THE DESIGN OF MAXIMALLY SMOOTH WAVELETS

Markus Lang
MeVis
Kitzbuehler Str. 1
D-28359 Bremen, Germany
email: mlang@mevis.uni-bremen.de

Peter N. Heller
Aware, Inc.
One Oak Park
Bedford, MA 01730
heller@aware.com

ABSTRACT

A method for designing more regular (more differentiable) scaling functions than the Daubechies wavelets is presented. The approach is to numerically minimize the Hölder exponent (a measure of smoothness) subject to constraints on the autocorrelation sequence of the scaling filter. The "maximally smooth" wavelets which result have a Hölder exponent which grows faster than that of the Daubechies orthogonal wavelets, as the filter length increases. This is achieved by relaxing some of the vanishing moment conditions.

1. INTRODUCTION

Smoothness of orthogonal scaling and wavelet functions is important both for theoretical and practical reasons. It is unsatisfactory that no general method exists that minimizes such a basic property as the regularity of the scaling function, although many methods are known for other design criteria such as maximum number of vanishing moments (Daubechies filters), minimization of some frequency domain error [9, 13], or mixtures of both [12, 8].

Wavelet smoothness is also important from a practical point of view. The discrete wavelet transform is realized as a cascade of wavelet filter banks, providing a multiresolution decomposition which has proven very successful for compression [1], denoising [3], and estimation [10]. If the wavelet basis functions (derived from the iterated discrete filter) are not smooth, then perceptually unacceptable artifacts or false details will appear in the reconstructed or denoised signal [11]. In fact, the cascade converges fast enough that the smoothness of the continuous-time limit $\varphi(t)$ is visually comparable to that of a six-level iterate [11]. Thus we can study smoothness of the finite iterates (which appear in real-world applications) by studying the infinite limit, with all the mathematical tools available in the continuous-time setting.

Although the classical orthogonal Daubechies filters give rise to relatively smooth scaling and wavelet functions they do not achieve the maximal degree of smoothness. This is not surprising since Daubechies chose the maximal number of vanishing moments as a design criterion, not smoothness or regularity.[1]

As mentioned above, the problem of designing maximally smooth wavelet functions is, to our knowledge, still an open question. Solutions are known for the special cases of length 4 and length 6 filters [2] where simple parametrizations can be found and only a one-dimensional search has to be performed. Another approach yields improved smoothness by prescribing zeros of the filter frequency response at certain frequencies [5]. This method, however, does not yield maximal smoothness.

A new approach is presented that attempts to maximize the smoothness of an orthogonal multiresolution analysis. It utilizes the autocorrelation sequence domain where 1) the dimension of the parameter space can be considerably reduced and 2) an *unconstrained* optimization problem results (as opposed to a *constrained* problem if one were to work in the impulse response domain). This method is used to maximize the Hölder smoothness, a time-domain generalization of differentiability.

The following section introduces notation, followed by a review of a popular measure for smoothness, Hölder regularity. In section 3 the problem is formulated and section 4 outlines its solution. The examples presented in the subsequent section clearly demonstrate that much smoother scaling functions than those corresponding to Daubechies filters can be designed by the new method.

2. PRELIMINARIES

Let $h_{0,k}$ be the length $K+1$ impulse response of a dyadic scaling filter with z-transform

$$H_0(z) = \sum_{k=0}^{K} h_{0,k} z^{-k} \qquad (1)$$

[1] There is, however, a relationship between smoothness and vanishing moments. We will elaborate on this below.

and the normalization $H_0(1) = 2$. The associated scaling function is defined by

$$\varphi(t) = \sqrt{2}\sum_{k=0}^{K} h_{0,k}\varphi(2t-k). \tag{2}$$

In the following we will only consider orthogonal wavelet systems. Let a_k be the length $2K+1$ (symmetric) autocorrelation sequence of $h_{0,k}$, with (noncausal) z-transform

$$A_0(z) = \sum_{k=-K}^{K} a_k z^{-k}, \tag{3}$$

and power spectral density $A(\omega) = A_0(e^{j\omega})$. Precisely $\frac{K+1}{2}$ out of the total of $2K+1$ coefficients a_k can be considered as free parameters because of the symmetry

$$a_k = a_{-k}, \quad k = 1, \ldots, K, \tag{4}$$

and orthonormality, i.e.,

$$a_{2k} = \delta_{0k}, \quad k = 0, \ldots, \frac{K-1}{2}, \tag{5}$$

where δ_{0k} denotes the Kronecker delta.

An important theorem [2, p. 154ff] states that if the shifted and scaled versions of the wavelet function constitute an orthonormal basis of wavelets and if the wavelet function is N times continuously differentiable then the wavelet function possesses $N+1$ vanishing moments or, equivalently, the function $H_0(z)$ can be factorized [2] according to

$$H_0(z) = 2\left(\frac{1+z}{2}\right)^{N+1} Q(z).$$

The converse is, however, in general not true. Nevertheless, Daubechies used maximal vanishing moments as a criterion for the construction of the first family of orthogonal wavelets [2]. Daubechies' wavelet functions do exhibit smoothness (number of continuous derivatives) that asymptotically increases linearly as a function of the filter length, with slope $\approx 0.2075\frac{K+1}{2}$ [2, 14]. In this paper it will be shown how one can systematically sacrifice some zeros of $H_0(z)$ at -1 (vanishing moments) to achieve smoother scaling functions and consequently smoother wavelet functions than those of Daubechies in the Hölder sense.

2.1. Hölder Smoothness

Hölder smoothness is a direct generalization of the notion of differentiability from the natural to the real numbers. A function f has Hölder smoothness of order $\alpha \in [0, 1)$ if, for any t, h,

$$|f(t+h) - f(t)| < c|h|^{\alpha}, \tag{6}$$

where $c < \infty$ is a constant independent of t and h. For a higher order of Hölder smoothness $s_H(f) = L + \alpha$, $L \in \mathbf{N}$, $\alpha \in [0, 1)$, the definition above is applied to the L-th derivative of f. It can be shown that a function f possesses L continuous derivatives if and only if it has Hölder smoothness of order $s_H(f) \geq L$. Thus the notion of Hölder smoothness is more appropriate than the Sobolev notion [4], if one is interested in the time-domain behavior of a function.

Rioul gives an algorithm for determining the Hölder smoothness [11] that can be easily implemented and yields a very accurate estimate for s_H.

3. THE PROBLEM FORMULATION

The problem to be solved can be stated as follows:

> **Maximize** $s_H(\varphi)$ over the coefficients $\mathbf{a} = [a_K a_{K-2} \cdots a_1]^T$ of $A(\omega)$ according to equations (3)-(5), and subject to
>
> $$A(\omega) \geq 0. \tag{7}$$

In order to compute the scaling filter (and not its autocorrelation sequence), one must find the spectral factor [2] of the solution $\mathbf{a}$.

4. THE SOLUTION

The solution employs a general purpose optimization algorithm requiring only a function that computes the order of smoothness s_H for any valid set of parameters. The approach outlined here would work as well for any other method of measuring smoothness, such as Sobolev differentiability [4], [6], [7]. Gradient information would improve the performance of the optimization routine, but is not available for the Hölder smoothness. A continuous derivative of s_H with respect to the coefficients a_k does not exist.

It is obviously desirable to reduce the number of free parameters to reduce the complexity of the problem. This is addressed by the first of the following two sections. The second section deals with the choice of a valid parameter set, i.e., a set that is compatible with the constraint (7).

4.1. Moments and Smoothness

According to the theorem cited above, for a given degree of smoothness N, the function $A_0(z)$ must have $2(N+1)$ zeros at $z = -1$. Thus the dimension of the parameter space can be reduced to $\frac{K-1}{2} - N$. Although we do not know the achievable order of smoothness for a given filter length $K+1$, we have a lower bound by measuring the Hölder smoothness of the corresponding Daubechies filter. For example, the D6 filter has 3 free parameters but $s_H \approx 1.08$, so that there is just $3 - N - 1 = 1$ parameter left for optimization.

4.2. Positivity of $A(\omega)$

The remaining free parameters are further constrained by the nonnegativity condition (7). It effectively constrains the possible root locations of $A_0(z)$. Single roots on the unit circle are not admissible; complex roots off the unit circle have to occur in quadruples, real roots must occur in mirror pairs, and roots on the unit circle must come in complex conjugate pairs with multiplicity two, thus fixing two, one, and two parameters, respectively. It is important to note that each of the three possible cases can be expressed by equations that are linear in the autocorrelation coefficients.

However, it is not clear *a priori* which of the different combinations of roots (e.g., in the case of two free parameters either one complex quadruple or a pair of double roots on the unit circle can be chosen) leads to the maximum smoothness. Thus one has to try all possible constellations.

4.3. The Algorithm

1. Choose the length $K+1$ of the scaling filter.
2. Determine the number of degrees of freedom, i.e., find how many vanishing moments can be relaxed.
3. Determine all possible constellations that could give rise to a nonnegative power spectrum.
4. For each possible constellation, numerically maximize the desired smoothness over the parameter space, i.e., the $\frac{K-1}{2} - N$ radii and angles. A general purpose optimization routine is used. It only requires a function that computes the order of smoothness, given the parameter values.

5. RESULTS

We carried out the above method for filter lengths ranging from 4 to 30. When the filter length was 10 or larger, we obtained filters with greater Hölder smoothness than the Daubechies orthogonal wavelets. For lengths 4, 6, and 8, the optimization procedure did not yield higher smoothness than the Daubechies' filters. Figure 1 shows the the Hölder exponent s_H as a function of $K+1$ for the scaling functions corresponding to Daubechies filters and to the new filters. Beyond length 8 the new method clearly improves the smoothness, with increasing success for larger filter lengths. A linear least-squares fit to each of the smoothness plots found the Daubechies filters with order less than 30 to have a Hölder exponent growing as $.1418(K+1)$, while the maximally smooth filters have a Hölder exponent growth of $.2024(K+1)$. We know [14] that asymptotically the Daubechies filters' Hölder exponents grow as $.1037(K+1)$. The asymptotic behavior of the maximal possible Hölder exponent for a given filter length, as the length increases, remains an open question.

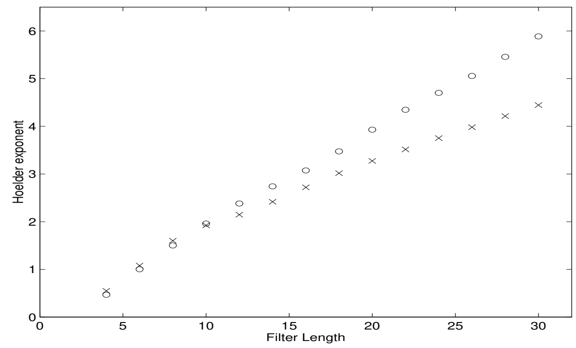

Figure 1: Hölder smoothness for Daubechies ($\times$) and optimized ($\circ$) filters as a function of filter length.

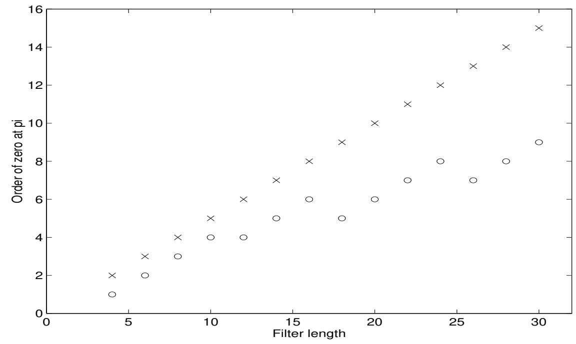

Figure 2: Order of zero at pi (number of vanishing moments) for Daubechies ($\times$) and maximally smooth ($\circ$) filters as a function of filter length.

In Figure 2 we have plotted the number of zeros at π for the Daubechies and maximally smooth scaling filters; this is equal to the number of vanishing moments of the corresponding wavelet system. Notice that while Daubechies filters of length $K+1$ have a zero of order $\frac{K+1}{2}$ at π, the maximally smooth filters have a zero of roughly half that order. In other words, we have given up half the vanishing moments of the Daubechies filter to obtain free parameters for optimizing the Hölder regularity. Notice that this behavior is not monotonic, probably because several different root constellations lead to near-optimal filters.

5.1. An Example

Consider the case of a length 26 ($K=25$) scaling filter that is optimized for s_H. The Daubechies filter corresponds to smoothness $s_H = 4.005$. After optimization of three distinct complex quadruples (off the unit circle) one gets a scaling function with smoothness $s_H = 5.06$, one more continuous derivative. Figure 3 depicts the scaling functions and their 4-th derivatives for each of these filters. While the scaling functions have a similar appearance, the higher degree of

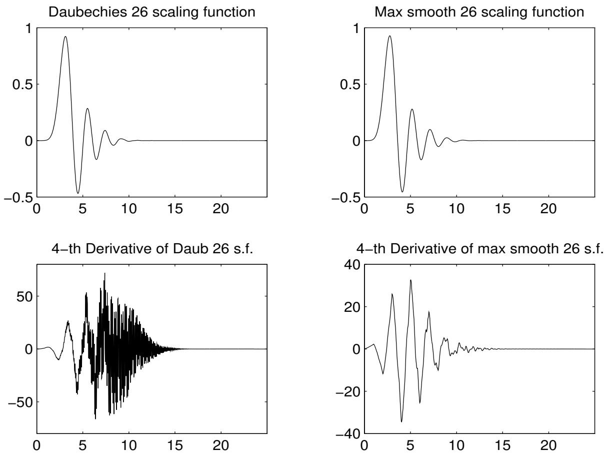

Figure 3: Scaling functions (top) and their 4-th derivatives (bottom) corresponding to 26-coefficient Daubechies and optimized filters.

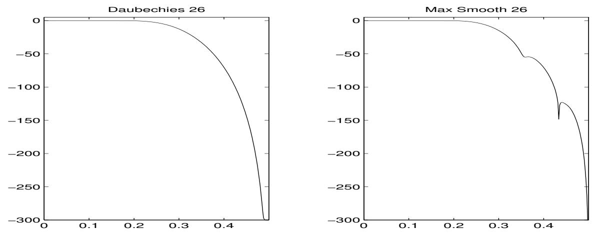

Figure 4: Frequency responses in dB of the 26-coefficient Daubechies (left) and max-smooth (right) filters.

smoothness of the derivative is quite apparent. The filter frequency responses are shown in Figure 4.

6. CONCLUSIONS

We have described the class of maximally smooth wavelets, filters characterized by the property that their associated scaling functions are as smooth as possible for the given filter length. Smoothness is measured here using the notion of Hölder regularity. One can relax some but not all of the vanishing wavelet moment conditions and use the resulting free parameters for filter design. An optimization algorithm was described that works with the autocorrelation of the scaling (lowpass) filter to range over the parameter space and find these maximally smooth filters. The resulting filters are generally smoother than Daubechies orthogonal wavelets, and have smoothness which grows faster than the Daubechies wavelets as the filter length increases. Future work [7] will apply the methods developed here to other measures of smoothness, such as differentiability in the Sobolev sense [4], [6].

7. REFERENCES

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, Image coding using the wavelet transform. *IEEE Trans. on Image Processing*, 2(2), April 1992.

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

[3] D. L. Donoho, De-noising by soft-thresholding. *IEEE Trans. Info. Theory*, 41, pp. 613–627, May 1995.

[4] T. Eirola, Sobolev characterization of solutions of dilation equations. *SIAM J. of Math.*, 23, pp. 1015–1030, July 1992.

[5] P. N. Heller, Lagrange M-th Band Filters and the Construction of Smooth M-band Wavelets. *Proc. IEEE-SP Intl. Symp. on Time-Frequency and Time-Scale Analysis,* Philadelphia, PA, 1994.

[6] P. N. Heller, J. M. Shapiro, and R. O. Wells, Jr., Optimally smooth symmetric quadrature mirror filters for image coding, in *Proc. SPIE 2491, Wavelet applications for dual use,* Orlando, FL, April, 1995.

[7] M. Lang and P. N. Heller, Maximally smooth wavelet filter banks and their design, article in preparation, 1995.

[8] M. Lang, I. W. Selesnick, J. E. Odegard, and C. S. Burrus, Constrained FIR filter design for 2-band filter banks and orthonormal wavelets. In *Sixth Digital Signal Processing Workshop*, pages 211–214, October 1994.

[9] F. Mintzer, Filters for distortion-free two-band multirate filter banks. *IEEE Trans. ASSP*, pages 626–630, June 1985.

[10] P. Moulin, Wavelet thresholding techniques for power spectrum estimation. *IEEE Trans. SP*, 42(11):3126–3136, November 1994.

[11] O. Rioul, Simple regularity criteria for subdivision schemes. *SIAM J. Math Anal.*, 23(6):1544–1576, November 1992.

[12] O. Rioul and P. Duhamel, A Remez exchange algorithm for orthonormal wavelets. *IEEE CAS II,* pp. 550-560, August 1994.

[13] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* Prentice Hall, Englewood Cliffs, NJ, 1992.

[14] H. Volkmer, On the regularity of wavelets, *IEEE Trans. on Information Theory,* 38 (1992), pp. 872-876.