Approximation Power of Biorthogonal Wavelet Expansions

Michael Unser, *Senior Member, IEEE*

***Abstract*— This paper looks at the effect of the number of vanishing moments on the approximation power of wavelet expansions. The Strang–Fix conditions imply that the error for an orthogonal wavelet approximation at scale $a = 2^{-i}$ globally decays as a^N, where N is the order of the transform. This is why, for a given number of scales, higher order wavelet transforms usually result in better signal approximations. We prove that this result carries over for the general biorthogonal case and that the rate of decay of the error is determined by the order properties of the synthesis scaling function alone. We also derive asymptotic error formulas and show that biorthogonal wavelet transforms are equivalent to their corresponding orthogonal projector as the scale goes to zero. These results strengthen Sweldens' earlier analysis and confirm that the approximation power of biorthogonal and (semi-)orthogonal wavelet expansions is essentially the same. Finally, we compare the asymptotic performance of various wavelet transforms and briefly discuss the advantages of splines. We also indicate how the smoothness of the basis functions is beneficial in reducing the approximation error.**

I. Introduction

For researchers working with multirate filterbanks, the mathematical theory of the wavelet transform brought about the new constraint of designing filterbanks with a certain number of zeros (multiplicity N) at $z = -1$ [1], [2]. Indeed, most families of wavelet bases are indexed by this important order parameter, which also represents the number of vanishing moments for the analysis wavelet [3], [4]. One of the initial justifications for selecting a zero of multiplicity N is that this condition is necessary for constructing regular wavelets with $N-1$ continuous derivatives (cf. [5, Corollary 5.5.4, p. 155]). Unfortunately, it is not sufficient, and the regularity index of most wavelet bases is generally smaller than $N-1$ [5]–[7]. The number of vanishing moments of the analysis wavelet also play a crucial role in the characterization of the local Hölder exponent of singularities [8], [9]. These are all reasons why the vanishing moments of the wavelet are generally believed to be useful in many applications. Two examples where these properties are especially relevant are image coding [10], [11] and the analysis of fractal random processes [12], [13].

Manuscript received May 25, 1995; revised September 15, 1995. The associate editor coordinating the review of this paper and approving it for publication was Prof. Banu Onaral.

The author is with the Biomedical Engineering and Instrumentation Program, National Center for Research Resources, National Institutes of Health, Bethesda, MD 20892-5766 USA (e-mail: unser@helix.nih.gov).

Publisher Item Identifier S 1053-587X(96)02421-8.

Beside the regularity of the basis functions themselves, there is also another compelling reason for using higher order wavelet decompositions, which takes its roots in approximation theory [14], [15]. This particular aspect of the wavelet theory is less well known in signal processing, but it is probably more directly relevant to this particular area of application. Specifically, if f is a smooth L_2 function (in the sense that its Nth derivative $f^{(N)}$ is square integrable) and if ϕ is an Nth-order orthogonal scaling function, then one has the following error bound (cf. [6]):

$$\|f - P_{2^{-i}} f\|_{L_2} \le C_\phi \cdot 2^{-iN} \cdot \|f^{(N)}\|_{L_2} \quad (1)$$

where $P_{2^{-i}} f$ denotes the approximation of f at scale $a = 2^{-i}$ (orthogonal projection):

$$P_{2^{-i}} f = \sum_{k\in Z} \langle f, \phi_{-i,k}\rangle \phi_{-i,k} = \sum_{j=-i+1}^{+\infty} \sum_{k\in Z} \langle f, \psi_{j,k}\rangle \psi_{j,k} \quad (2)$$

using the short-form notation $\varphi_{j,k} = 2^{-j/2}\varphi(2^{-j}x - k)$. In other words, the order N controls the rate of decay of the approximation error, suggesting that higher order wavelet approximations should generally require less terms (or scales) to approximate a smooth function within a certain error tolerance. This characteristic $O(a^N)$ decay of the error as a function of the scale is also apparent in the pointwise and asymptotic estimates that have been derived recently [16]–[19]. Sweldens and Piessens (SP) obtained the same estimates for biorthogonal wavelet expansions as well and proved that the asymptotic error only depends on the order properties of the primary representation space and not on how the complementary wavelet subspaces are chosen [19]. These results strongly suggest that the same kind of extension should also be possible for Strang's initial L_2-bound (1).

In this paper, we start by briefly explaining how these results relate to the Strang–Fix theory developed in the early 1970's, long before the invention of the wavelet transform. In Section III, we then show how to obtain the corresponding L_2-error bound for the more general biorthogonal case. Biorthogonal transforms may be preferable in some applications because of the freedom that they leave in the specification of the analysis filters. In Section IV, we extend the SP asymptotic analysis and determine the limit of the error as a goes to zero. The main improvements over SP's earlier results are as follows.

1) The computation is more direct in the sense that it avoids using wavelet expansions.

2) The present error estimate is sharper (smaller constant) and asymptotically exact.
3) The setting is more general.

In the process, we also provide practical formulas that permit an easy calculation of the bound constants. We then compare the performance of various wavelet transforms. Finally, we conclude by proposing a new explanation of why the smoothness of the basis functions has a reducing effect on the approximation error.

II. Preliminary Notions

We take a more general perspective than the usual multiresolution formulation and consider signal representations in terms of rescaled translates of an (almost) arbitrary generating function φ. At this stage, φ is not required to have the multiresolution property, and the scale parameter (or sampling step) a can be arbitrary (not necessarily a power of two).

Definition 2.1: An Nth-order generating function is a function $\varphi \in L_2$ such that

i) $0 < A \leq \sum_{k\in Z} |\hat{\varphi}(\omega + 2\pi k)|^2 \leq B < +\infty$ (Riesz basis condition)

ii) $\hat{\varphi}(0) = 1$ and $\hat{\varphi}^{(m)}(2\pi k) = 0, \quad k \in Z, k \neq 0$ for $(m = 0, \cdots, N-1)$ (Order property)

where $\hat{\varphi}(\omega)$ is the Fourier transform of φ, and $\hat{\varphi}^{(m)}(\omega)$ denotes its mth derivative with respect to ω.

The corresponding signal representation space at scale a is

$$V_a(\varphi) = \left\{ f_a(x) = a^{-1/2} \sum_{k\in Z} c_a(k)\varphi(x/a - k) \colon c_a \in l_2 \right\}. \quad (3)$$

Condition i) ensures that $V_a(\varphi)$ is a well-defined subspace of L_2 and that each function $f_a \in V_a(\varphi)$ has a unique representation in terms of its coefficients $c_a(k)$ [20]. Condition ii) is an indirect statement of the fact φ that reproduces all polynomials of degree $N-1$ [6], [14].

An other equivalent way of expressing the order property ii) is to consider the periodization of the function $x^m\varphi(x)$ whose Fourier transform is $j^m\hat{\varphi}^{(m)}(\omega)$. Using the property that a periodization along the time dimension corresponds to a sampling in the frequency domain, we obtain the following Fourier series: representations

$$\sum_{k\in Z} (x-k)^m \varphi(x-k) = j^m \sum_{k\in Z} \hat{\varphi}^{(m)}(2\pi k)e^{j2\pi kx} = m_\varphi^m, \quad (m = 0, \cdots, N-1) \quad (4)$$

where the right-most constant corresponds to the Fourier component at the origin since all others are zero as a result of Condition ii). This constant turns out to be the mth-order moment of the generating function, which is defined as follows:

$$m_\varphi^m := \int_{-\infty}^{+\infty} x^m \varphi(x)dx = j^m \hat{\varphi}^{(m)}(0). \quad (5)$$

In particular, for $m = 0$, we have $m_\varphi^0 = \hat{\varphi}(0) = 1$, which gets translated into the following "partition of unity" property

$$\sum_{k\in Z} \varphi(x-k) = 1. \quad (6)$$

Given an arbitrary function $f \in L_2$, we compute its minimum error approximation in $V_a(\varphi)$ (orthogonal projection at scale a) as

$$(P_a f)(x) = a^{-1} \sum_{k\in Z} \langle f, \mathring{\varphi}(x/a - k)\rangle \cdot \varphi(x/a - k). \quad (7)$$

where $\mathring{\varphi} \in V_1(\varphi)$ is the unique semi-orthogonal dual of φ [20], and where a^{-1} is the appropriate inner-product normalization factor. The order property ii) has some important implication on the approximation power of the representation.

Theorem 2.2 (Strang-Fix): If φ is an Nth-order generating function with suitable decay, then the minimum approximation error at step size a for an arbitrary function f (sufficiently smooth) is bounded as follows:

$$\inf_{f_a \in V_a(\varphi)} \|f - f_a\| \leq C_\varphi \cdot a^N \cdot \|f^{(N)}\| \quad (8)$$

where C_φ is a constant that is independent of f.

This result is an abbreviated form of the Strang–Fix conditions, which are even stronger in the sense that the implication goes both ways [14]. Strang and Fix initially assumed that φ is compactly supported, but their result has also been extended for noncompact φ with polynomial decay [21], [22].

What is somewhat counterintuitive with Theorem 2.2 is that the rate of decay of the error does not depend on the smoothness (or regularity) of φ. The only determining factor is the multiplicity of the zeros of $\hat{\varphi}$ at $\omega = 2k\pi$, $k \in Z \setminus \{0\}$ (cf. Condition ii)). One point of this paper will be to demonstrate that smoothness has some importance as well because of the way in which it affects the magnitude of the constant involved.

In the particular case of the wavelet transform, the generator φ has the additional multiresolution property; the scale is also restricted to powers of two.

Definition 2.3: An Nth-order scaling function is an Nth-order generating function that satisfies the additional two-scale relation

$$\varphi(x/2) = \sum_{k\in Z} h(k)\varphi(x-k). \quad (9)$$

By applying this refinement equation *ad infinitum*, we get the equivalent infinite product representation of the Fourier transform of the scaling function φ

$$\hat{\varphi}(\omega) = \prod_{i=1}^{+\infty} H(e^{j\omega/2^i}) \quad (10)$$

where $H(z)$ denotes the z-transform of the refinement filter h. It is then possible to establish the following proposition.

Proposition 2.4 If $H(e^{j\omega})$ has zeros of multiplicity N at $\omega = \pi$ (or if $H(z) = 2^{-N}(1+z)^N \cdot Q(z)$, where $Q(z)$ is a stable transfer function), then φ is an Nth-order scaling function.

Since the infimum in (8) corresponds to the orthogonal projection, we can therefore directly apply this result to get Strang's wavelet bound (1), which is obviously also valid for semi-orthogonal representations.

III. Error Bounds for Biorthogonal Expansions

Here, we are concerned with biorthogonal wavelet expansions that are obtained by oblique projection (instead of orthogonal) [23], [24]. Using essentially the same formulation as above, we define the general rescaled oblique projection operator as in [25]

$$(\tilde{P}_a f)(x) = a^{-1} \sum_{k\in Z} \langle f, \tilde{\varphi}(x/a - k)\rangle \cdot \varphi(x/a - k) \quad (11)$$

where φ and $\tilde{\varphi}$ are two biorthogonal generating functions such that

$$\langle \varphi(x-k), \tilde{\varphi}(x-l)\rangle = \delta_{k,l}. \quad (12)$$

Specifically, $\tilde{P}_a f$ is the projection of f onto $V_a(\varphi)$ perpendicular to $V_a(\tilde{\varphi})$. The essential difference with (7) is that $\tilde{\varphi}$ (unlike $\mathring{\varphi}$) is not necessarily in $V_1(\varphi)$. We can now state our first result.

Theorem 3.1: If φ is an Nth-order generating function with appropriate decay, then the oblique projection error at scale a for an arbitrary function f (sufficiently smooth) is bounded as follows:

$$\inf_{f_a\in V_a(\varphi)} \|f - f_a\| \le \|f - \tilde{P}_a f\| \le C_{\varphi,\tilde{\varphi}} \cdot a^N \cdot \|f^{(N)}\| \quad (13)$$

where $C_{\varphi,\tilde{\varphi}}$ is a constant that is independent of f.

Proof: We have chosen to derive this result as a corollary of the Strang–Fix conditions (8). For this purpose, we make use of a rescaled version of Theorem 3 in [25], which provides a direct bound between the orthogonal and oblique projection errors

$$\|f - P_a f\| \le \|f - \tilde{P}_a f\| \le \frac{1}{\cos\theta}\|f - P_a f\| \quad (14)$$

where $\cos\theta$ is given by

$$\cos\theta = \operatorname*{ess\,inf}_{\omega\in[0,\pi)} \frac{\sum_{k\in Z} \overline{\hat{\varphi}(\omega+2\pi k)} \cdot \hat{\tilde{\varphi}}(\omega+2\pi k)}{\sqrt{\sum_{k\in Z}|\hat{\varphi}(\omega+2\pi k)|^2 \cdot \sum_{k\in Z}|\hat{\tilde{\varphi}}(\omega+2\pi k)|^2}}. \quad (15)$$

Since the generating functions are biorthogonal, we can easily show that $\cos\theta \ge (B\tilde{B})^{-1} > 0$, where B and $\tilde{B}$ are the upper frame bounds for φ and $\tilde{\varphi}$ in Definition 2.1, respectively. Hence, combining (8) with (14), we get

$$\|f - P_a f\| \le \|f - \tilde{P}_a f\| \le (\cos\theta)^{-1} \cdot \|f - P_a f\| \le C_{\varphi,\tilde{\varphi}} \cdot a^N \cdot \|f^{(N)}\|$$

with a finite constant $C_{\varphi,\tilde{\varphi}} = C_\varphi / \cos\theta$. □

A direct implication of this result is that we can approximate any L_2 function as closely as we wish, provided that we use a sampling step that is sufficiently small. The theorem also indicates that the rate of convergence depends on the order properties of the synthesis function φ alone and that the analysis function has essentially no influence (except on the magnitude of the constant). We can also particularize this result to the case where φ and $\tilde{\varphi}$ are both scaling functions and obtain the generalization of (1) for biorthogonal wavelet expansions

$$\left\| f - \sum_{j=-i+1}^{+\infty} \sum_{k\in Z} \langle f, \tilde{\psi}_{j,k}\rangle \psi_{j,k} \right\| = \left\| f - \sum_{k\in Z} \langle f, \tilde{\varphi}_{-i,k}\rangle \varphi_{-i,k} \right\| \le C_{\varphi,\tilde{\varphi}} \cdot 2^{-iN} \cdot \|f^{(N)}\|. \quad (16)$$

In this sense, orthogonal, semi-orthogonal, and biorthogonal wavelet expansions are all essentially equivalent. Note that in the semi-orthognal case, $\cos\theta = 1$ so that we end up with Strang's initial estimate. The advantage of biorthogonal expansions is that they offer more freedom in the design of the wavelet filters. For instance, these can all be symmetric FIR, which is typically not possible in the (semi-)orthogonal case [2], [5].

IV. Asymptotic Performance Analysis

The general error bound shows that up to a constant factor $(1/\cos\theta)$, oblique and orthogonal projection operators are qualitatively equivalent. However, one should note that this estimate corresponds to a worse-case scenario and that the loss of performance in a practical situation may be much less than what is suggested by this bound. In this section, we will be more quantitative and provide a precise characterization of the asymptotic error. We will also use these formulas to compare the performance of various wavelet transforms.

A. Asymptotic Error Characterization and Equivalence

For smaller values of the scale, we can use a Taylor series analysis method to get the limiting form of the approximation error, which also exhibits the characteristic $O(a^N)$ behavior. The conclusion of this analysis is that in most practical cases, biorthogonal projection operators are asymptotically optimal in the sense that they have exactly the same limit behavior as their corresponding orthogonal projector. This result is expressed by the following theorem.

Theorem 4.1: The Nth-order oblique projection operator (11) is asymptotically equivalent to the corresponding orthogonal projector P_φ in the sense that both

$$\|f - \tilde{P}_a f\| = C_\varphi^- \cdot a^N \cdot \|f^{(N)}\| + O(a^{N+1})$$

and

$$\|f - P_a f\| = C_\varphi^- \cdot a^N \cdot \|f^{(N)}\| + O(a^{N+1}) \quad \text{as } a \to 0$$

provided that either of the two following conditions are satisfied: a) $\sum_{k\in Z} \tilde{\varphi}(x-k) = 1$ (i.e., $\tilde{\varphi}$ is a first-order generating function), or b) N is odd plus φ and $\tilde{\varphi}$ are both symmetrical. The optimal constant C_φ^-, which does not depend on $\tilde{\varphi}$, is given by

$$C_\varphi^- = \frac{1}{N!}\left(\sum_{k\neq 0} |\hat{\varphi}^{(N)}(2\pi k)|^2\right)^{1/2}. \tag{17}$$

Note that condition a) is the minimum requirement for constructing meaningful scaling functions and (admissible) wavelets. Our result is therefore consistent with SP's earlier report that the asymptotic error for biorthogonal wavelet expansions is entirely determined by the properties of the approximation space (cf. [19, Theorem 2, p. 398]). We note, however, that our specification of the optimal constant C_φ^- leads to a sharper L_2-estimate, which is asymptotically exact. It is also not difficult to show that the value of this constant remains the same if we use any another equivalent generating function $\varphi_{\rm eq} \in V_1(\varphi)$ such that $\hat{\varphi}_{\rm eq}(0) = 1$.

Condition b) also indicates that there are asymptotically optimal solutions that do not require the partition of unity condition on the analysis side. For instance, one may consider a single resolution deconvolution system that uses a symmetric Gaussian on the analysis side (acquisition device) and a centered B-spline of even degree n on the synthesis side. The corresponding oblique projector can be implemented exactly by using an additional digital filtering module, as described in [25].

The detailed proof of Theorem 4.1, which also provides some insight into the approximation process, is presented in the next section. Those not interested in derivations can directly proceed to Section IV-C.

B. Proof of Theorem 4.1 and Related Results

This section presents an alternative to the asymptotic analysis in [19]. The main improvement is the use of a more direct formulation in which the approximation error is expressed using the reproducing kernel rather than an infinite sequence of wavelet terms. This results in a number of simplifications and lends itself more easily to the derivation of exact bound constants.

We start with a result that is essentially equivalent to lemma 5 in [19].

Lemma 4.2: Let $\eta(y) = \langle \varphi(x-y), \tilde{\varphi}(x)\rangle$ be the cross-correlation function between the Nth-order generating function φ and its dual $\tilde{\varphi}$. Then, we have the following moment properties:

$$m_\eta^m = \sum_{k=0}^{m} \binom{m}{k} (-1)^k m_\varphi^k m_{\tilde{\varphi}}^{m-k} = \begin{cases} 1, & m = 0 \\ 0, & m = 1, \cdots, N+\tilde{N}-1 \end{cases} \tag{18}$$

where $\tilde{N}$ is the order of $\tilde{\varphi}$.

Proof: To derive the central identity, we write the Fourier transform of η and differentiate m times applying the product rule:

$$\hat{\eta}^{(m)}(\omega) = \sum_{l=0}^{m} \binom{m}{l} \overline{\hat{\varphi}^{(l)}(\omega)}\, \hat{\tilde{\varphi}}^{(m-l)}(\omega) = \sum_{l=0}^{m} \binom{m}{l} (-1)^l \hat{\varphi}^{(l)}(-\omega) \hat{\tilde{\varphi}}^{(m-l)}(\omega).$$

Note that the right-hand side expression is obtained by using the Hermitian symmetry of the Fourier transform of the real valued function φ. Rewriting this equation for $\omega = 0$ and using the moment relation (5), we obtain the following identity:

$$m_\eta^m = \sum_{k=0}^{m} \binom{m}{k} (-1)^k m_\varphi^k m_{\tilde{\varphi}}^{m-k} \tag{19}$$

which is true irrespective of the value of m. Next, we express the biorthogonality condition in the Fourier domain

$$\langle \varphi(x-k), \hat{\varphi}(x-l)\rangle = \delta_{k,l} \Leftrightarrow \sum_{k\in Z} \overline{\hat{\tilde{\varphi}}^*(\omega+2\pi k)} \hat{\varphi}(\omega + 2\pi k) = 1. \tag{20}$$

and take the mth derivative at $\omega = 0$, which yields

$$\sum_{k\in Z} \sum_{l=0}^{m} \binom{m}{l} (-1)^l \hat{\varphi}^{(l)}(-2\pi k)\, \hat{\tilde{\varphi}}^{(m-l)}(2\pi k) = 0.$$

If $m < \tilde{N} + N$, where $\tilde{N}$ is the order of the dual function, then, for $k \neq 0$, there will always be at least one term in the product that is zero, i.e., $\hat{\varphi}^{(m)}(2\pi k) = 0$ for $m = 0, \cdots, N-1$. Hence, the only remaining terms in the sum are those at the origin, which implies that

$$\sum_{l=0}^{m} \binom{m}{l} (-1)^l \hat{\varphi}^{(l)}(0)\, \hat{\tilde{\varphi}}^{m-l}(0) = \begin{cases} 1, & m = 0 \\ 0, & m = 1, \cdots, N+\tilde{N}-1. \end{cases} \tag{21}$$

The last step is to rewrite this identity in terms of moments which yields the desired result. □

For notation convenience, we rewrite the oblique projection operator (11) as

$$(\tilde{P}_a f)(x) = a^{-1} \int_{-\infty}^{+\infty} f(y) K\left(\frac{x}{a}, \frac{y}{a}\right) dy \tag{22}$$

where $K(x,y)$ is the reproducing kernel

$$K(x,y) = \sum_{k\in Z} \tilde{\varphi}(y-k)\varphi(x-k). \tag{23}$$

Using the results in Lemma 4.2, we can now show that $K(x,y)$ has the following cancellation properties.

Proposition 4.3: If φ has an Nth order of approximation, then

$$e_0(x) = \int_{-\infty}^{+\infty} K(x,y)dy = 1 \tag{24}$$

$$e_m(x) = \int_{-\infty}^{+\infty} (y-x)^m K(x,y)dy = 0 \quad m = 1,\cdots,N-1 \tag{25}$$

and

$$e_N(x) = m_\eta^N - (-1)^N m_\varphi^N + (-1)^N \sum_{k\in Z}(x-k)^N \varphi(x-k) \tag{26}$$

where m_η^N is the Nth moment of the cross-correlation function η defined in Lemma 4.2.

Proof: Using the Binomial theorem, we expand $e_m(x)$ as follows:

$$\begin{aligned} e_m(x) &= \int_{-\infty}^{+\infty} \sum_{k\in Z}[(y-k)-(x-k)]^m \\ &\quad \times \varphi(x-k)\tilde{\varphi}(y-k)\cdot dy \\ &= \sum_{k\in Z}\sum_{l=0}^{m} \binom{m}{l}(-1)^l (x-k)^l \varphi(x-k) \\ &\quad \times \int_{-\infty}^{+\infty} (y-k)^{m-l}\tilde{\varphi}(y-k)\cdot dy \\ &= \sum_{l=0}^{m}\binom{m}{l}(-1)^l \sum_{k\in Z}(x-k)^l\varphi(x-k) \\ &\quad \times \int_{-\infty}^{+\infty} y^{m-l}\tilde{\varphi}(y)\cdot dy. \end{aligned}$$

Note that we need some decay on φ (e.g., $\varphi(x) \le C\cdot|1+x|^{-(N+1+\epsilon)}$ with $0<\epsilon$ to justify the various permutations of the infinite sum. For $m<N$, we rewrite this equation in terms of moments using (4) and apply Lemma 4.2 to show that

$$\begin{aligned} e_m(x) &= \sum_{l=0}^{m}\binom{m}{l}(-1)^l m_\varphi^l m_{\tilde{\varphi}}^{m-l} \\ &= \begin{cases} 1, & m=0 \\ 0, & m=1,\cdots,N-1. \end{cases} \end{aligned}$$

For $m=N$, we can also identify moments and isolate the last nonconstant term on the right

$$\begin{aligned} e_N(x) &= \sum_{l=0}^{N-1}\binom{N}{l}(-1)^l m_\varphi^l m_{\tilde{\varphi}}^{N-l} + \\ &\quad (-1)^N m_{\tilde{\varphi}}^0 \sum_{k\in Z}(x-k)^N\varphi(x-k). \end{aligned}$$

Finally, using (19) with $m=N$ and the fact that $m_\varphi^0 = m_{\tilde{\varphi}}^0 = 1$, we replace the first term by its equivalent expression $m_\eta^N - (-1)^N m_\varphi^N$, which yields the desired result. □

Interestingly, the functions $e_m(x)$ also provide a direct measure of the approximation error for the monomial x^m. Specifically, it can be shown that

$$\begin{aligned} (-1)^m e_m(x) &= x^m - \int_{-\infty}^{+\infty} y^m K(x,y)dy \\ &= x^m - \tilde{P}_1 x^m, \quad \text{for } m=0,\cdots,N. \end{aligned}$$

Thus, the cancellation properties in Proposition 4.3 imply a perfect reproduction of polynomials of degree $n=N-1$. The error for the lowest degree monomial that cannot be approximated exactly is $(-1)^N e_N(x)$. This observation provides the connection with the approach of SP, who express this error in term of the "monowavelet" function $\tau_0(x)$, which comes from the periodization of a monomial multiplied with a wavelet [18], [19].

We then use Proposition 4.3 to establish the following standard pointwise estimate.

Proposition 4.4: If φ has an Nth-order of approximation and $f\in C^{N+1}$, then

$$f(x) - (\tilde{P}_a f)(x) = -\frac{a^N}{N!} f^{(n)}(x) e_N\left(\frac{x}{a}\right) + O(a^{N+1}). \tag{27}$$

Proof: Using (22) and (24), we write the approximation error as

$$f(x) - (\tilde{P}_a f)(x) = \int_{-\infty}^{+\infty} (f(x)-f(y))\frac{1}{a}K\left(\frac{x}{a},\frac{y}{a}\right)dy.$$

Next, we replace $f(y)$ by its Nth-order Taylor series around x

$$\begin{aligned} f(x) &- (\tilde{P}_a f)(x) \\ &= -\sum_{m=1}^{N}\frac{f^{(m)}(x)}{m!}\int_{-\infty}^{+\infty}(y-x)^m \frac{1}{a}K\left(\frac{x}{a},\frac{y}{a}\right)dy \\ &\quad - \int_{-\infty}^{+\infty} R_{N+1}(y)\frac{1}{a}K\left(\frac{x}{a},\frac{y}{a}\right)dy \\ &= -\sum_{m=1}^{N}\frac{a^m\cdot f^{(m)}(x)}{m!}e_m\left(\frac{x}{a}\right) \\ &\quad - \int_{-\infty}^{+\infty} R_{N+1}(y)\frac{1}{a}K\left(\frac{x}{a},\frac{y}{a}\right)dy \end{aligned}$$

where the rest of the Taylor series is $R_{N+1}(y) = (y-x)^{N+1} f^{N+1}(\xi)/(N+1)!$ with $x\le\xi\le y$. As a consequence of the order property (cf. Proposition 4.3), the first $N-1$ error terms are zero (polynomial cancellation). This leads to the following estimate:

$$\left| f(x) - (\tilde{P}_a f)(x) + \frac{a^N\cdot f^{(N)}(x)}{N!}e_N\left(\frac{x}{a}\right)\right| \le C\cdot a^{N+1}.$$

where

$$C = \frac{\|f^{(N+1)}\|_\infty}{(N+1)!}\sup_x \int_{-\infty}^{+\infty}|y-x|^{N+1}|K(x,y)|dy$$

which proves the desired result. Note that the right-hand side supremum is well defined if φ and $\tilde{\varphi}$ both decay faster than $O(x^{(N+1+\epsilon)})$. □

We are now in the position to prove our main result.

Proof of Theorem 4.1: For a sufficiently small, the $O(a^{N+1})$ error terms become negligible, and we can use the pointwise estimate in Proposition 4.4 to obtain the asymptotic form of the L_2-error

$$\lim_{a\to 0}\left(\frac{\|f(x)-(\tilde{P}_a f)(x)\|}{a^N}\right)=\frac{1}{N!}\lim_{a\to 0}\left\|f^{(N)}(x)\cdot e_N\left(\frac{x}{a}\right)\right\|. \tag{28}$$

To evaluate this limit, we start by computing the Fourier series expansion of $e_N(x)$ given by (26), noting that the second nonconstant term corresponds to the periodized version of $g(x)=(-x)^N\varphi(x)$ whose continuous Fourier transform is $\hat{g}(\omega)=(-j)^N\hat{\varphi}^{(N)}(\omega)$. This yields

$$e_N(x)=m_\eta^N+(-j)^N\sum_{\substack{k\in Z\\ k\neq 0}}\hat{\varphi}^{(N)}(2\pi k)e^{j2\pi kx}. \tag{29}$$

Next, we replace the function $f^{(N)}\in L_2$ in (28) by its best (least square) piecewise constant approximation with step size a, which we denote by $f_a^{(N)}$. This is a valid substitution in the sense that

$$\lim_{a\to 0}\left\|f^{(N)}(x)e_N\left(\frac{x}{a}\right)\right\|_{L_2}=\lim_{a\to 0}\left\|f_a^{(N)}(x)e_N\left(\frac{x}{a}\right)\right\|_{L_2}. \tag{30}$$

This equivalence directly follows from the inequality

$$\lim_{a\to 0}\left\|f^{(N)}(x)e_N\left(\frac{x}{a}\right)-f_a^{(N)}e_N\left(\frac{x}{a}\right)\right\|_{L_2} \le \|e_N\|_\infty \underbrace{\lim_{a\to 0}\|f^{(N)}-f_a^{(N)}\|_{L_2}}_{=0}=0$$

where we use the fact that $e_N(x)\in L_\infty$ and the property that piecewise constant splines are dense in L_2 (cf. Theorem 2.2 with $N=1$ as $a\to 0$). Let us now concentrate on the computation of the quantity

$$\left\|f_a^{(N)}(x)e_N\left(\frac{x}{a}\right)\right\|_{L_2}^2 =\sum_{k\in Z}\int_{k\cdot a}^{(k+1)\cdot a}|f_a^{(N)}(ka)|^2\left|e_N\left(\frac{x}{a}\right)\right|^2dx.$$

Using the fact that the function $e_N(x/a)$ is periodic with periodicity a, we find that

$$\begin{aligned}\left\|f_a^{(N)}(x)e_N\left(\frac{x}{a}\right)\right\|_{L_2}^2 &=\left(\frac{1}{a}\int_0^a\left|e_N\left(\frac{x}{a}\right)\right|^2dx\right)\cdot\left(\sum_{k\in Z}a|f_a^{(N)}(ka)|^2\right)\\ &=E_N^2\cdot\|f_a^{(N)}\|_{L_2}^2\end{aligned} \tag{31}$$

where E_N^2 is the mean square value of e_N given by (cf. (29))

$$\begin{aligned}E_N^2&=\frac{1}{a}\int_0^a|e_N(x/a)|^2dx=\int_0^1|e_N(x)|^2dx\\ &=\left(m_\eta^N\right)^2+\sum_{k\neq 0}|\hat{\varphi}^{(N)}(2\pi k)|^2.\end{aligned} \tag{32}$$

Combining (30) and (31), we then rewrite (28) as

$$\begin{aligned}\lim_{a\to 0}\left(\frac{\|f(x)-(\tilde{P}_a f)(x)\|}{a^N}\right)&=\left(\frac{E_N}{N!}\right)\cdot\lim_{a\to 0}\left\|f_a^{(N)}(x)\right\|_{L_2}\\ &=C_{\varphi,\tilde{\varphi}}^-\cdot\left\|f^{(N)}\right\|\end{aligned}$$

where we use the fact that $f_a^{(N)}\to f^{(N)}$ as a goes to zero. Thus, the constant is given by

$$C_{\varphi,\tilde{\varphi}}^-=\frac{1}{N!}\left(\left(m_\eta^N\right)^2+\sum_{k\neq 0}|\hat{\varphi}^{(N)}(2\pi k)|^2\right)^{1/2}.$$

The only term in this formula that depends on $\tilde{\varphi}$ is m_η^N, and the constant is obviously minimum for $m_\eta^N=0$. This moment vanishes when $\tilde{N}\geq 1$ (cf. Lemma 4.2) or when the second condition b) in Theorem 4.1 is met. In particular, the lower limit is achieved with the least square solution for which $\tilde{\varphi}=\mathring{\varphi}\in V_1(\varphi)$ and $\tilde{N}=N$. □

C. Computation of the Bound Constant

Theorem 4.1 provides us with an explicit characterization for the minimum bound constant C_φ^-, which can be the basis for a quantitative evaluation. Although there are cases such as splines where the scaling function $\hat{\varphi}(\omega)$ is known explicitly and (17) is applicable directly, we also need a mechanism to compute the required derivatives when the multiresolution representation is specified indirectly through the refinement filter h.

To derive such a relation, we rewrite the two scale (9) in the Fourier domain and differentiate N times, applying the chain rule

$$\hat{\varphi}^{(N)}(\omega)=\sum_{m=0}^{N}\binom{N}{m}\frac{1}{2^m}H^{(m)}(e^{j\omega/2})\cdot\frac{1}{2^{N-m}}\hat{\varphi}^{(N-m)}(\omega/2). \tag{33}$$

Using the property $H^{(m)}(e^{j(2l+1)\pi})=0$ for $m=0,\cdots,N-1$, we obtain a direct expression for the odd indexed derivatives in (17)

$$\begin{aligned}\hat{\varphi}^{(N)}(2\pi k)&=\hat{\varphi}^{(N)}(2\pi(2l+1))\\ &=\frac{\Pi_h^N}{2^N}\cdot\hat{\varphi}(\pi(2l+1)),\quad \text{(for } k=2l+1 \text{ odd)}\end{aligned} \tag{34}$$

where Π_h^N is the value of the first nonzero derivative of H

$$\Pi_h^N=H^{(N)}(e^{j\pi})=\left.\frac{\partial^N H(z)}{\partial z^N}\right|_{z=-1}. \tag{35}$$

To compute the right-most term in (34) for cases other than splines, we can use an approximation of the product (10) with a finite number of terms.

Likewise, by using the order property ii) and the fact that $H(e^{j2\pi l})=1$ in (33), we determine the remaining even indexed coefficients in (17) using the recursive rule

$$\begin{aligned}\hat{\varphi}^{(N)}(2\pi k)&=\hat{\varphi}^{(N)}(4\pi l)\\ &=\frac{1}{2^N}\cdot\hat{\varphi}^{(N)}(2\pi l),\quad \text{(for } k=2l \text{ even)}.\end{aligned} \tag{36}$$

TABLE I
RESCALED BOUND CONSTANT FOR DIFFERENT WAVELET FAMILES

N	Daubechies	closest-to-linear phase	coiflets	spline	Deslauriers-Dubuc
1	0.2887	0.2887		0.2887	
2	0.2236	0.2236	0.2124	0.07454	0.07454
3	0.2988	0.2988		0.03450	
4	0.5557	0.5557	0.4953	0.02182	0.1871
5	1.316	1.316		0.01734	
6	3.779	3.779	3.231	0.01655	1.212
7	12.74	12.74		0.01844	
8	49.35	49.35	40.92	0.02347	15.06
9	215.8	215.8		0.03362	

TABLE II
RELATIVE MAGNITUDE OF FIRST NON-VANISHING MOMENT (i.e., $Q(e^{j\pi}) = \Pi_h^N \cdot (2^N/N)$) FOR DIFFERENT WAVELET FAMILES

N	Daubechies	closest-to-linear phase	coiflets	spline	Deslauriers-Dubuc
1	1.	1.		1.	
2	1.732	1.732	1.646	1.	1.
3	3.162	3.162		1.	
4	5.916	5.916	5.272	1.	3.
5	11.225	11.225		1.	
6	21.49	21.49	18.38	1.	10.
7	41.42	41.42		1.	
8	80.22	80.22	66.51	1.	35.
9	156.	156.		1.	

For B-splines of degree $N-1$, $|\hat{\varphi}_N(\omega)| = |\sin(\omega/2)/(\omega/2)|^N$ and $H_N(z) = z^{-k_0} \cdot 2^{-N}(1+z)^N$, where k_0 is a suitable integer shift. Thus, the value of Π_h^N is simply $(N!)/2^N$, and it is not difficult to use these relations to show that

$$C_N^- = \frac{2}{N!}\left(\sum_{k=1}^{+\infty} \frac{(N!)^2}{(2\pi k)^{2N}}\right)^{1/2} = \sqrt{\frac{|B_{2N}|}{(2N)!}} \qquad (37)$$

where $|B_{2N}|$ is Bernouilli's number of order $2N$. Note that these constants are valid for any spline-based orthogonal [26], [27], semi-orthogonal [28]–[30], or biorthogonal [23] wavelet transforms.

D. Comparison of Multiresolution Representations

It is interesting to compute and compare the constants C_φ^- for various representation spaces, following the footsteps of SP [19]. The comparison obviously only makes sense for representation spaces that have the same order of approximation N and, hence, the same rate of convergence. We considered the same families of wavelet transforms as these authors, and our results in Table I are given in term of the rescaled constant $A_N^- = C_\varphi^- \cdot N!$ to facilitate the comparison. Except for the fact that our constant A_N^- is slightly below the sup-estimate A_N reported in [19], the conclusions are essentially the same.

The spline wavelets have by far the smallest constant. Both types of Daubechies wavelets (extremal phase and closest-to-linear phase) have the same performance because the corresponding filters differ by a phase term only. If we use splines as our reference, we see that the relative performance of most families worsens for higher order N. Part of this effect can be attributed to the increased magnitude of the first nonvanishing moment Π_h^N (cf. Table II). For Daubechies wavelets with $N > 2$, the performance degradation is at least 2^N, which confirms an earlier finding of SP. In other words, the approximation quality of Daubechies wavelets at a given resolution will be no better than that of splines at half the resolution. Thus, for signals that are sufficiently smooth, one should be able to get away with one less level of resolution using spline wavelets. Considerations of this nature may turn out to be quite relevant for data compression.

E. Discussion

The results in Table I clearly indicate that some representations—splines, in particular—are more favorable than others for approximating smooth functions. The ingredients that are important for good asymptotic performance can be identified by having a closer look at (34) and (35). Using the basic factorization $H(z) = 2^{-N}(1+z)^N \cdot Q(z)$ (cf. Proposition 2.4), we can compute the quantity Π_h^N defined by (35) as follows:

$$\Pi_h^N = \sum_{k=0}^{N} \binom{N}{k} H_N^{(k)}(e^{j\omega}) Q^{(N-k)}(e^{j\omega})\big|_{\omega=\pi}$$
$$= H_N^{(N)}(e^{j\omega}) \cdot Q(e^{j\omega})\big|_{\omega=\pi} = \frac{N!}{2^N} \cdot Q(e^{j\pi})$$

where $H_N(z) = 2^{-N}(1+z)^N$ is the refinement filter for splines of order N. Substituting this expression in (34), we find that for k odd

$$\hat{\varphi}^{(N)}(2\pi k) = \left(\frac{N!}{2^{2N}}\right) \cdot Q(e^{j\pi}) \cdot \hat{\varphi}(\pi k). \qquad (38)$$

Thus, in order to have a small asymptotic constant, it is preferable to have $Q(e^{j\pi})$ small and $\hat{\varphi}$ decaying rapidly. This last property is primarily dependent on the regularity of the scaling function. Specifically, if $\varphi \in C^m$ (m times continuously differentiable), then its Fourier transform decays at least as $O(\omega^{-m})$. This can be shown simply by considering the Fourier transform of $\varphi^{(m)}(x)$, which is given by $(j\omega)^m\hat{\varphi}(\omega)$. If $\varphi^{(m)} \in L_1$, then

$$|\omega|^m|\hat{\varphi}(\omega)| = \left|\int_{-\infty}^{+\infty} \varphi^{(m)}(x)e^{-j\omega x}dx\right|$$
$$< \int_{-\infty}^{+\infty} |\varphi^{(m)}(x)|dx = C$$

which implies that $|\hat{\varphi}(\omega)| < C \cdot |\omega|^{-m}$. Note that the decay of $\hat{\varphi}$ is also dependent on the magnitude of $Q(e^{j\omega})$, which is typically greater than $Q(e^{j\omega})|_{\omega=0} = 1$.

Splines with $Q(z) = 1$ are extremely attractive from both of these perspectives; but are they really optimal? We can show that this is not the case in an absolute sense by proposing a first-order counterexample $H(z) = (1+3z+2z^2)/6$ with $Q(z) = (1+2z)/3$. The corresponding $|H(e^{j\omega})|$ is always positive except for a first-order zero at $\omega = \pi$, which implies that the lower frame bound A is positive (cf. Condition i) in Definition 2.1). Moreover, because $|Q(e^{j\omega})| \leq 1$, the Fourier transform of this scaling function decays at least like $O(\omega^{-1})$, which implies that the upper frame bound B is finite. Hence, φ is a valid first-order scaling function; its C_φ^- constant is 0.0596, which is four times better than a B-spline of degree zero (Haar function). Higher order functions that are smoother than splines can also be constructed by considering the N-fold convolution of $H(z)$ [31]. While this example demonstrates our point, it is of limited practical value because, for the same complexity, we can also implement a piecewise linear spline with $H_2(z) = 2^{-2}(1+2z+z^2)$, which buys us one more order of convergence. In fact, B-splines are optimal if we include the filter length constraint in the design. They are the shortest Nth-order scaling functions. Splines are also the most regular functions among the examples in Table I, which explains why they are so good at approximating functions.

Although most of this paper has emphasized the significance of the order properties of the representation that determine the approximation rate, it is satisfying to find that regularity has some importance as well because of the way in which it controls the magnitude of the constants involved.

V. Conclusion

The object of this paper has been discuss an important consequence of the order properties of wavelets, which is not often emphasized in signal processing. The main point is that the approximation error at scale $a = 2^i$ decays like $O(a^N)$, where N is the order of the representation space. Thus, in principle, one should be able to use fewer levels of resolution to approximate a signal using a higher order wavelet transform. The performance index N also corresponds to the number of vanishing moments of the analysis (or dual) wavelet. This characteristic behavior of the error gets translated into error bounds and asymptotic formulas that are valid for all biorthogonal wavelet expansions and convolution-based oblique projectors in general. One important feature is that the analysis function has almost no influence on the convergence properties, except perhaps on the magnitude of the constant involved. In fact, Sweldens had already shown that biorthogonal wavelet expansions are asymptotically optimal in the sense that the error is the same as in the orthogonal case. Here, we provided an extended error analysis and identified more general oblique projection operators that share the same optimality properties. While the convergence rate is entirely determined by the zero properties of the synthesis function, we have demonstrated that the regularity of the basis functions also has some reducing effect on the error via the constants involved. In this respect, it appears that splines are especially favorable. Specifically, by using splines instead of Daubechies' wavelets, one can potentially achieve the same quality of approximation with a representation at half the resolution.

The general L_2 error bounds that have been described are valid for any signal that is N times continuously differentiable. However, for larger values of the scale, these results are at best qualitative. One can get much more quantitative by using the constants of the present asymptotic analysis, but these results are only rigorously applicable for functions that are very smooth or once the scale gets sufficiently small.

Acknowledgment

The author would like to thank W. Sweldens for his prompt and encouraging responses to email messages, for providing the wavelet filter coefficients for the calculations in Tables I and II, and for his improvement of the proof of Theorem 4.1, which was incorporated in the final manuscript (cf. (30) to (31)). He would also like to thank G. Strang and I. Daubechies for very helpful discussions and expert advice as well as A. Aldroubi, whose stimulating comments led to the construction of a first-order scaling function with improved asymptotic performance and regularity.

References

[1] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, no. 4, pp. 11–38, Oct. 1991.
[2] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[3] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.
[4] S. G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.
[5] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
[6] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Rev.*, vol. 31, pp. 614–627, 1989.
[7] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, pp. 1544–1576, Nov. 1992.
[8] S. G. Mallat, "Multiresolution approximations and wavelet orthogonal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, no. 1, pp. 69–87, 1989.
[9] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 617–643, Mar. 1992.
[10] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 205–220, Apr. 1992.
[11] O. Rioul, "On the choice of wavelet filters for still image compression," in *Proc. Int. Conf. Acoust., Speech Signal Processing*, Minneapolis, Aug. 1993, pp. 550–553.
[12] P. Flandrin, "Wavelet analysis and synthesis of fractional brownian motion," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 910–917, Mar. 1992.
[13] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 910–917, Mar. 1992.
[14] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," in *Constructive Aspect of Functional Analysis*. Rome, Italy: Edizioni Cremonese, 1971, pp. 796–830.
[15] B. Jawerth and W. Sweldens, "An overview of wavelet based multiresolution analyses," *SIAM Rev.*, vol. 36, no. 3, pp. 377–412, Sept. 1994.
[16] S. E. Kelly, M. A. Kon, and L. A. Raphaël, "Pointwise convergence of wavelet expansions," *Bull. Amer. Math. Soc.*, vol. 30, no. 1, pp. 87–94, Jan. 1994.
[17] S. Cambanis and E. Masry, "Wavelet approximation of deterministic and random signals: Convergence properties and rates," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1013–1029, July 1994.
[18] W. Sweldens and R. Piessens, "Quadrature formulae and asymptotic error expansions for wavelet approximations of smooth functions," *SIAM J. Numer. Anal.*, vol. 31, no. 4, pp. 1240–1264, Aug. 1994.
[19] ——, "Asymptotic error expansions for wavelet approximations of smooth functions II," *Numer. Math.*, vol. 68, no. 3, pp. 377–401, 1994.
[20] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Funct. Anal. Optimization*, vol. 15, nos. 1–2, pp. 1–21, Feb. 1994.
[21] W. A. Light, "Recent developments in the Strang-Fix theory for approximation orders," *in Curves and Surfaces*, P. J. Laurent, A. L.

Méhauté, and L. L. Schumaker, Eds. Boston, MA: Academic, 1991, pp. 285–292.
[22] R.-Q. Jia and J. Lei, "Approximation by multi-integer translates of functions having global support," *J. Approx. Theory*, vol. 72, no. 2–23, 1993.
[23] A. Cohen, I. Daubechies, and J. C. Feauveau, "Bi-orthogonal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.
[24] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, no. 9, pp. 2207–2232, Sept. 1992.
[25] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.
[26] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions," *Commun. Math. Phys.*, vol. 110, pp. 601–615, 1987.
[27] P.-G. Lemarié, "Ondelettes à localization exponentielles," *J. Math. Pures Appl.*, vol. 67, no. 3, pp. 227–236, 1988.
[28] M. Unser, A. Aldroubi, and M. Eden, "On the asymptotic convergence of B-spline wavelets to Gabor functions," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 864–872, Mar. 1992.
[29] C. K. Chui and J. Z. Wang, "On compactly supported spline wavelets and a duality principle," *Trans. Amer. Math. Soc.*, vol. 330, no. 2, pp. 903–915, 1992.
[30] M. Unser, A. Aldroubi, and M. Eden, "A family of polynomial spline wavelet transforms," *Signal Processing*, vol. 30, no. 2, pp. 141–162, Jan. 1993.
[31] A. Aldroubi and M. Unser, "Families of multiresolution and wavelet spaces with optimal properties," *Numerical Functional Anal. Optimization*, vol. 14, nos. 5–6, pp. 417–446, 1993.

Michael Unser (M'89–SM'94) was born in Zug, Switzerland, on April 9, 1958. He received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology, Lausanne, Switzerland.

He is currently a Visiting Scientist with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, USA, which he joined in 1985. His research interests include the application of image processing and pattern recognition techniques to various biomedical problems, multiresolution algorithms, wavelet transforms, and the use of splines in signal processing.

Dr. Unser is the author of more than 60 published journal papers. He serves as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and is a member of the Image and Multidimensional Signal Processing Committee of the IEEE Signal Processing Society. He is also on the editorial board of *Signal Processing, Pattern Recognition*, and was an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1992 to 1995. He co-organized the 1994 IEEE-EMBS Workshop on Wavelets in Medicine and Biology, and serves as a conference chair for SPIE's Wavelet Applications in Signal and Image processing, which has been held annually since 1993. He received the research prize of the Brown–Boweri Corporation (Switzerland) for his thesis in 1984, and the IEEE Signal Processing Society's 1995 Best Senior Paper Award (IMDSP technical area) for a Transactions paper with A. Aldroubi and M. Eden on B-spline signal processing..