Wavelet Theory Demystified

Michael Unser, *Fellow, IEEE,* and Thierry Blu, *Member, IEEE*

***Abstract*—In this paper, we revisit wavelet theory starting from the representation of a scaling function as the convolution of a B-spline (the regular part of it) and a distribution (the irregular or residual part). This formulation leads to some new insights on wavelets and makes it possible to rederive the main results of the classical theory—including some new extensions for fractional orders—in a self-contained, accessible fashion. In particular, we prove that the B-spline component is entirely responsible for five key wavelet properties: order of approximation, reproduction of polynomials, vanishing moments, multiscale differentiation property, and smoothness (regularity) of the basis functions. We also investigate the interaction of wavelets with differential operators giving explicit time domain formulas for the fractional derivatives of the basis functions. This allows us to specify a corresponding dual wavelet basis and helps us understand why the wavelet transform provides a stable characterization of the derivatives of a signal. Additional results include a new peeling theory of smoothness, leading to the extended notion of wavelet differentiability in the L_p-sense and a sharper theorem stating that smoothness implies order.**

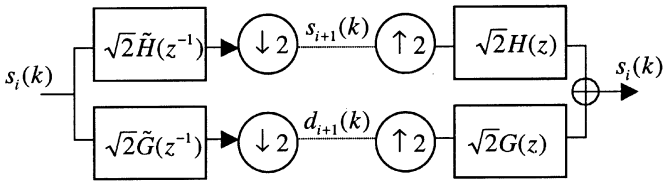
***Index Terms*—Approximation order, Besov spaces, Hölder smoothness, multiscale differentiation, splines, vanishing moments, wavelets.**

I. Introduction

THE mechanics of the wavelet transform are usually well understood by engineers working in signal processing. In essence, the system boils down to a two channel filterbank as shown in Fig. 1 [1], [2]. In the decomposition (or analysis) step, the digital signal $s_i(k)$ is split into two half-size sequences $s_{i+1}(k)$ and $d_{i+1}(k)$ by filtering it with a conjugate pair of lowpass and highpass filters ($\tilde{H}(z^{-1})$ and $\tilde{G}(z^{-1})$, respectively) and down-sampling the results thereafter. The signal is then reconstructed (synthesis step) by up-sampling, filtering, and summation of the components, as shown on the right-hand side of the block diagram; note that the analysis and synthesis procedures are flow-graph transposes of each other. A standard analysis shows that this kind of two channel decomposition is one-to-one and reversible provided that the z-transforms of the filters satisfy the perfect reconstruction (PR) equations also given in Fig. 1 [1], [3]. In the tree-structured wavelet transform, the decomposition step is further iterated on the lowpass component $s_{i+1}(k)$.

While the block diagram in Fig. 1 constitutes a valid description of the algorithm—different transforms simply correspond to different sets of PR filters—it tells us little about wavelets per se and their fundamental properties. To gain a deeper understanding, one needs to consider the continuous-time interpretation of the transform that is favored by mathematicians [4]–[6]. This representation involves continuously-defined basis functions (wavelets) that are rescaled and shifted versions of each other; it also gives rise to a filterbank implementation that is equivalent to the block diagram in Fig. 1—Mallat's fast wavelet algorithm [4], [7]. However, one of the important points of the continuous-time formulation is that the filters cannot be completely arbitrary. A key constraint is that the lowpass filter $H(z)$—also called the *refinement filter*—must be factorizable as $H(z) = 2^{-\gamma}\left(1+z^{-1}\right)^{\gamma} Q(z)$, which is an expression that involves some number of "regularity" factors $(1+z^{-1})$ as well as a stable residual term $Q(z)$ satisfying the lowpass constraint $Q(1) = 1$. A nontrivial factor $Q(z)$ is generally necessary for obtaining orthogonal or biorthogonal wavelets that can be implemented by means of FIR filters [8], [9]. The presence of the regularity term $\left(1+z^{-1}\right)^{\gamma}$ is essential for theoretical reasons. It is responsible for a number of key wavelet properties such as order of approximation, vanishing moments, reproduction of polynomials, and smoothness of the basis functions. If one excepts the vanishing moments, these are aspects of wavelets that are often not so well understood by signal processing practitioneers, mainly because the connection with the digital filterbank interpretation is not obvious. This is rather unfortunate as many consider these properties as the very core of wavelet theory (cf. [6, Ch. 7]); they are almost unavoidable if one wants to give precise explanations as to why wavelets work so well for approximating piecewise-smooth signals and why they are such a nice tool for characterizing singularites [10].

The purpose of this paper is to demystify some of these concepts by linking them to a factorization theorem involving the convolution between a B-spline (the regular part of the wavelet) and a distribution (the irregular or residual part). Indeed, we will show that it is the B-spline part—and nothing else—that brings



Conditions for a perfect reconstruction:

$$\tilde{H}(z^{-1})H(z) + \tilde{G}(z^{-1})G(z) = 1$$
$$\tilde{H}(z^{-1})H(-z) + \tilde{G}(z^{-1})G(-z) = 0$$

Fig. 1. Block diagram of the two-channel filterbank that forms the basis for the wavelet transform algorithm. From a purely digital signal processing point of view, the only constraints on the choice of the filters are the two perfect reconstruction equations.

Manuscript received December 12, 2001; revised September 24, 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Helmut Boelcskei.

The authors are with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland (e-mail: Michael.Unser@epfl.ch).

Digital Object Identifier 10.1109/TSP.2002.807000

1053-587X/03$17.00 © 2003 IEEE

in all the desired mathematical properties. Although the effect of the regularity factors is well understood by mathematicians working with wavelets, we are not aware of any deliberate effort to explain these properties from the perspective of B-splines. While it could be argued that this is essentially a matter of interpretation (regularity factors are equivalent to B-splines), we believe that the present formulation makes the whole matter more transparent and accessible. Our only prerequisite here is to have a complete understanding of the properties of B-splines, which is much easier than for other wavelets, since these are the only ones that have explicit formulas. We then use relatively simple manipulations to show that these properties carry over to all scaling functions through the convolution relation.

A. Scope and Organization of the Paper

This paper is written for researchers in signal processing who have a basic knowledge of wavelets and filterbanks and who would like to improve their understanding of the more theoretical aspects of wavelets that are usually confined to the mathematical literature. The paper is largely self-contained but assumes some familiarity with Mallat's multiresolution formulation of the wavelet transform that rests on the fundamental notion of scaling function [4]. From a theoretical point of view, the scaling function is more important than the wavelet because it provides the elementary building blocks for the decomposition; it is responsible for the key properties of the transform—and this is precisely what this paper is all about.

The presentation is organized as follows. In Section II, we recall the main definitions and mathematical concepts that are required for the L_2 formulation of the wavelet transform. We also spend some time describing the fractional B-splines, which play a central role in our formulation. In Section III, we rederive—and often extend—the key theorems of wavelet theory starting from the B-spline factorization of the scaling function. Finally, in Section IV, we investigate the smoothness issue and the interaction of wavelets with fractional differential operators. An important aspect of our formulation is that it allows for noninteger orders γ, which are not covered by traditional wavelet theory. This brings in two novelties:

1) the extension of the theory for fractional wavelets such as those introduced in [11];
2) a new "peeling" theory of smoothness which generalizes an interpretation of integer orders of differentiability given by Strang [6].

The motivation here is to come up with a more intuitive—and perhaps even more general—understanding of the concept of wavelet smoothness. Another important goal of this paper is to clarify the notion that the wavelet representation is stable with respect to (fractional) differentiation.

II. Scaling Functions and Wavelets in L_2

This section presents a brief review of the L_2 formulation and interpretation of the wavelet transform. It also contains a short primer on fractional B-splines.

A. Continuous Function Spaces and Notations

The continuous-time version of the wavelet transform applies to functions $f(x)$, $x \in R$ that are square integrable. The space of those functions is denoted by L_2; it is the Hilbert space that corresponds to the inner product

$$\langle f, g \rangle_{L_2} = \int_{-\infty}^{+\infty} f(x) g(x) dx \tag{1}$$

where the integral is taken in Lebesgue's sense. The energy of a function $f(x)$ is given by its squared L_2-norm: $\|f\|_{L_2}^2 = \langle f, f \rangle_{L_2}$; thus, the notation $f \in L_2$ is equivalent to the statement that $\|f\|_{L_2}^2$ is finite. More generally, one defines the L_p spaces for $1 \leq p \leq +\infty$ (where L stands for Lebesgue) as the set of functions whose L_p-norm

$$\|f\|_{L_p} = \left(\int_{-\infty}^{+\infty} |f(x)|^p \, dx \right)^{1/p} \tag{2}$$

is finite. These are Banach spaces; there is no corresponding inner product, except, of course, for $p = 2$.

The Fourier transform of $f(x)$ is denoted by $\hat{f}(\omega)$. If $f \in L_1$, it is given by $\int_{-\infty}^{+\infty} f(x) e^{-j\omega x} dx$. This definition can be extended for functions $f \in L_p$, as well as for generalized functions $\mu \in S'$, where S' stands for Schwarz' space of tempered distributions on $\mathbb{R}$ [12].

In wavelet theory, one usually considers some wavelet function $\psi(x)$ that generates a basis of L_2. What is meant here is that every function $f \in L_2$ can be represented in a unique and stable fashion using the expansion

$$f(x) = \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} d_i(k) \psi_{i,k}(x). \tag{3}$$

where the basis functions $\psi_{i,k}(x) = 2^{-i/2} \psi(x/2^i - k)$ are obtained by dilation (index i) and translation (index k) of the wavelet $\psi(x)$. The wavelet coefficients $d_i(k)$ are obtained by forming the (double infinite) sequence of inner products

$$d_i(k) = \left\langle f, \tilde{\psi}_{i,k} \right\rangle_{L_2} \tag{4}$$

where $\left\{ \tilde{\psi}_{i,k} \right\}_{i \in \mathbb{Z}, k \in \mathbb{Z}}$ is the biorthogonal basis of $\{\psi_{i,k}\}_{i \in \mathbb{Z}, k \in \mathbb{Z}}$ such that $\left\langle \tilde{\psi}_{i,k}, \psi_{j,l} \right\rangle = \delta_{i-j} \cdot \delta_{k-l}$. The biorthogonal basis is also generated by a single template $\tilde{\psi}(x)$—the so-called biorthogonal wavelet [8]. In practice, the inner products are never evaluated literally as in (4). They are computed much more efficiently using Mallat's algorithm, which is based on a hierarchical application of the filterbank decomposition in Fig. 1 [4]. The wavelet transform specifies the choice of a particular set of perfect reconstruction filters, and vice versa [cf. (6) and (8)].

B. Scaling Functions

The continuous-time interpretation of the wavelet decomposition algorithm in Fig. 1 is based on the fundamental concept of scaling function [4]. The scaling function $\varphi(x)$ is either given explicitly—as is the case with the B-splines—or it is derived

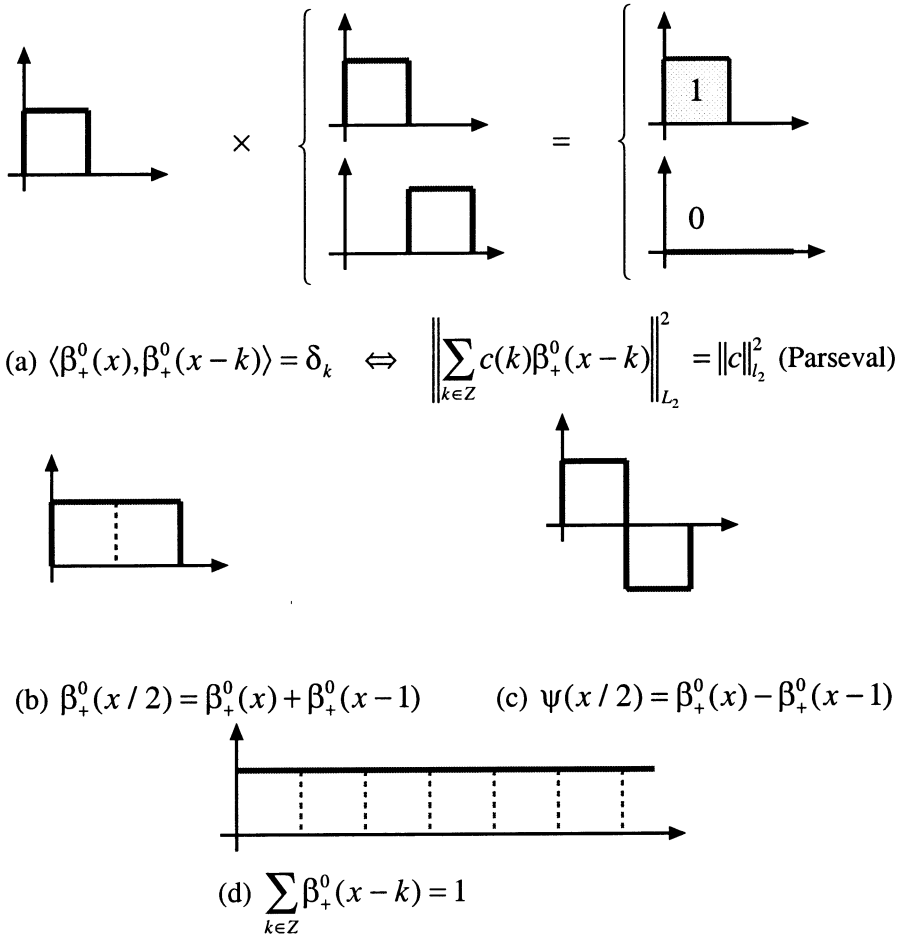


Fig. 2. Illustration of the key properties for the Haar case $\varphi(x) = \beta_+^0(x)$ (B-spline of degree zero). (a) Orthogonality. (b) Two-scale relation. (c) Wavelet relation. (d) Partition of unity.

indirectly from the *refinement filter* $H(z)$ as solution of the two-scale relation [9]. The difficulty with the latter approach is that not all filters do generate valid scaling functions [13], [14]. The mathematical requirements are the following.

Definition: $\varphi(x)$ is an *admissible scaling function* of L_2 if and only if it satisfies the following three conditions:

i) Riesz basis criterion; there exist two constants $A > 0$ and $B < +\infty$ such that

$$A \leq \sum_{n \in \mathbb{Z}} |\hat{\varphi}(\omega + 2\pi n)|^2 \leq B \tag{5}$$

ii) Refinability

$$\varphi\left(\frac{x}{2}\right) = 2 \sum_{k \in \mathbb{Z}} h(k) \varphi(x-k) \tag{6}$$

iii) Partition of unity

$$\sum_{k \in Z} \varphi(x-k) = 1. \tag{7}$$

These three conditions are satisfied by the Haar function $\beta_+^0(x)$—or B-spline of degree zero—as illustrated in Fig. 2. Condition i) ensures that φ generates a stable basis for the basic function space

$$V_{(0)} = \left\{ s(x) = \sum_{k \in \mathbb{Z}} c(k) \varphi(x-k),\ c \in l_2 \right\}.$$

In other words, there is an equivalence between the L_2-norm of the functions in $V_{(0)}$ and the ℓ_2-norm of their coefficients

$$\forall c \in l_2,\ A \cdot \|c\|_{l_2}^2 \leq \left\| \sum_k c(k) \varphi(x-k) \right\|_{L_2}^2 \leq B \cdot \|c\|_{l_2}^2$$

where A and B are the Riesz bounds of φ. The basis is orthonormal if and only if $A = B = 1$, as is precisely the case with the example in Fig. 2(a).

The two-scale relation (6), which is equivalent to $\varphi(x/2) \in V_{(0)}$, is the key to the multiresolution structure of the transform. It allows us to define the coarse-to-fine sequence of embedded subspaces $\{0\} \subset \cdots \subset V_{(i)} \subset \cdots \subset V_{(1)} \subset V_{(0)} \subset V_{(-1)} \subset \cdots \subset L_2$ such that $f_i(x) \in V_{(i)} \Leftrightarrow f_i(2^i x) \in V_{(0)}$.

The last, more technical partition of unity condition is necessary and sufficient (under mild conditions) for the approximation error to vanish as the wavelet size approaches zero [15]. It ensures that the multiresolution decomposition is dense in L_2.

We can prove that the above specification of a valid scaling function is equivalent to the axiomatic definition of a multires-

olution analysis of L_2 given by Mallat [7]. Hence, we have the guarantee that there exists some corresponding wavelet $\psi(x)$ that generates a Riesz basis of L_2. The wavelet is expressed as a linear combination of shifted versions of the scaling function:

$$\psi\left(\frac{x}{2}\right) = 2\sum_{k\in\mathbb{Z}} g(k)\varphi(x-k) \tag{8}$$

where the weights $g(k)$ correspond to the impulse response of the wavelet synthesis filter in Fig. 1. The dual scaling function $\tilde{\varphi}$ and wavelet $\tilde{\psi}$ are defined in a similar fashion using the dual versions of (6) and (8), respectively; these involve the analysis filters $\tilde{h}$ and $\tilde{g}$ in Fig. 1.

C. Fractional B-Splines

The simplest example of a scaling function is the B-spline of degree 0 (cf. Fig. 2). This function can be constructed by taking the difference of two step functions shifted by one. This yields the following mathematical representation:

$$\beta_+^0(x) := x_+^0 - (x-1)_+^0 = \Delta_+ x_+^0 \overset{\mathfrak{F}}{\longleftrightarrow} \hat{\beta}_+^0(\omega) = \left(\frac{1-e^{-j\omega}}{j\omega}\right) \tag{9}$$

where x_+^α denotes the one-sided power function $x_+^\alpha = \max\{0,x\}^\alpha$; Δ_+ is the causal[1] finite difference (convolution) operator whose impulse response is $\delta(x) - \delta(x-1)$. By convolving β_+^0 with itself n times, one generates the classical B-splines which are piecewise polynomials of degree n[16]. These are also valid scaling functions simply because the convolution of two scaling functions is a scaling function as well. It is still possible to go one step further by considering fractional convolution products that yield the fractional B-splines [11]:

$$\beta_+^\alpha(x) := \frac{\Delta_+^{\alpha+1} x_+^\alpha}{\Gamma(\alpha+1)} \overset{\mathfrak{F}}{\longleftrightarrow} \hat{\beta}_+^\alpha(\omega) = \left(\frac{1-e^{-j\omega}}{j\omega}\right)^{\alpha+1}. \tag{10}$$

The right-hand side of this formula is easy to understand: It is the $(\alpha+1)$th power of the Fourier transform in (9). The corresponding definition of the fractional power of a complex number z is $z^\alpha = |z|\, e^{j\alpha \arg(z)}$ with $\arg(z) \in [-\pi,\pi[$. The time-domain formula on the left-hand side can be obtained by inverse Fourier transformation (cf. [11]). The key operator here is the fractional causal finite difference $\Delta_+^{\alpha+1}$, which is best defined in the Fourier domain

$$\Delta_+^{\alpha+1} f(x) \overset{\mathfrak{F}}{\longleftrightarrow} \left(1-e^{-j\omega}\right)^{\alpha+1} \hat{f}(\omega). \tag{11}$$

To obtain its time domain representation, we expand $\left(1-e^{-j\omega}\right)^{\alpha+1}$ using Newton's generalized binomial theorem and perform an inverse Fourier transform

$$\Delta_+^{\alpha+1} f(x) = \sum_{k=0}^{+\infty} a_{\alpha+1}(k) f(x-k)$$
$$\text{with } a_{\alpha+1}(k) = (-1)^k \binom{\alpha+1}{k} \tag{12}$$

[1]the subscript "+" is used to signify that the corresponding function or convolution operator is causal; this is to distinguish them from the centered version of these operators which are often used in spline theory.

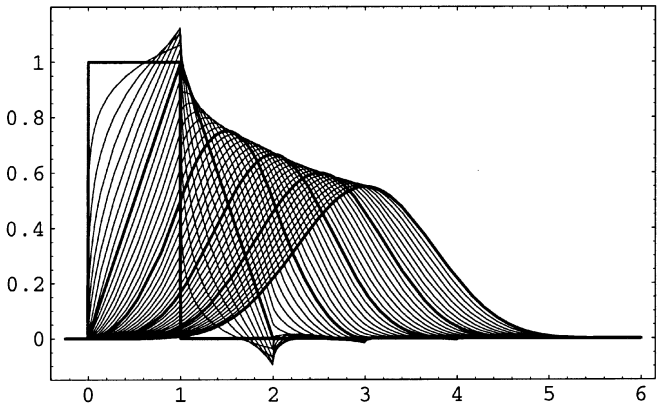


Fig. 3. Fractional B-splines $\beta_+^\alpha(x)$ for $\alpha \geq 0$. The polynomial ones (α integer) are represented using thicker lines.

where $\binom{u}{v} = \Gamma(u+1)/(\Gamma(v+1)\Gamma(u-v+1))$; $\Gamma(u+1)$ is Euler's Gamma function, which interpolates the factorial, i.e., $\Gamma(k+1) = k!$ for k integer. If we specialize these formulae for the more standard integer case ($\alpha = n$), then (10) yields the classical finite difference formula for the polynomial B-splines (cf. [17]).

The fractional B-splines are shown in Fig. 3; they provide a progressive transition between the polynomial ones (α integer) displayed in thicker lines. These functions have a number of interesting properties that are briefly summarized here—for more details, see [11].

- They generate valid Riesz bases for $\alpha > -1/2$. In particular, this means that they are square integrable, i.e., $\beta_+^\alpha \in L_2, \forall \alpha > -1/2$.
- They belong to L_p for all $p \geq 1$ whenever $\alpha \geq 0$, and for all $1 \leq p < 1/-\alpha$ if $\alpha < 0$.
- They satisfy the convolution relation: $\beta_+^{\alpha_1} * \beta_+^{\alpha_2} = \beta_+^{\alpha_1+\alpha_2+1}$. This comes as a direct consequence of their definition.
- They reproduce polynomials of degree $n = \lceil \alpha \rceil$, where $\lceil \alpha \rceil$ denotes the ceiling of α. In particular, they satisfy the partition of unity for $\alpha > -1$.
- They satisfy the two-scale relation. Their refinement filter is the generalized binomial. This is established easily by expressing the two-scale relation in the Fourier domain and evaluating the following ratio:

$$H_+^\alpha(e^{j\omega}) = \frac{\hat{\beta}_+^\alpha(2\omega)}{\hat{\beta}_+^\alpha(\omega)} = \left(\frac{1}{2}\cdot\frac{1-e^{-2j\omega}}{1-e^{-j\omega}}\right)^{\alpha+1} = \left(\frac{1+e^{-j\omega}}{2}\right)^{\alpha+1}. \tag{13}$$

In fact, the fractional B-splines share all the properties of the polynomial ones (cf. [18]) with two notable exceptions: They are not strictly positive nor compactly supported unless α is integer. For α noninteger, they decay like $x^{-\alpha-2}$ when $x \to +\infty$. Their lack of compact support is also apparent if we look at their refinement filter $h_+^\alpha(k) = 2^{-(\alpha+1)}\binom{\alpha+1}{k}$, which is infinitely supported whenever α is noninteger, although it decays like $k^{-(\alpha+2)}$, $k \to +\infty$.

D. Differentiation and Smoothness

Our primary motivation for considering fractional splines instead of conventional ones is that the enlarged family happens to be closed under fractional differentiation. This will prove extremely useful for understanding and characterizing the smoothness properties of scaling functions and wavelets.

Our fractional derivative operator ∂^γ is defined in the Fourier domain

$$\partial^\gamma f(x) \overset{\mathfrak{F}}{\longleftrightarrow} (j\omega)^\gamma \hat{f}(\omega). \tag{14}$$

It coincides with the usual derivative operator when γ is integer. Applying the operator to the fractional B-splines yields the following explicit differentiation formula

$$\partial^\gamma \beta_+^\alpha(x) = \Delta_+^\gamma \beta_+^{\alpha-\gamma}(x) \tag{15}$$

where Δ_+^γ is the fractional finite difference operator defined by (12). Thus, the fractional derivative of a B-splines happens to be spline with a corresponding decrease of the degree. The derivation of this result is rather straightforward if we work in the Fourier domain

$$\partial^\gamma \beta_+^\alpha(x) \overset{\mathfrak{F}}{\longleftrightarrow} (j\omega)^\gamma \cdot \left(\frac{1-e^{-j\omega}}{j\omega}\right)^{\alpha+1} = \left(1-e^{-j\omega}\right)^\gamma \left(\frac{1-e^{-j\omega}}{j\omega}\right)^{\alpha+1-\gamma}.$$

Again, this is an extension of the well-known differentiation formula for the polynomial B-splines.

Thus, if we differentiate $\beta_+^\alpha(x)$ α times, we end up with a function $\partial^\alpha \beta_+^\alpha = \Delta_+^\alpha \beta^0(x)$ that is piecewise constant and bounded (because $a_\alpha(k) \in l_1$ for $\alpha > -1$). The boundedness of the derivative is an indication that β_+^α is α-Hölder continuous. In fact, $\beta_+^\alpha(x)$ is infinitely differentiable everywhere except at the knots (integers), where it has discontinuities of order α (e.g., $(x-k)_+^\alpha$). Hölder smoothness is a pointwise measure of continuity that is often used to characterize wavelets [19]. In this context, the Sobolev smoothness, which is a more global measure of differentiability in the L_2-sense, is also important [6], [14], [20]. The critical Sobolev exponent for the fractional B-splines is $s_{\max} = \alpha + 1/2$. In other words, $\beta_+^\alpha(x)$ has $s_{\max}$ derivatives in the L_2-sense, i.e., $\left\|\partial^s \beta_+^\alpha\right\|_{L_2} < +\infty$ for $0 \le s < \alpha + 1/2$.

III. Wavelet Theory Revisited

Our goal in this section is to reformulate wavelet theory using a nontraditional point of view. Our argumentation is entirely based on a B-spline factorization theorem (Section III-B), which is intimately related to the crucial notion of approximation order. We will use this representation to derive the most important wavelet properties using relatively straightforward manipulations. In other words, we will show that the order of approximation of the transform (Section III-A), the reproduction of polynomials property (Section III-C), the vanishing moments of the wavelet (Section III-D), and the multiscale differentiation property of the wavelet (Section III-E) can all be attributed to the B-spline that lies hidden within.

A. Order of Approximation

A fundamental idea in wavelet theory is the concept of multiresolution analysis. There, one usually considers a sequence of embedded subspaces with a dyadic scale progression, i.e., $a = 2^i, i \in Z$. Specifically, the approximation space at scale $a = 2^i$ (or at resolution level i) is defined as

$$V_a = \left\{ s_a(x) = \sum_{k \in \mathbb{Z}} c(k) \varphi\left(\frac{x}{a} - k\right) : c(k) \in l_2 \right\}. \tag{16}$$

Because of the two-scale relation, the subspaces have the following inclusion property: $V_{2a} \subset V_a$. Given some input function $f \in L_2$, one considers its approximation in V_a. The best approximation in the least-squares sense (minimum L_2-norm) is given by the orthogonal projection into V_a

$$\forall f \in L_2,\ P_a f = \arg \min_{s_a \in V_a} \|f - s_a\|_{L_2} \in V_a. \tag{17}$$

Because of the Riesz basis property, the approximation $P_a f$ of f in V_a is always well-defined and unique (Projection theorem). A necessary condition for being able to construct wavelet bases of L_2 is that the approximation improves as the scale a gets closer to zero. In other words, we want to be able to approximate any L_2-function as closely as desired by selecting a scale (or sampling step) that is sufficiently small. It is therefore crucial to understand and characterize the rate of decay of the approximation error as a function of the scale. This leads to the notion of order of approximation, which plays a fundamental role in wavelet theory [6], [21], [22].

Definition: A scaling function φ has order of approximation γ if and only if

$$\forall f \in W_2^\gamma,\ \|f - P_a f\|_{L_2} \le C \cdot a^\gamma \|\partial^\gamma f\|_{L_2} \tag{18}$$

where C is a constant that may depend on φ but not on f; the condition $f \in W_2^\gamma$ (Sobolev space of order γ) ensures that $\|\partial^\gamma f\|_{L_2}$, which is the L_2-norm of the γth derivative of f, is finite so that the bound is well defined.

This definition means that for smooth signals, the L_2-approximation error should decay like the γth power of the scale. All popular families (Daubechies, Battle-Lemarié, etc.) satisfy the order constraint (29) and are indexed by their order γ, which is always a positive integer. The least sophisticated transform—Haar with $\varphi(x) = \beta_+^0(x)$—has a first order of approximation ($\gamma = 1$), which is enough for the approximation error to vanish as $a \to 0$. However, one usually prefers to work with higher order approximations which result in much faster convergence; typically, $\gamma = 4$ (e.g., the 9/7 Daubechies and cubic spline wavelets). While traditional wavelet theory requires γ to be integer, our definition is also valid for noninteger orders $\gamma \in R^+$ and applicable to the recent fractional extensions of wavelets [11].

The way in which the order of approximation translates into filter design constraints is made explicit by the following theorem.

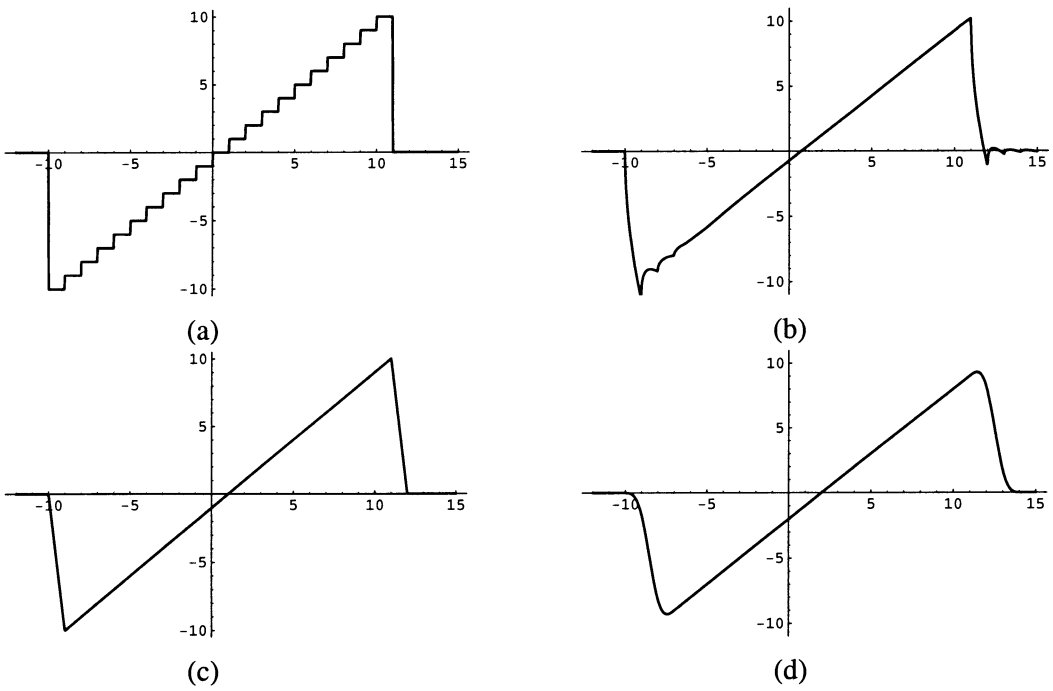


Fig. 4. Plot of $\sum_{k=-10}^{+10} k\beta_+^\alpha(x-k)$ for various values of α. (a) $\alpha = 0$ (piecewise constant). (b) $\alpha = 0.5$ (fractional spline). (c) $\alpha = 1$ (linear spline). (d) $\alpha = 3$ (cubic spline). This illustrates the property that the B-splines are able to reproduce a first order polynomial as soon as $\alpha > 0$ (i.e., cases b, c, and d, but not a).

Theorem 1: A *valid scaling function* $\varphi(x)$ has a γth-order of approximation if and only if its refinement filter $H(z)$ can be factorized as

$$H(z) = \underbrace{\left(\frac{1+z^{-1}}{2}\right)^{\gamma}}_{\text{spline part}} \cdot \underbrace{Q(z)}_{\text{distributional part}} \tag{19}$$

where $Q(z)$ is stable, i.e., $|Q(e^{j\omega})| < C$.

The restricted version of this theorem for φ compactly supported—or equivalently $H(z)$ FIR—with γ integer is a standard result in wavelet theory (cf. [6]). The important point here is that the present version holds for any real $\gamma > 0$ with minimal restriction on φ (Riesz basis). The proof is technical and will be published elsewhere [23].

B. B-Spline Factorization Theorem

The idea here is to interpret Theorem 1 in terms of B-splines. If we were to consider the first term of the factorization alone, then the refinement filter would generate a fractional B-spline of degree $\alpha = \gamma - 1$ [cf. (13)]. The $Q(z)$ filter alone, on the other hand, would generate some distribution $\varphi_0 \in S'$, which is typically not a true function (in fact, we will see later on that it cannot have any Sobolev regularity at all). The only thing we can say about φ_0 is that it is lowpass (because $Q(1) = 1$). Thus, because of the convolutional nature of the two-scale relation, we obtain the following corollary, which constitutes the main result on which we will build our wavelet theory.

Theorem 2 (B-Spline Factorization): φ is a valid scaling function of order γ if and only if its Fourier transform $\hat{\varphi}$ can be factorized as

$$\hat{\varphi}(\omega) = \hat{\beta}_+^{\gamma-1}(\omega) \cdot \hat{\varphi}_0(\omega)$$

where $\hat{\beta}_+^{\gamma-1}(\omega)$ is the Fourier transform of a fractional B-spline as given by (10), and $\hat{\varphi}_0(\omega)$ is a true function of ω bounded on every closed interval. Because of our assumptions on φ, this corresponds to a well-defined convolution product in the time domain

$$\varphi(x) = \left(\beta_+^{\gamma-1} * \varphi_0\right)(x) \text{ with } \varphi_0 \in S' \text{ and } \int \varphi_0(x)dx = 1. \tag{20}$$

It is therefore always possible to express a scaling function as the convolution between a B-spline and a distribution. What Theorem 2 also tells us is that the B-spline part is entirely responsible for the approximation order of the transform. We will now use the convolution relation (20) to show that the B-spline part brings in three other very useful properties.

C. Reproduction of Polynomials

We have already mentioned that the B-splines of degree α have the ability to reproduce polynomials of degree lesser or equal to $n = \lceil \alpha \rceil$. Practically, this means that we can generate all polynomials by taking suitable linear combination of B-splines. In particular, one can construct the following series of polynomials for $p = 0, \ldots, n = \lceil \alpha \rceil$:

$$\sum_{k \in \mathbb{Z}} k^p \beta_+^\alpha(x-k) = a_{p,0}x^p + \cdots + a_{p,p} \text{ with } a_{p,0} = 1 \tag{21}$$

which form a basis of the polynomials of degree n because $a_{p,0} \neq 0$. The ability of B-splines to reproduce a straight line ($p = 1$) is illustrated in Fig. 4.

The fact that (21) holds for the conventional case $\alpha = n$, with n integer, is not too surprising because the B-splines are themselves piecewise polynomial of degree n. The noninteger

case [cf. Fig. 4(b)] is less intuitive but is only truly relevant for fractional wavelets, which exhibit a few exotic properties [11].

The main point of this section is that the polynomial reproduction property is preserved through the convolution relation (20).

Proposition 3: Let φ_0 be any distribution such that $\int \varphi_0(x)dx = 1$ and $\int x^i \varphi_0(x)dx < +\infty$, $i = 1, \ldots, n$. Then, $\varphi(x) = \beta_+^\alpha * \varphi_0(x)$ reproduces the polynomials of degree lesser or equal to $n = \lceil \alpha \rceil$.

When γ is integer, Proposition 3 is a B-spline reformulation of a standard result in wavelet theory [5], [6]. What is novel here is the extension for $\gamma = \alpha + 1$ noninteger, which is nontrivial.

Proof: We assume that the residual filter q has sufficient (inverse polynomial) decay for the moments of φ_0 to be bounded up to order n. In other words, we want $\hat{\varphi}_0(\omega)$ to be n times differentiable at the origin. This mild technical requirement—which is much weaker than compact support—ensures that the convolutions and manipulations below are well-defined in the distributional sense.

We start by showing that the convolution between φ_0 and the monomial x^p produces a polynomial $q_p(x)$ of degree p

$$\begin{aligned}\varphi_0(x) * x^p &= \int_{-\infty}^{+\infty} (x-u)^p \varphi_0(u)du \\ &= \sum_{k=0}^{p} \binom{p}{k} x^{p-k}(-1)^k \underbrace{\int_{-\infty}^{+\infty} u^k \varphi_0(u)du}_{m_{\varphi_0}^k} \\ &= x^p + b_1 x^{p-1} + \cdots + b_p = q_p(x) \\ &\quad \text{with} \quad b_k = (-1)^k \binom{p}{k} m_{\varphi_0}^k.\end{aligned}$$

The leading coefficient is one because $Q(1) = 1 \Leftrightarrow m_{\varphi_0}^0 = \int \varphi_0(x)dx = \hat{\varphi}_0(0) = 1$. We then use this result to evaluate the sum

$$\begin{aligned}\sum_{k\in\mathbb{Z}} k^p \varphi(x-k) &= \sum_{k\in\mathbb{Z}} k^p \left(\varphi_0 * \beta_+^\alpha\right)(x-k) \\ &= \varphi_0 * \sum_{k=0}^{p} a_{p,k} x^{p-k} \quad \text{[using (20) and (21)]} \\ &= \sum_{k=0}^{p} a_{p,k} q_{p-k}(x) = x^p + c_1 x^{p-1} + \cdots + c_p.\end{aligned}$$

In addition, here, the leading coefficient is one implying that the sequence of these polynomials for $p = 0, \ldots, n$ is a basis for the polynomials of degree n. □

This polynomial reproduction property is illustrated in Fig. 5 for the Daubechies scaling function of order 2. Although no one will question the fact that the linear B-splines reproduce the constant and the ramp, it is much less obvious that the fractal-like Daubechies functions are endowed with the same property.

For sake of completeness, we mention the existence of a converse implication which goes back to the Strang–Fix theory of approximation.

Theorem 4: If a function $\varphi(x)$ reproduces the polynomials of degree n and is compactly supported, then $\varphi(x) = \beta_+^n * \varphi_0(x)$.

This result was first conjectured by Strang in [24] and was later proven by Ron in a more abstract framework [25]. A recent proof that is more accessible to an engineering audience is available in [26]. Note that Theorem 4—the converse part of Proposition 3—is only applicable in the conventional framework: φ compactly supported, and $n = \gamma - 1$. We have already encountered some counter examples with the fractional B-splines (i.e., $n = \lceil \gamma - 1 \rceil \neq \gamma - 1$), which are not compactly supported.

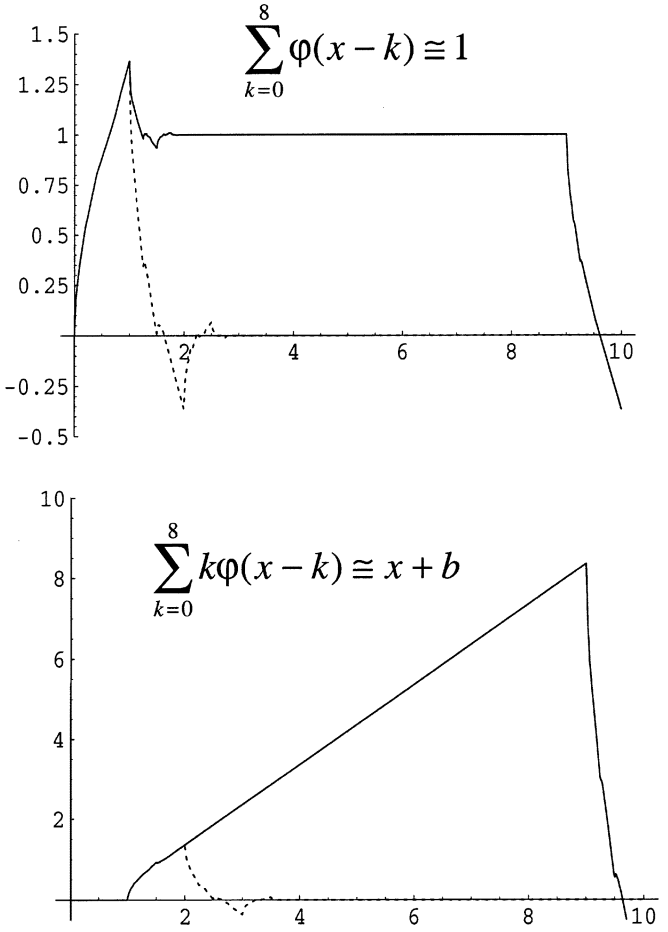


Fig. 5. Illustration of the reproduction of polynomials of degree 0 and 1 using Daubechies scaling functions of order $\delta = 2$.

In addition, note that the results in this section do not require the two-scale relation. As such, they are also applicable outside the wavelet framework—for instance, in the context of interpolation [27].

D. Vanishing Moments

A wavelet transform is said to have $n+1$ vanishing moments if the analysis wavelet $\tilde{\psi}$ is such that $\int x^p \tilde{\psi}(x)dx = 0$, $p = 0, \ldots, n$. These vanishing moments are undoubtedly one of the better known and most useful properties of wavelets [10]. In particular, they give the transform the ability to kill all polynomials of degree n [5], [21]. This property translates into a sparse representation of piecewise smooth signals because the wavelet coefficients will be essentially zero over all regions where the signal is well approximated by a polynomial, e.g., the first few terms of its Taylor series. This will produce streams of zero coefficients that can be coded with very few bits [28], [29]. The vanishing moments also allow the characterization of singularities based on the decay of the coefficients across scale—the

so-called persistence across scale property of the wavelet transform [10], [30], [31]. The study of this decay plays an important role in the analysis of fractal and multifractal signals [32].

The vanishing moments are nothing but an indirect manifestation of the ability of the scaling function to reproduce polynomials.

Proposition 5: If the scaling function $\varphi(x)$ reproduces the polynomials of degree n, then the analysis wavelet $\tilde{\psi}$ has $n+1$ vanishing moments.

This is a standard result in wavelet theory (cf. [5], [6]) that is rederived here for sake of completeness.

Proof: The polynomial reproduction property is equivalent to saying that the polynomials are in the linear span of $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$. Since the wavelet $\tilde{\psi}$ is perpendicular to $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ by construction (cf. the general biorthogonal wavelet formulation of Cohen-Feauveau-Daubechies [8]), it is therefore perpendicular to the polynomials; in particular, the monomials x^p with $p=0,\ldots,n$. □

Thus, by combining Propositions 3 and 5, we can claim that it is the B-spline component once again that is entirely responsible for the vanishing moments of $\tilde{\psi}$. The above argument can also be applied to the synthesis side of the transform. In other words, if the dual scaling function can be factorized as $\tilde{\varphi}(x) = \beta^{\tilde{\gamma}-1} * \tilde{\varphi}_0(x)$—meaning that $\tilde{\varphi}$ is of order $\tilde{\gamma}$—then the synthesis wavelet ψ will have exactly $\tilde{n}+1 = \lceil\tilde{\gamma}\rceil$ vanishing moments (general biorthogonal case). When the analysis and synthesis spaces are identical (orthogonal and semi-orthogonal cases), the number of vanishing moments are the same on both sides.

E. Multiscale Differentiator

We will now show that another consequence of the B-spline factor is that the analysis wavelet $\tilde{\psi}$ essentially behaves like a γth-order differentiator. The proof of this property rests almost entirely on the perfect reconstruction property of the filterbank in Fig. 1. In the sequel, we will assume that the four filters $\tilde{H}(z^{-1})$, $H(z)$, $\tilde{G}(z^{-1})$, and $G(z)$ are stable in the sense that their Fourier transforms are bounded. We also recall that the perfect reconstruction property has an equivalent l_2-formulation that is expressed by four biorthogonality relations between the various filters pairs, which can be written in matrix form, cf. [33]

$$\underbrace{\begin{pmatrix} G(z^{-1}) & G(-z^{-1}) \\ H(z^{-1}) & H(-z^{-1}) \end{pmatrix}^T}_{\mathbf{M}(z^{-1})^T} \underbrace{\begin{pmatrix} \tilde{G}(z) & \tilde{G}(-z) \\ \tilde{H}(z) & \tilde{H}(-z) \end{pmatrix}}_{\tilde{\mathbf{M}}(z)} = \mathbf{I}. \quad (22)$$

These are strong constraints that have a direct implication on the form of the wavelet filter $\tilde{G}(z)$.

Proposition 6: Under the constraint of a stable perfect reconstruction filterbank, the order condition (19) is equivalent to the following factorization of the wavelet analysis filter:

$$\tilde{G}(z) = (1-z)^{\gamma} \cdot P(z) \quad (23)$$

where the filter $P(z)$ is stable, i.e., $\left|P(e^{j\omega})\right| < C$. Another equivalent formulation is $\tilde{G}(e^{j\omega}) = O\left(|\omega|^{\gamma}\right)$.

Proof: Let us denote by $D(z)$ and $\tilde{D}(z)$ the determinant of the matrices $\mathbf{M}(z)$ and $\tilde{\mathbf{M}}(z)$ in (22), respectively. Then, clearly, $D(z^{-1})\tilde{D}(z) = 1$. By inverting $\mathbf{M}(z^{-1})^T$, we get

$$\tilde{G}(z) = \frac{H(-z^{-1})}{D(z^{-1})} = H(-z^{-1})\tilde{D}(z). \quad (24)$$

As a result, imposing (19) yields (23), with $P(z) = 2^{-\gamma}Q(-z^{-1})\tilde{D}(z)$, which is bounded on the unit circle by the stability hypothesis. Conversely, since $H(z) = \tilde{G}(-z^{-1})D(z)$, imposing (24) with the assumption that $P(z)$ is stable implies (19) with $Q(z) = 2^{\gamma}P(-z^{-1})D(z)$, which is stable as well. Finally, because $2/\pi \le \left|(1-e^{-j\omega})/j\omega\right| \le 1$ for $\omega \in [-\pi,\pi]$, we note that the factorization (23) with $P(z)$ stable is equivalent to the condition $G(e^{j\omega}) \le C\cdot|\omega|^{\gamma} = O(|\omega|^{\gamma})$. □

To get a better feel for this result, we consider (23) and take the limit to obtain the asymptotic form of the filter as $\omega \to 0$. Using the fact that $1-e^{j\omega} = -j\omega + O(\omega^2)$ and assuming that $P(e^{j\omega})$ is continuous, we obtain

$$\tilde{G}(e^{j\omega}) = B_0(-j\omega)^{\gamma}, \text{ as } \omega \to 0 \quad (25)$$

where $B_0 = P(1) = -2^{-\gamma}Q(-1)/G(-1) \neq 0$ because $G(-1) < \infty$ (stability) and $Q(-1) \neq 0$ (φ is of maximum order γ).

The transfer function of the analysis wavelet is obtained by taking the Fourier transform of (8) with $g = \tilde{g}$ and $\psi = \tilde{\psi}$:

$$\hat{\tilde{\psi}}(\omega) = \tilde{G}(e^{j\omega/2}) \cdot \hat{\tilde{\varphi}}\left(\frac{\omega}{2}\right). \quad (26)$$

Its low-frequency response obviously depends on the behavior of $\tilde{G}(e^{j\omega})$ near the origin.

Theorem 7: Let φ and $\tilde{\varphi}$ be two valid biorthogonal scaling functions with $\hat{\varphi}(\omega)$ and $\hat{\tilde{\varphi}}(\omega)$ continuous at $\omega = 0$. Then, φ is of order γ (i.e., $\varphi = \beta_+^{\gamma-1} * \varphi_0$) if and only if

$$\hat{\tilde{\psi}}(\omega) = O\left(|\omega|^{\gamma}\right).$$

The proof is given in the Appendix. If we now assume that $P(e^{j\omega})$ is continuous as well, we can easily obtain the asymptotic version of the result by plugging (25) into (26) and making use of the property $\lim_{\omega\to 0}\hat{\tilde{\varphi}}(\omega) = 1$. This yields

$$\begin{aligned} \hat{\tilde{\psi}}(\omega) =& C_{\gamma}\cdot(-j\omega)^{\gamma}, \text{ as } \omega \to 0 \\ \text{with } C_{\gamma} =& 2^{-\gamma}P(1) = -2^{-2\gamma}\frac{Q(-1)}{G(-1)} \neq 0. \end{aligned} \quad (27)$$

The practical implication of Theorem 7 is that the wavelet transform acts as a multiscale differentiator. Specifically, the wavelet coefficients will correspond to the samples of the γth derivative of a smoothed version of the input signal $f(x)$:

$$\langle f(x), \tilde{\psi}(x-u)\rangle = \partial^{\gamma}\left\{\phi * f\right\}(u)$$

where the smoothing kernel is defined by its frequency response $\hat{\phi}(\omega) = \hat{\tilde{\psi}}^*(\omega)/(j\omega)^{\gamma}$; it is necessarily lowpass and bounded because of (27) and Theorem 7. This kernel essentially limits the bandwidth of the signal being analyzed with two practical benefits. First, it regularizes the differentiation process by reducing its noise amplification effect, and second, it attenuates the signal components above the Nyquist frequency so that the differentiated signal is well represented by its samples (or wavelet coefficients).

In the classical case where $\gamma = L$ is an integer, there is an equivalence between (27) and the vanishing moment property. This can be shown by determining the Taylor series of $\hat{\tilde{\psi}}(\omega)$ around $\omega = 0$: $\hat{\tilde{\psi}}(\omega) = C_L \cdot (-j\omega)^L + O(\omega^{L+1})$, where $C_L = \int x^L \tilde{\psi}(x)dx/L!$ is proportional to the magnitude of the first nonvanishing moment.

In the fractional case, there is no Taylor series of order L and, therefore, no direct equivalence with the vanishing moments property. The unique property of this kind of wavelets is that they give access to fractional orders of differentiation, which can be very useful in some applications [34], in particular, when dealing with fractal-like signals.

IV. Wavelet Differentiability and Integrability

One of the primary reasons for the mathematical successes of wavelets is their stability with respect to differentiation and integration. In the conclusion of his classical monograph, Meyer writes: "everything takes place as if the wavelets $\psi(x/a)$ were eigenvectors of the differential operator ∂^α, with corresponding eigenvalue $a^{-\alpha}$" [35]. We will now use our B-spline formalism to shed some light on this important aspect of wavelet theory.

As starting point, we use the B-spline factorization (20) together with the B-spline convolution property ($\beta_+^{\gamma-1} = \beta_+^{\gamma-1-s} * \beta_+^{s-1}$) to manipulate the basic scaling function biorthogonality relation as follows:

$$\begin{aligned}\langle \tilde{\varphi}(x), \varphi(x-k)\rangle &= \langle \underbrace{(\beta_-^{s-1} * \tilde{\varphi})}_{\tilde{\varphi}_{\tilde{\gamma}+s}}(x), \underbrace{(\beta_+^{\gamma-1-s} * \varphi_0)}_{\varphi_{\gamma-s}}(x-k)\rangle \\ &= \delta_k \end{aligned} \tag{28}$$

where $\beta_-^\alpha(x) = \beta_+^\alpha(-x)$ denotes the anticausal B-spline of degree α. This formula suggests that the new pair of functions $(\tilde{\varphi}_{\tilde{\gamma}+s}, \varphi_{\gamma-s})$ should also generate a valid biorthogonal basis. In the sequel, we will show that these scaling functions play a crucial theoretical role for they provide the building blocks for the fractional integrals and fractional derivatives of the analysis and synthesis wavelets, respectively. For this purpose, we will first investigate the extent to which φ (or equivalently ψ) is differentiable and propose a peeling formulation of smoothness that provides some new insights on the various notions of wavelet regularity. We will also prove that the B-spline component is entirely responsible for the smoothness of the basis functions. Finally, in Section IV-C, we will specify the biorthogonal wavelet basis that is associated with the fractional differentiation operator.

A. Smoothness: Peeling Theory

In this subsection, we characterize the fractional derivatives of the scaling function and show that the presence of the B-spline component is absolutely necessary for these to be well-defined. The argument is entirely based on the convolution property of B-splines (cf. Section II-C). Specifically, we rewrite the B-spline factorization (20) in terms of the function $\varphi_{\gamma-r}$ that are already encountered in (28):

$$\varphi_\gamma(x) = \beta_+^{r-1} * \varphi_{\gamma-r}(x) \text{ where } \varphi_{\gamma-r}(x) = \beta_+^{\gamma-r-1} * \varphi_0(x) \tag{29}$$

with a new subscript notation that makes the order of the functions explicit, e.g., $\varphi = \varphi_\gamma$. We then use this relation to compute the rth derivative of φ for $0 \le r \le \gamma$:

$$\partial^r \varphi_\gamma(x) = \partial^r \beta_+^{r-1} * \varphi_{\gamma-r}(x).$$

Using (15) with $\gamma = r$ and $\alpha = r-1$ and noting that $\beta^{-1}(x) = \delta(x)$, which is consistent with (10), we get

$$\partial^r \varphi_\gamma(x) = \Delta_+^r \varphi_{\gamma-r}(x). \tag{30}$$

This explicit time-domain differentiation formula is known for r integer, but its extension for arbitrary r is new to the best of our knowledge. This result indicates that the function $\varphi_{\gamma-r}$ generates a basis for representing the fractional derivatives of the scaling function and of the wavelet; indeed, by linearity, we have that $\partial^r \{\psi(x/2\} = 2\sum_k g(k)\partial^r \varphi(x-k)$. The question that then arises is the following: How far can we differentiate, or equivalently, how much B-spline can we "peel off" before the so-called residue $\varphi_{\gamma-r}$ in (29) really blows up? What we mean by "blowing up" can be made mathematically precise by requiring that some L_p-norm of the residue $\varphi_{\gamma-r}$ remains finite. This particular interpretation turns out to be intimately related to the concept of smoothness in a general L_p-sense and leads to the following smoothness characterization theorem.

Theorem 8: If $\varphi_\gamma(x) = \beta_+^{r-1} * \varphi_{\gamma-r}(x)$ with $\varphi_{\gamma-r} \in L_p$, then $\partial^r \varphi_\gamma \in L_p$, i.e., φ_γ has r derivatives in the L_p sense.

Proof: First, we rewrite (30) as

$$\partial^r \varphi_\gamma(x) = \Delta_+^r \varphi_{\gamma-r}(x) = \sum_{k\in\mathbb{Z}} a_r(k)\varphi_{\gamma-r}(x-k)$$

where the coefficients of the finite difference operator are given by (12). We also recall that $a_r(k) \approx O(1/k^{r+1})$, which implies that the sequence $a_r(k)$ is absolutely summable (i. e., $a_r \in l_1$ for $r > 0$). We then use the above formula in conjunction with Minkowsky's inequality to obtain the following bound for the L_p-norm of the derivative

$$\|\partial^r \varphi_\gamma\|_{L_p} \le \sum_{k\in\mathbb{Z}} |a_r(k)| \cdot \|\varphi_{\gamma-r}(x-k)\|_{L_p} = \|a_r\|_{l_1} \cdot \|\varphi_{\gamma-r}\|_{L_p}$$

which proves the desired results. □

Theorem 8 provides an explicit link between the smoothness properties of wavelets and the B-spline factorization. It is also interesting because it yields a general and coherent approach to the concept of smoothness, i.e., fractional differentiability in the L_p-sense. For $p = 2$, the present definition of smoothness is equivalent to the widely used Sobolev regularity [36]. Another interesting case is $p \to \infty$ because it penalizes the worst case (max norm); this is very close to Hölder regularity, even though the latter is a measure of continuity rather than of differentiability [19], [37].

Our peeling theory of smoothness is illustrated in Fig. 6 for the case of Daubechies' scaling function of order 2. It is clear from the graph that the residue $\varphi_{\gamma-r}$ becomes rougher as a higher order B-spline component gets pulled out. These various plots were obtained by running a Fourier version [38] of the cascade algorithm with ten levels of iteration on the residual factor of the refinement filter corresponding to $\varphi_{\gamma-r}$ in (29). Past the limiting case ($r = 0.55$) [37], $\varphi_{\gamma-r}$ is no longer bounded, and it does not make much sense to attempt to represent it graphically.

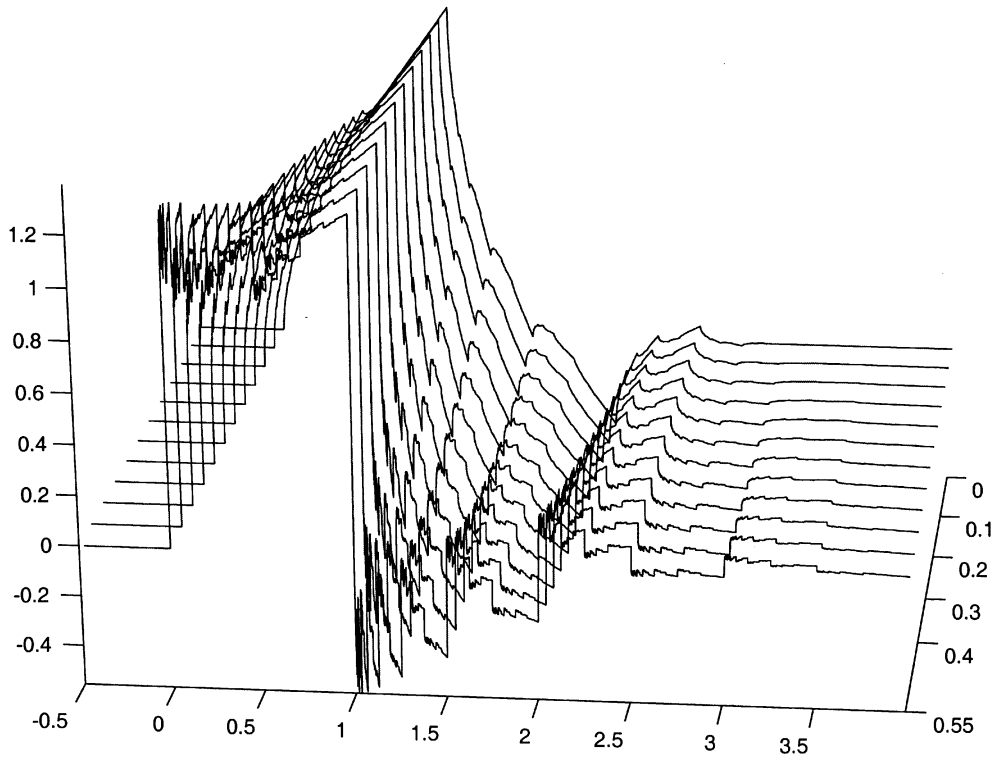


Fig. 6. Residual factor $\varphi_{\gamma-r}(x)$ in (29) as a function of r for Daubechies' scaling function of order $\gamma = 2$. The critical value $r_\infty = 0.55$ corresponds to the limiting case where $\varphi_{0,r}(x)$ is bounded (r derivatives in L_∞-sense). It also yields the Hölder exponent.

Note that there are computational techniques for determining this critical Hölder exponent (cf. Section IV-B).

To get some further insight on our L_p notion of smoothness, we now consider the example of the fractional B-splines. The borderline cases here are the B-splines of negative degree; these can be shown to belong to the following functional spaces:

$$\beta_+^{-u}(x) \in L_p, \text{ for } p > \frac{1}{u} \text{ and } 0 \le u < 1.$$

For the fractional splines, we have $\gamma = \alpha + 1$, so that we can write (29) as $\beta_+^\alpha = \beta_+^{r-1} * \beta_+^{\alpha-r}$, where $\beta_+^{\alpha-r} = \varphi_{\gamma-r}$ is the corresponding residue. Applying Theorem 8, we conclude that β^α has up to $s(p) = \alpha + 1/p$ derivatives in the L_p-sense with the strict equality only being reached for $p = \infty$. Indeed, we have already mentioned that $\partial^\alpha \beta_+^\alpha(x)$ is bounded (α derivatives in L_∞), which is consistent with the fact that β_+^α is α-Hölder continuous.

There is also a converse part to Theorem 8, albeit only within the restricted framework of wavelets.

Theorem 9: If φ is a valid scaling function such that $\partial^s \varphi \in L_2$, then $\varphi(x) = \beta_+^{s-1} * \varphi_0(x)$ with $\varphi_0 \in L_2$.

The proof is more technical and is given in the Appendix . Since the order of the B-spline factor β^{s-1} in Theorem 9 is s, the above statement is equivalent to saying that the order of a valid scaling function φ is necessarily greater or equal to its critical Sobolev exponent $s_{\max}$. In other words, we have that $\partial^s \varphi \in L_2 \Rightarrow \gamma \ge s, \forall s < s_{\max}$. This is a new result that extends a classical theorem in wavelet theory stating that C^m-continuity (m integer) implies some minimum approximation order: $\varphi \in C^m \Rightarrow \gamma \ge m$ [5]. Note that our result is slightly more generous (i.e., it yields more order); this follows from the Sobolev inequality (for compactly supported functions) $s_{\max} - 1/2 \le \alpha < s_{\max}$, where $\alpha \ge m$ is the Hölder exponent of φ. In fact, our bound is sharp, as demonstrated by the example below.

The B-splines of order γ are very regular, but they fall short of the maximum possible Sobolev smoothness by 1/2. Yet we can perturb the B-splines to achieve the maximum possible smoothness by taking $Q(z) = \left(1 + \varepsilon + z^{-1}\right)/(2 + \varepsilon)$ in (19), where $0 < \varepsilon < 1$ is small but nonzero. Since we are dealing with a two-tap residual, the Hölder (alias L_∞) and Sobolev (alias L_2) exponents can be computed explicitly using the technique outlined in [19] and [37]. Specifically, we find that

$$\begin{aligned}
\alpha &= \gamma - \log_2\left(\frac{2+2\varepsilon}{2+\varepsilon}\right) \\
&= \gamma - \frac{\varepsilon}{2\ln 2} + O(\varepsilon^2), \text{ for } 0 \le \varepsilon < 1 \\
s_{\max} &= \gamma - \frac{1}{2}\log_2\left(2 \cdot \frac{2+2\varepsilon+\varepsilon^2}{(2+\varepsilon)^2}\right) \\
&= \gamma - \frac{\varepsilon^2}{8\ln 2} + O(\varepsilon^4), \text{ as } \varepsilon \to 0^+.
\end{aligned}$$

Thus, by letting ε tend to 0^+, we are able to saturate the above inequalities such that $\gamma = s_{\max} = \alpha$, which proves that the bound in Theorem 9 is sharp. If we now push it a little further by letting $\varepsilon = 0$, we increase the B-spline factor by one which makes us jump to $s_{\max} = \gamma + 1/2$, while the Hölder continuity remains at $\alpha = s(\infty) = \gamma$, but at the same time, the order is also increased by one when compared with the previous degenerate case.

The functions that saturate the inequality (i.e., $s_{\max} = \gamma$) are the only ones for which the residual in (29) can be in L_2. Indeed,

if φ_0 would have some residual Sobolev smoothness, it would be possible, by Theorem 9, to factor out a larger order B-spline that would contradict the assumption that φ is of maximum order γ. For all other cases, φ_0 is not in L_2 and is typically only a distribution; for instance, $\varphi_0(x) = \delta(x) \in S'$ in the case of the B-splines. We can therefore safely state that φ_0 in (20) has no Sobolev regularity at all, which also means that there is no smoothness possible without the B-spline part of the wavelet.

While smoothness implies approximation order, there is generally no such implication in the reverse direction [39]. In other words, the presence of the B-spline factor is not enough to guarantee that the scaling function is in L_2, not to mention its differentiability. Indeed, one can conceive of very irregular wavelets where the distributional part φ_0 is so rough that it counterbalances the inherent smoothness of the B-spline factor. This is the reason why smoothness is relatively difficult to control when applying conventional filter design procedures; it is a property that is usually checked *a posteriori*.

B. Determining Wavelet Smoothness

The peeling theory provides an interpretation of wavelet smoothness that is appealing intuitively. However, we should not be fooled by the apparent simplicity of the argument. Wavelet smoothness is and remains one of the most difficult theoretical aspects of wavelet theory. While Theorem 8 provides a mathematical criterion for testing differentiability in the L_p-sense, it does not give a numerical method for determining the critical exponents $s(p)$ for a given filter $H(z)$. As far as we know, an exact computation is only possible for determining the Sobolev index (i.e., L_2-smoothness) of compactly supported basis functions. The preferred method is based on the determination of the spectral radius of the so-called transition operator; in practice, this amounts to computing the maximum eigenvalue of a reduced transition matrix associated with the residual filter $Q(z)$ [6], [20]. Various techniques have also been proposed to estimate the Hölder exponent (alias L_∞-smoothness) [37], [19]. These methods are not exact anymore (except when φ is symmetrical) but can provide tight upper and lower bounds.

Determining fractional orders of smoothness in norms other than L_2 is more challenging mathematically, especially since there is no single universal definition of fractional differentiation in the time domain. Mathematicians have turned the difficulty by testing the appartenance of the function to some "smoothness" spaces. Villemoes [20] proposed to study wavelet smoothness using Besov spaces, which leads to the determination of a critical Besov exponent that is qualitatively similar to the notion of L_p-smoothness that we have considered here. His approach is based on the fact that wavelets provide unconditional bases for Besov spaces [35]. More recently, Micchelli and Sauer have proposed to extend the Hölder notion of continuity to L_p-space by introducing what they call generalized Lipschitz spaces [40]. Their formulation is general but also rather involved—about 90 pages of dense mathematics. By contrast, our approach to measuring wavelet smoothness is much closer to a classical L_p-Sobolev analysis, except that the traditional method for $p \neq 2$ is restricted to integer orders of differentiation. What makes the Sobolev technique applicable here is our explicit time-domain formula (30) for fractional differentiation, which is specific to scaling (or refinable) functions.

C. Biorthogonal Wavelet Basis for Fractional Differentiation

The stability of wavelets with respect to differentiation is a prerequisite for using them to characterize smoothness spaces, in particular, for proving that they provide unconditional bases for Sobolev, Hölder, and Besov spaces [35]. Although these functional aspects of wavelet theory are beyond the scope of this paper, we want to make an interesting connection by identifying and characterizing the biorthogonal wavelet basis associated with the fractional differentiation operator.

The idea is simple and is based on the following manipulation of the wavelet biorthogonality relation:

$$\langle \tilde{\psi}(x), \psi(x-k) \rangle = \langle \partial_-^{-s} \tilde{\psi}(x), \partial^s \psi(x-k) \rangle = \delta_k \qquad (31)$$

where $\partial_-^{-s} \longleftrightarrow (-j\omega)^{-s}$ denotes the anticausal fractional integration operator; it is the adjoint of $\partial^{-s} \longleftrightarrow (j\omega)^{-s}$, which is the inverse of the fractional differentiation operator ∂^s. Based on the results in Section IV-A, we obtain the explicit form of the "derivative" wavelet

$$\psi^{(s)}(x) := \partial^s \psi(x) = 2 \sum_{k \in \mathbb{Z}} g_s(k) \varphi_{\gamma - s}(2x - k)$$

where $\varphi_{\gamma-s}$ is defined by (28), and $G_s(z) = 2^s (1 - z^{-1})^s G(z)$. We now close the loop by showing that the biorthogonal partner[2] of $\varphi_{\gamma-s}$ in (29), that is, the scaling function

$$\tilde{\varphi}_{\tilde{\gamma}+s} = \beta_-^{s-1} * \tilde{\varphi}_{\tilde{\gamma}}$$

provides the complementary basis for expanding $\partial_-^{-s} \tilde{\psi}$, which is the s-order fractional integral of the analysis wavelet. This manipulation is performed in the Fourier domain starting from (26)

$$\partial_-^{-s} \tilde{\psi}(x) \overset{\mathfrak{F}}{\longleftrightarrow} 2^{-s} \overbrace{\frac{\left(1 - e^{j\omega/2}\right)^s}{\left(-\frac{j\omega}{2}\right)^s}}^{\hat{\beta}_-^{s-1}(\omega/2)} \times \left(1 - e^{j\omega/2}\right)^{\gamma - s} P(e^{j\omega/2}) \hat{\tilde{\varphi}}\left(\frac{\omega}{2}\right).$$

The key here is the result of Proposition 6, which allows us to extract the relevant B-spline factor as long as $s \leq \gamma$. After inverse Fourier transformation, we end up with a stable, explicit representation

$$\tilde{\psi}^{(-s)}(x) = \partial_-^{-s} \tilde{\psi}(x) = 2 \sum_{k \in \mathbb{Z}} \tilde{g}_{-s}(k) \tilde{\varphi}_{\tilde{\gamma}+r}(2x - k)$$

where $\tilde{G}_{-s}(z) = 2^{-s}(1 - z)^{\gamma - s} P(z)$. Indeed, we can invoke Young's convolution inequality (cf. [42])

$$\|\tilde{\varphi}_{\gamma+s}\|_{L_p} \leq \left\|\beta_-^{s-1}\right\|_{L_1} \cdot \|\tilde{\varphi}_\gamma\|_{L_p}$$

which proves that $\tilde{\varphi}_\gamma \in L_p \Rightarrow \tilde{\varphi}_{\gamma+s} \in L_p$ (since $\beta_-^{s-1} \in L_1$, $s > 0$). Hence, we have established that the new wavelet

[2]This is a suggestive denomination that we are happy to borrow from Vaidyanathan [41].

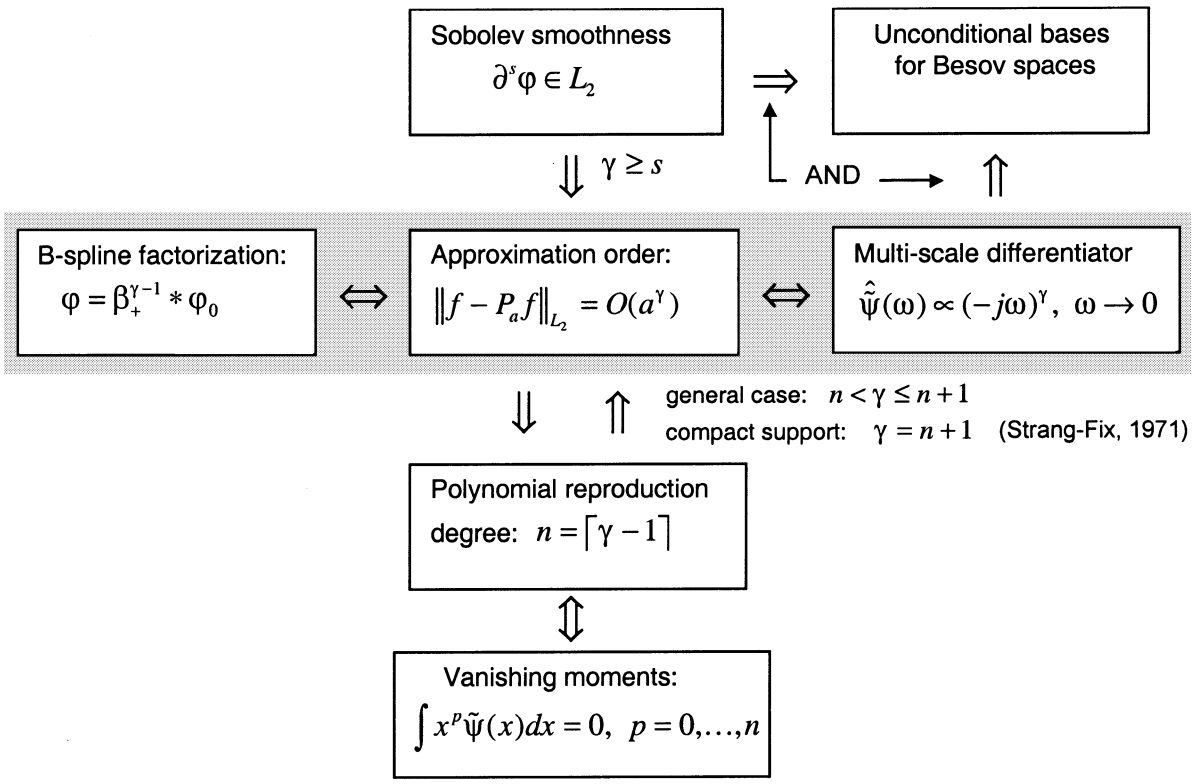


Fig. 7. Summary of wavelet theory: main properties and equivalences.

pair $(\tilde{\psi}^{(-s)}, \psi^{(s)})$ generates a biorthogonal basis that is in L_p provided that $\partial^s \varphi \in L_p$ and $\tilde{\varphi} \in L_p$. Note that it is essential here to work with wavelets because the fractional integral of a scaling function is generally not in L_p (since $\hat{\varphi}(0) \neq 0$).

Finally, by using the rescaling property $\partial^s \{f(x/a)\} = a^{-s} \partial^s \{f\}(x/a)$, we obtain the differentiation formulas for the basis functions:

$$\partial^s \{\psi_{i,k}\} = 2^{-is} \psi_{i,k}^{(s)} \text{ and } \partial_-^{-s} \left\{\tilde{\psi}_{i,k}\right\} = 2^{is} \tilde{\psi}_{i,k}^{(-s)} \quad (32)$$

which are the "eigen-relations" to which Meyer was alluding, even though he did not write them down explicitly. Of course, the qualifying statement is not rigorously correct—the important point is that there are basis functions with the same wavelet structure on both sides of the identity. The practical relevance of these "differential" wavelets is that they give us a direct way of gauging the fractional derivative of a signal based on its wavelet coefficients in the original basis. Specifically, we have that $\langle f, \tilde{\psi}_{i,k} \rangle = \langle \partial^s f, \partial_-^{-s} \tilde{\psi}_{i,k} \rangle = 2^{is} \langle \partial^s f, \tilde{\psi}_{i,k}^{(-s)} \rangle$, which implies that the original wavelet coefficients divided by 2^{is} are the coefficients of $\partial^s f$ in the modified wavelet basis $\{\psi_{i,k}^{(s)}\}$. This provides a strong hint as to why the Sobolev norm of a signal can be measured from the l_2-norm of its rescaled wavelet coefficients. The Besov case is analogous with l_p-norms being used instead; the argument there is more involved and relies on some Riesz-type L_p-norm equivalences [35]. In addition, note that the wavelets that have just been specified are fractional ones, which, in themseves, may serve as an *a posteriori* justification for our extended formulation.

V. Conclusion

In this paper, we proposed a new spline-based interpretation and derivation of some fundamental wavelet properties. Our argumentation entirely rests on the representation of a scaling function as the convolution between a B-spline and a distribution: $\varphi_\gamma = \beta^{\gamma-1} * \varphi_0$. The B-spline constitutes the regular part of the wavelet, and its presence is necessary (and sufficient) for the transform to have approximation order $\gamma > 0$ (cf. Theorem 2). In particular, this implies that one cannot build a multiresolution wavelet basis of L_2 without any B-spline factor. We have also shown that the key properties of B-splines (smoothness, differentiability, reproduction of polynomial, etc.) get carried over to the scaling function in an essentially mechanical fashion through the convolution relation. These results and the various equivalences that have been established are summarized in Fig. 7. The advantage of approaching the problem from this end is that these basic properties are relatively easy to understand for the B-splines because of their simple analytical form in both the time and frequency domain. This particular point of view leads to the conclusion that the primary mathematical wavelet properties—order of approximation, reproduction of polynomials, vanishing moments, and smoothness of the basis function—are entirely due to the B-spline component. Another interesting property that has been explained from the B-spline perspective is the stability of the wavelet basis with respect to fractional differentiation. Our key observation is that there is a direct correspondence between the process of moving a B-spline factor from one side to another in a pair of biorthogonal scaling functions and the exchange of fractional integrals/derivatives on their wavelet counterparts.

If these fundamental properties are to be attributed to the B-spline part exclusively, then what is the purpose of the distribution φ_0 corresponding to the residual filter $Q(z)$? This part is essential for imposing additional properties—such as orthonormality—and, more important, to balance the localization properties (size) of the analysis and synthesis filters in Fig. 1. B-splines are optimal in terms of size and smoothness, but they are not orthogonal. To construct a pure spline wavelet transform, one needs to orthogonalize the B-splines [43], [44] or to specify a dual pair of spline functions [45]–[47]; in both cases, this is equivalent to selecting a distributional part of the form $\varphi_0(x) = \sum_{k\in\mathbb{Z}} p(k)\delta(x-k)$, where $p(k)$ is an appropriate digital filter. However, the implementation of these (semi-orthogonal) spline transforms requires IIR filters, which is often considered a handicap. Early on, Daubechies and others [8], [9] have shown that the only way to construct wavelets and basis functions that are compactly supported on both sides (analysis and synthesis) is by careful selection of the factor $Q(z)$, eventually moving it to the analysis side, which yields biorthogonal spline wavelets [8]. This explains why all popular wavelet families (Daubechies, Coiflets, Cohen-Daubechies-Feauveau, 9/7, etc.) include nontrivial distributional factors $\tilde{\varphi}_0(x)$ and/or $\varphi_0(x)$, which we like to view as the price to pay for the additional but more classical filter design constraints (e.g., FIR filterbank and orthogonality).

APPENDIX
PROOFS

A. Proof of Theorem 7

If φ is of order γ, then $\tilde{G}(e^{j\omega}) = O(|\omega|^\gamma)$ by Proposition 6. This, together with the wavelet scaling (26) and the boundedness of $\hat{\tilde{\varphi}}(\omega)$ (consequence of the Riesz condition), implies that $\hat{\tilde{\psi}}(\omega) = O(|\omega|^\gamma)$.

Conversely, if $\hat{\tilde{\psi}}(\omega) = O(|\omega|^\gamma)$, then $\tilde{G}(e^{j\omega})\hat{\tilde{\varphi}}(\omega) = O(|\omega|^\gamma)$. Now, using the assumption that $\hat{\tilde{\varphi}}(\omega) > 0$ in a neighborhood of $\omega = 0$ (consequence of the continuity of $\hat{\tilde{\varphi}}(\omega)$ at $\omega = 0$ since $\hat{\tilde{\varphi}}(0) = 1$) and the stability of the filter $\tilde{G}(z)$, we can claim that $\tilde{G}(e^{j\omega}) = O(|\omega|^\gamma)$. This is also equivalent to (23), where $P(z)$ is stable, as seen in the proof of Proposition 6. Thus, from Proposition 6, this is equivalent to the γ-order property of φ. □

B. Proof of Theorem 9

We require one (very mild) technical assumption, namely, that the refinement filter $H(e^{j\omega})$ be continuous at $\omega = 0$. To prove Theorem 9, we define

$$B(\omega) = \sum_{k\in\mathbb{Z}} \left|\frac{\omega+2k\pi}{1-e^{-j\omega}}\right|^{2s} |\hat{\varphi}(\omega+2k\pi)|^2. \qquad (33)$$

This is a true function (possibly taking the value $+\infty$ for some ω) because it is a sum of positive functions. We observe that because of (6), $B(\omega)$ satisfies the scaling relation

$$B(2\omega) = |Q_0(e^{j\omega})|^2 B(\omega) + |Q_0(-e^{j\omega})|^2 B(\omega+\pi) \qquad (34)$$

where we have defined $Q_0(z) = 2^s H(z)/(1+z^{-1})^s$. In particular, this implies that

$$|Q_0(e^{j\omega})|^2 B(\omega) \le B(2\omega).$$

We now prove that $B(\omega)$ is integrable over $[-\pi,\pi]$.

First, we observe that $B(\omega)$ is integrable over any closed subset I of $[-\pi,\pi]$ that does not include 0. This is because $|1-e^{-j\omega}|^{-2s}$ is bounded over I, which implies that $B(\omega)$ is bounded by $\sum_{k\in\mathbb{Z}} |\omega+2k\pi|^{2s}|\hat{\varphi}(\omega+2k\pi)|^2$ up to a multiplicative constant. Now, because $|\omega|^s|\hat{\varphi}(\omega)|$ is in L_2

$$\sum_{k\in\mathbb{Z}} |\omega+2k\pi|^{2s}|\hat{\varphi}(\omega+2k\pi)|^2$$

is in $L_1([-\pi,\pi])$ and, thus, is in $L_1(I)$.

Second, the continuity of $|H(e^{j\omega})|$—and of $|1+e^{-j\omega}|^s$—at 0 implies that for any $\varepsilon > 0$—here, we choose $\varepsilon < 1/2$—we can find $0 < \omega_0 \le \pi$ such that for all $|\omega| \le \omega_0$, $|Q_0(e^{-j\omega})| \ge 2^{-\varepsilon}$. Consequently

$$\forall |\omega| \le \omega_0, \quad 2^{-2\varepsilon} B(\omega) \le B(2\omega).$$

Then, we define $b_n = \int_{I_n} B(\omega)\,d\omega$, where $I_n = 2^{-n}\omega_0 \times ([-1,-1/2]\cup[1/2,1])$. b_n are bounded quantities because the I_n's are closed subsets of $[-\pi,\pi]\setminus\{0\}$. We thus have $b_{n+1} \le 2^{2\varepsilon-1} b_n$, i.e., $b_n \le 2^{n\varepsilon-n} b_0$ or, by summing up all the contributions

$$\int_{|\omega/\omega_0|\in[1/2^n,1]} B(\omega)\,d\omega \le \frac{b_0}{1-2^{2\varepsilon-1}}$$

which is bounded, independently of n. Letting n tend to infinity, we conclude that $B(\omega)$ is integrable in $[-\omega_0,\omega_0]$ (Fatou's theorem), in particular, in the neighborhood of 0. Thus, $B(\omega)$ is fully integrable in $[-\pi,\pi]$.

It is now a simple matter to state that the function $\hat{\varphi}_0(\omega) = \hat{\varphi}(\omega)/\hat{\beta}^{s-1}(\omega)$ is in L_2. This is because

$$\int |\hat{\varphi}_0(\omega)|^2\,d\omega = \sum_{k\in\mathbb{Z}} \int_{-\pi}^{\pi} |\hat{\varphi}_0(\omega+2k\pi)|^2\,d\omega = \int_{-\pi}^{\pi} B(\omega)\,d\omega < \infty$$

where $B(\omega)$ is the same as defined by (33) and has just been shown to be integrable in $[-\pi,\pi]$. □

REFERENCES

[1] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction technique," *IEEE Acoust., Speech, Siganl Process. Mag.*, vol. 4, pp. 4–20, July 1987.

[2] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, pp. 11–38, Oct. 1991.

[3] M. Vetterli, "A theory of multirate filter banks," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-35, pp. 356–372, Mar. 1987.

[4] S. G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.

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

[6] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge, 1996.

[7] 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.

[8] A. Cohen, I. Daubechies, and J. C. Feauveau, "Bi-orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

[9] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.

[10] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1998.
[11] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Rev.*, vol. 42, no. 1, pp. 43–67, Mar. 2000.
[12] Y. Katznelson, *An Introduction to Harmonic Analysis*. New York: Dover, 1976.
[13] I. Daubechies and J. C. Lagarias, "Two-scale difference-equations .1. Existence and global regularity of solutions," *SIAM J. Math. Anal.*, vol. 22, no. 5, pp. 1388–1410, 1991.
[14] G. Strang, "Eigenvalues of $(\downarrow 2)H$ and convergence of the cascade algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 233–238, Feb. 1996.
[15] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, pp. 569–587, Apr. 2000.
[16] I. J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, vol. 4, pp. 45–99, 1946.
[17] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, vol. 16, pp. 22–38, Nov. 1999.
[18] C. de Boor, *A Practical Guide to Splines*. New York: Springer-Verlag, 1978.
[19] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, pp. 1544–1576, Nov. 1992.
[20] L. F. Villemoes, "Wavelet analysis of refinement equations," *SIAM J. Math. Anal.*, vol. 25, no. 5, pp. 1433–1460, 1994.
[21] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Rev.*, vol. 31, pp. 614–627, 1989.
[22] M. Unser, "Approximation power of biorthogonal wavelet expansions," *IEEE Trans. Signal Processing*, vol. 44, pp. 519–527, Mar. 1996.
[23] T. Blu and M. Unser, "Wavelet regularity and fractional orders of approximation," *SIAM J. Math. Anal.*, submitted for publication.
[24] G. Strang and G. Fix, "A fourier analysis of the finite element variational method," in *Constructive Aspects of Functional Analysis*. Rome, Italy: Edizioni Cremonese, 1971, pp. 793–840.
[25] A. Ron, "Factorization theorems for univariate splines on regular grids," *Israel J. Math.*, vol. 70, no. 1, pp. 48–68, 1990.
[26] T. Blu, P. Thévenaz, and M. Unser, "Complete parametrization of piecewise-polynomial interpolation kernels," *IEEE Trans. Image Processing*, submitted for publication.
[27] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.*, vol. 19, pp. 739–758, July 2000.
[28] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.
[29] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243–250, June 1996.
[30] S. Jaffard, "Pointwise smoothness, two-microlocalization and wavelet coefficients," *Publicacions Matemàtiques*, vol. 35, pp. 155–168, 1991.
[31] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, pp. 617–643, Mar. 1992.
[32] A. Arneodo, F. Argoul, E. Bacry, J. Elezgaray, and M. J. F., *Ondelettes, Multifractales et Turbulence*. Paris, France: Diderot, 1995.
[33] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 2207–2232, Sept. 1992.
[34] F. J. Herrmann, "Singularity characterization by monoscale analysis: Application to seismic imaging," *Appl. Comput. Harmon. Anal.*, vol. 11, no. 1, pp. 64–88, July 2001.
[35] Y. Meyer, *Ondelettes et Opérateurs I : Ondelettes*. Paris, France: Hermann, 1990.
[36] T. Eirola, "Sobolev characterization of solutions of dilation equations," *SIAM J. Math. Anal.*, vol. 23, pp. 1015–1030, 1992.
[37] I. Daubechies and J. C. Lagarias, "Two-scale difference-equations .2. Local regularity, infinite products of matrices and fractals," *SIAM J. Math. Anal.*, vol. 23, no. 4, pp. 1031–1079, 1992.
[38] T. Blu and M. Unser, "The fractional spline wavelet transform: Definition and implementation," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Istanbul, Turkey, 2000, pp. 512–515.
[39] A. Cohen, I. Daubechies, and A. Ron, "How smooth is the smoothest function in a given refinable space?," *Appl. Comput. Harmon. Anal.*, vol. 3, no. 1, pp. 87–90, 1996.
[40] C. A. Micchelli and T. Sauer, "Regularity of multiwaveletes," *Adv. Comput. Math.*, vol. 7, pp. 455–545, 1997.
[41] P. P. Vaidyanathan and B. Vrcelj, "Biorthogonal partners and applications," *IEEE Trans. Signal Processing*, vol. 49, pp. 1013–1027, May 2001.
[42] E. M. Stein and G. Weiss, *Fourier Analysis on Euclidean Spaces*. Princeton, NJ: Princeton Univ. Press, 1971.
[43] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions," *Commun. Math. Phys.*, vol. 110, pp. 601–615, 1987.
[44] P.-G. Lemarié, "Ondelettes à localization exponentielle," *J. Math. Pures et Appl.*, vol. 67, no. 3, pp. 227–236, 1988.
[45] 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.
[46] M. Unser, A. Aldroubi, and M. Eden, "On the asymptotic convergence of B-spline wavelets to Gabor functions," *IEEE Trans. Inform. Theory*, vol. 38, pp. 864–872, Mar. 1992.
[47] ——, "A family of polynomial spline wavelet transforms," *Signal Process.*, vol. 30, no. 2, pp. 141–162, Jan. 1993.

Michael Unser (M'89–SM'94–F'99) 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 (EPFL), Lausanne, Switzerland.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD. He is now Professor and Head of the Biomedical Imaging Group at EPFL. His main research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing. He is the author of 100 published journal papers in these areas.

Dr. Unser is Associate Editor-in-Chief for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He is on the editorial boards of several other journals, including IEEE SIGNAL PROCESSING MAGAZINE, *Signal Processing*, IEEE TRANSACTIONS ON IMAGE PROCESSING (from 1992 to 1995), and IEEE SIGNAL PROCESSING LETTERS (from 1994 to 1998). He serves as regular chair for the SPIE Conference on Wavelets, which has been held annually since 1993. He was general co-chair of the first IEEE International Symposium on Biomedical Imaging, Washington, DC, 2002. He received the 1995 Best Paper Award and the 2000 Magazine Award from the IEEE Signal Processing Society.

Thierry Blu (M'96) was born in Orléans, France, in 1964. He received the "Diplôme d'ingénieur" from École Polytechnique, Paris, France, in 1986 and from Télécom Paris (ENST), in 1988. In 1996, he received the Ph.D degree in electrical engineering from ENST for a study on iterated rational filterbanks applied to wideband audio coding.

He is currently with the Biomedical Imaging Group, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, on leave from France Télécom National Center for Telecommunications Studies (CNET), Issy-les-Moulineaux, France. His research interests include (multi)wavelets, multiresolution analysis, multirate filterbanks, approximation and sampling theory, mathematical imaging, etc.