,1})^n = (M-1)^n (m_{0,1})^n$$

for all non-negative n.

Proof: Clearly $d_{0,0} = \mu_{0,0}/\sqrt{M} = 1 = m_{0,0}$. Now from Eqn. 11 for $n = 1$ for $l = 0$ we get, $M m_{0,1} = d_{0,0} m_{0,1} + d_{0,1} m_{0,0}$, and therefore $d_{0,1} = (M-1) m_{0,1}$. Now assume that for all n, $0 \leq n \leq k, d_{0,n} = (d_{0,1})^n$. By the induction hypothesis, $m_{0,n} = (m_{0,1})^n$ and $d_{0,n} = (M-1)^n (m_{0,1})^n$. Now invoking Eqn. 11 for $l = 0$ and $n+1$, and using the fact that $d_{0,n+1} = (d_{0,1})^{n+1}$ we get

$$M^{n+1} m_{0,n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} d_{0,i} m_{0,n+1-i} \tag{12}$$
$$= \sum_{i=0}^{n+1} \binom{n+1}{i} (d_{0,1})^i (m_{0,1})^{n+1-i} + m_{0,n+1} - m_{0,1}^{n+1}$$
$$= (d_{0,1} + m_{0,1})^{n+1} + m_{0,n+1} - m_{0,1}^{n+1}$$
$$= M^{n+1} (m_{0,1})^{n+1} + m_{0,n+1} - (m_{0,1})^{n+1}$$

and hence the result follows. The converse also follows similarly. In particular, note that $m_{0,2} = m_{0,1}^2$ if and only if $d_{0,2} = d_{0,1}^2$. □

The Lemma has the following interesting consequence. If we shift $\psi_0(t)$ to its center of mass, then for the new function $m_{0,1} = 0$. This would imply that for the new function, for all $n \leq k$ in the Lemma, $m_{0,n} = m_{0,1}^n = 0$!

3. THE FOURIER TRANSFORM AND DISCRETE MOMENTS

In this section we show the relationship between the discrete moments of a sequence and the magnitude squared of its Fourier transform. We then show that the Fourier transform of the scaling vector satisfies certain properties and this allows us to conclude our main result about the moments of the scaling vector.

Let $H_0(\omega)$ be the Fourier transform of the sequence $h_0(k)$.

$$H_0(\omega) = \sum_{k=0}^{N-1} h_0(k) e^{-i\omega k} \tag{13}$$

Then the magnitude squared of the Fourier transform of $h_0(k)$ is given by

$$|H_0(\omega)|^2 = \sum_{k,l} h_0(k) h_0(l) e^{i(k-l)\omega}. \tag{14}$$

Differentiating n times on both sides with respect to ω, and evaluating at $\omega = 0$ we get

$$\left[\left(\frac{d}{d\omega} \right)^n |H_0(\omega)|^2 \right]_{\omega=0} \stackrel{\text{def}}{=} a(n) = i^n \sum_{k,l} h_0(k) h_0(l) (k-l)^n. \tag{15}$$

For odd n from the symmetry it is clear that the right hand side evaluates to zero. Hence all the odd derivatives of $|H_0(\omega)|^2$ are zero. The even derivatives are related to the the discrete moments of h_0. Indeed, for $n = 2p$,

$$a(2p) = i^{2p} \sum_{k,l} h_0(k) h_0(l) \sum_{j=0}^{2p} \binom{2p}{j} k^{(2p-j)} (-l)^j$$
$$= (-1)^p \sum_{j=0}^{2p} \binom{2p}{j} \left(\sum_k h_0(k) k^{2p-j} \right) \left(\sum_l h_0(l) (-l)^j \right)$$
$$= (-1)^p \sum_{j=0}^{2p} \binom{2p}{j} (-1)^j \mu_{0,2p-j} \mu_{0,j}. \tag{16}$$

Now, for a general multiplicity M, K-regular orthonormal wavelet basis, it can be shown ([5]) that the Fourier transform of the scaling vector satisfies

$$|H_0(\omega)|^2 = M + O(|\omega|^{2K}) \tag{17}$$

for ω close to the origin. From this fact and the previous discussion it readily follows that,

$$\left[\left(\frac{d}{d\omega} \right)^{2p} |H_0(\omega)|^2 \right]_{\omega=0} = M\delta(p) \tag{18}$$

for $p = 0, 1, 2, \ldots, K-1$ provided $K \geq 1$. Notice that for $p = 0$, Eqn. 2 was used. Thus we have a set of K equations relating the first $2K-1$ moments of h_0. It is obvious that this information is far from sufficient to know all of the first $2K-1$ moments. Notice that for $K = 1$, the maximum value of p is 0 and the relationship merely states that $|H_0(0)|^2 = M$. For $K \geq 2$, the maximum value of p is greater than or equal to 1. In particular, for $p = 1$ we obtain the main result in this paper. For $p = 1$ we have

$$2\mu_{0,2}\mu_{0,0} - 2\mu_{0,1}\mu_{0,1} = 0 \tag{19}$$

$$\mu_{0,2} = \mu_{0,1}^2/\sqrt{M} \tag{20}$$

$$d_{0,2}^2 = d_{0,1}^2 \tag{21}$$

Now using Lemma.3, we get the following theorem.

Theorem 1 *For compactly supported, multiplicity M, K-regular, orthonormal wavelet bases with $K \geq 2$ (i.e, except for the Haar case), the moments of the scaling function satisfy $m_{0,2} = (m_{0,1})^2$.*

This result was observed by W.M.Lawton in his numerical investigations with wavelets in the multiplicity 2 case [7] and was proved at that time by one of us [3].

We now show how Eqn. 17 is true in the multiplicity 2 case. In the multiplicity M case the result is relatively more difficult and the reader is referred to [5]. Let h_0 be a multiplicity 2, K regular scaling vector. Then from Eqn. 1 for $M = 2$ it follows that

$$|H_0(\omega)|^2 + |H_0(\omega+\pi)|^2 = 2 \tag{22}$$

In [2], Daubechies shows that $|H_0(\omega)|^2/2$ is necessarily of the form (where $c(\cdot)$ and $s(\cdot)$ denotes the cosine and sine functions respectively),

$$\left(c(\frac{\omega}{2})\right)^{2K}\left[\sum_k \binom{K-1+k}{k}\left(s(\frac{\omega}{2})\right)^{2k} + \left(s(\frac{\omega}{2})\right)^{2K} R(c(\omega))\right] \tag{23}$$

where R is an *odd* polynomial satisfying a certain growth condition. If $R = 0$, then N, the length of the sequence h_0, is equal to $2K$. Also, from Eqn. 22 we have

$$|H_0(\omega)|^2/2 = 1 - |H_0(\omega+\pi)|^2/2.$$

That is, $|H_0(\omega)|^2/2$ is given by

$$1-\left(s(\frac{\omega}{2})\right)^{2K}\left[\sum_k \binom{K-1+k}{k}\left(c(\frac{\omega}{2})\right)^{2k} + \left(c(\frac{\omega}{2})\right)^{2K} R(-c(\omega))\right] \tag{24}$$

and hence around $\omega = 0$,

$$|H_0(\omega)|^2 = 2 + O(|\omega|^{2K}) \tag{25}$$

This idea of using the behavior of $|H_0(\omega)|^2$ in the vicinity of the origin was suggested by I. Daubechies [1]. Tables. 1-3 give the moments of the scaling functions and scaling vectors for different values of regularity K. The relationships between $m_{0,n}$ and $d_{0,n}$, and Eqn. 16 can be verified from these tables.

4. PROJECTION ONTO THE FINEST SCALE $W_{0,J}$

Let $\psi_{l,j,k} = M^{j/2}\psi_l(M^j t - k)$ and let $W_{l,j} = Span\,\{\psi_{l,j,k}\}$. Then, there is a natural multiresolution analysis [6],

$$\ldots W_{0,j} \subset W_{0,j+1} \ldots \subset L^2(\mathbb{R}) \tag{26}$$

$$W_{0,j+1} = \bigoplus_{l=0}^{M-1} W_{l,j} \tag{27}$$

Now consider the computation of the wavelet transform of a function $f(t)$ from its uniformly spaced samples. Starting with the scaling expansion coefficients at any given scale,

Table 1. The moments of $\psi_0(t)$

$\boxed{M = 2 \text{ and } N = MK}$

N	k	$m_{0,k}$	$d_{0,k}$
4	0	1.0000000e+00	1.0000000e+00
	1	6.3397460e-01	6.3397460e-01
	2	4.0192379e-01	4.0192379e-01
	3	1.3109156e-01	-6.1121593e-01
	4	-3.0219333e-01	-4.2846097e+00
	5	-1.0658728e+00	-1.6572740e+01
6	0	1.0000000e+00	1.0000000e+00
	1	8.1740117e-01	8.1740117e-01
	2	6.6814467e-01	6.6814467e-01
	3	4.4546004e-01	-1.5863308e-01
	4	1.1722635e-01	-1.8579194e+00
	5	-4.6651091e-02	3.7516197e+00
8	0	1.0000000e+00	1.0000000e+00
	1	1.0053932e+00	1.0053932e+00
	2	1.0108155e+00	1.0108155e+00
	3	9.0736037e-01	2.5392023e-01
	4	5.8377181e-01	-2.0440853e+00
	5	6.3077524e-02	-2.4420547e+00
10	0	1.0000000e+00	1.0000000e+00
	1	1.1939080e+00	1.1939080e+00
	2	1.4254164e+00	1.4254164e+00
	3	1.5802598e+00	8.5092254e-01
	4	1.4513041e+00	-2.0317424e+00
	5	8.1371053e-01	-5.9644946e+00

Table 2. The moments of $\psi_0(t)$

$\boxed{M = 3 \text{ and } N = MK}$

N	k	$m_{0,k}$	$d_{0,k}$
6	0	1.0000000e+00	1.0000000e+00
	1	6.2084713e-01	1.2416943e+00
	2	3.8545116e-01	1.5418046e+00
	3	1.1024925e-01	-1.4410320e+00
	4	-3.3859274e-01	-2.7622103e+01
9	0	1.0000000e+00	1.0000000e+00
	1	7.8515128e-01	1.5703026e+00
	2	6.1646253e-01	2.4658501e+00
	3	3.8154196e-01	1.2077966e+00
	4	5.8194455e-02	-1.0654826e+01
12	0	1.0000000e+00	1.0000000e+00
	1	9.5286399e-01	1.9057280e+00
	2	9.0794979e-01	3.6317991e+00
	3	7.5580853e-01	4.0782740e+00
	4	4.0761249e-01	-8.4815717e+00

Table 3. The moments of $\psi_0(t)$

$M = 5$ and $N = MK$

N	k	$m_{0,k}$	$d_{0,k}$
10	0	1.0000000e+00	1.0000000e+00
	1	6.0961180e-01	2.4384472e+00
	2	3.7162654e-01	5.9460247e+00
	3	9.3544517e-02	-1.9933553e+00
	4	-3.6313857e-01	-2.3590840e+02
15	0	1.0000000e+00	1.0000000e+00
	1	7.5803488e-01	3.0321395e+00
	2	5.7461687e-01	9.1938700e+00
	3	3.3138863e-01	1.4957413e+01
	4	1.4262918e-02	-7.2169885e+01
20	0	1.0000000e+00	1.0000000e+00
	1	9.0920717e-01	3.6368287e+00
	2	8.2665767e-01	1.3226523e+01
	3	6.4125671e-01	3.4419647e+01
	4	2.8205206e-01	-2.4109276e+01

there are efficient numerical algorithms to compute the scaling and wavelet coefficients at all coarser scales using cascaded unitary filter banks [6, 4]. Hence usually f is projected onto the finest scale of interest, say $W_{0,J}$, and then the wavelet coefficients at all coarser scales are then computed from it.

Since only the samples of f, which we assume without loss of generality to be at a spacing of M^{-J}, are given, some assumption about the smoothness of f must be made in order to approximate/interpolate f from sample data. One could use polynomial or spline interpolation/approximation. Since the basis functions for $W_{0,J}$, i.e $\psi_{0,J,k}$ are compactly supported and concentrated around $M^{-J}k$ a useful approach is to use a Taylor series approximation of f in the neighborhood of $M^{-J}k$ in order to compute $\langle \psi_{0,J,k}, f\rangle$. Since the support of ψ_0 is $[0, \frac{N-1}{M-1}]$, the support of $\psi_{0,J,k}$ is $[M^{-J}k, M^{-J}(k + \frac{N-1}{M-1})]$. Now consider the Taylor series expansion of $f(t)$ around the center of mass of $\psi_{0,J,k}$. Then $f(M^{-J}(k+t))$ is

$$f\left(\frac{k+m_{0,1}}{M^J}\right) + \left(\frac{t-m_{0,1}}{M^J}\right) f^{(1)}\left(\frac{k+m_{0,1}}{M^J}\right) + \ldots$$

Now if the scaling vector h_0 is K-regular with $K \geq 2$, then from Theorem 1 $m_{0,2} = m_{0,1}^2$, and hence it follows that,

$$\begin{aligned}
\langle \psi_{0,J,k}, f\rangle &= \int_{\mathbb{R}} f(t) M^{J/2} \psi_0(M^J t - k)\, dt \\
&= M^{-J/2} \left\{ \int_R f\left(\frac{t+k}{M^J}\right) \psi_0(t)\, dt \right\} \\
&= M^{-J/2} \left\{ f\left(\frac{k+m_{0,1}}{M^J}\right) + O\left((1/M^J)^3\right) \right\}
\end{aligned}$$

The last step is obtained by invoking the Taylor series expansion and using the relationships between the moments. Hence the sample of a scaled version of f, namely $f(M^{-J}(k+m_{0,1}))$ themselves give a third order approximation to the scaling expansion coefficients. Increasing the sampling rate by a factor of M, reduces the error by a factor of M^3! This is precisely why in most applications the samples of the function provide a good approximation. It can also be argued that for sufficiently large J the scaling function approaches the Dirac delta measure and hence the samples of the function can be considered to be the scaling coefficients.

5. CONCLUSION

This paper derives a set of relationships among the moments of the scaling function. One of the relationships is helpful in explaining why the samples of a function themselves form a very good approximation to the scaling expansion coefficients; more precisely, the approximation is third order. This raises an interesting question: can the extra degree of freedom associated with construction of a multiplicity M, K regular non-minimal length scaling vector (for example in the multiplicity 2 case, there is the choice of an *odd* polynomial R) be exploited to make higher moments of the scaling vector be powers of the first moment? Clearly, these K-regular scaling vectors would be longer than the minimal length, but the extra accuracy in the representation of a function in the wavelet basis associated with that scaling vector may more than compensate for the increase in the support of the resulting scaling function. This would have implications in the numerical approximation of scaling coefficients which matters in any application using wavelets, be it in signal processing or in the solution of differential equations.

Acknowledgments

This material is based upon work supported by AFOSR under grant number 90-0334 which was funded by DARPA. The authors extend their thanks to W.M.Lawton, H.L.Resnikoff, and J.E.Odegard for helpful suggestions and comments.

REFERENCES

[1] I. Daubechies. Private Communication.

[2] I. Daubechies. Orthonormal Bases of Compactly Supported Wavelets. *Communications in Pure and Applied Mathematics*, (41):909–996, 1988.

[3] R. A. Gopinath. Wavelet Transforms and Time-Scale Analysis of Signals. Master's thesis, Rice University, Houston, TX-77251, 1990.

[4] R. A. Gopinath and C. S. Burrus. Efficient computation of the wavelet transforms. In *Proceedings of ICASSP, '90, Albuquerque,New Mexico*. IEEE, 1990.

[5] R. A. Gopinath and C. S. Burrus. State-space Approach to Multiplicity M Orthonormal Wavelet Bases. Technical Report CML TR91-22, Computational Mathematics Laboratory, Rice University, 1991. Submitted to IEEE Trans. on SP.

[6] R. A. Gopinath and C. S. Burrus. Wavelets and Filter Banks. In C. K. Chui, editor, *Wavelets: A Tutorial in Theory and Applications*, pages 603–654. Academic Press, 1992.

[7] W. M. Lawton. Private communication.