Wavelets, Fractals, and Radial Basis Functions

Thierry Blu and Michael Unser, *Fellow, IEEE*

***Abstract*—Wavelets and radial basis functions (RBFs) lead to two distinct ways of representing signals in terms of shifted basis functions. RBFs, unlike wavelets, are nonlocal and do not involve any scaling, which makes them applicable to nonuniform grids. Despite these fundamental differences, we show that the two types of representation are closely linked together . . . through fractals. First, we identify and characterize the whole class of self-similar radial basis functions that can be localized to yield conventional multiresolution wavelet bases. Conversely, we prove that for any compactly supported scaling function $\varphi(x)$, there exists a one-sided central basis function $\rho_+(x)$ that spans the same multiresolution subspaces. The central property is that the multiresolution bases are generated by simple translation of ρ_+ without any dilation. We also present an explicit time-domain representation of a scaling function as a sum of harmonic splines. The leading term in the decomposition corresponds to the fractional splines: a recent, continuous-order generalization of the polynomial splines.**

***Index Terms*—Central basis functions, fractals, fractional splines, localization, multiresolution, radial basis functions, refinement filter, scaling functions, self-similarity, splines, tempered distributions, two-scale difference equation, wavelets.**

I. Introduction

RADIAL basis functions (RBFs) constitute a powerful tool for working with data that is nonuniformly sampled. They are used extensively for scattered data interpolation [1]. Thin-plate splines, in particular, have some remarkable variational properties: They minimize a Laplacian energy functional which makes the solution invariant to rigid-body coordinate transformations [2]. They are often used in medical imaging for the estimation of deformation fields based on the specification of fiducial points (or landmarks) [3]–[5]. Radial basis functions are also frequently applied to neural networks; they have been proposed as an efficient mean for establishing a multidimensional mapping between some input feature space and some target response given a collection of (noisy) input–output pairs (learning by example) [6], [7]. Their use has been justified through regularization theory in connection with their energy minimization properties [8]–[10].

A radial basis function approximation in p-dimensions ($\mathbf{x} \in \mathbb{R}^p$) has the generic form

$$f(\mathbf{x}) = \sum_{k \in \mathbb{Z}} a_k \rho(\|\mathbf{x} - \mathbf{x}_k\|) \tag{1}$$

where $\rho(r)$: $\mathbb{R}_+ \rightarrow \mathbb{R}$ is a univariate function and where $\|\mathbf{x} - \mathbf{x}_k\|$ denotes the Euclidean distance between the p-vectors $\mathbf{x}$ and $\mathbf{x}_k$. The basis functions in (1) depend only on the distance to their corresponding grid point $\mathbf{x}_k$ and are thus called *radial*. The a_k are weighting coefficients that are typically determined by fitting the function to some data. In the classical interpolation problem, the function f is determined such that $f(\mathbf{x}_k) = f_k$, where the f_k are some data values; in this case, there is exactly one linear constraint per basis function, and the corresponding linear system of equations is invertible under relatively mild conditions [11]. The better-known examples of radial basis functions are $\rho(r) = r$ (linear or membrane spline), $\rho(r) = r^2 \log r$ (thin-plate spline), and $\rho(r) = \sqrt{c^2 + r^2}$ (Hardy's multiquadric).

Here, we are interested in discussing the connection between such radial basis functions and wavelets. In this latter context, the generic multiresolution approximation of a function $f(x)$ at scale $a = 2^{-i}$ has the form

$$f_i(x) = \sum_{k \in \mathbb{Z}} c_i(k) \varphi(x/a - k) \tag{2}$$

where the basis functions at level $i \in \mathbb{Z}$ are dilated by the factor a and spaced accordingly. While (1) and (2) both involve shifts of some basic template function—$\rho(\|\mathbf{x}\|)$, and $\varphi(2^i x)$, respectively—there are some crucial differences that need to be emphasized. First, although it is easy to define a multiresolution hierarchy of radial basis functions by simply dropping or adding some points, there is no provision for scaling in (1). Second, the generating functions used in both cases are fundamentally different: $\varphi(x)$ is a relatively tame, well-localized function which is at least square-integrable; $\rho(r)$, on the other hand, typically extends over the whole axis and is unbounded at infinity. Conversely, $\rho(r)$ has a convenient closed form expression, whereas $\varphi(x)$ usually has no such formula (it is defined implicitly through a infinite recusion) [12]–[14]. Third, the radial basis function framework is ideally suited for a nonuniform setting, whereas conventional wavelet theory is restricted to uniform grids.

Yet, there is an interesting link between both types of representations. This was recognized early on by mathematicians working in approximation theory and has been the basis for several interesting generalizations of wavelets [15]. The first step in this realization was the construction multivariate scaling functions (or pre-wavelets) that are linear combinations of some radial basis functions [16], [17]. Utreras specified the class of refinable radial basis functions as those whose Fourier transform is $\hat{\rho}(\omega) = |\omega|^{-\gamma}$ [17]; this is a condition that is unnecessarily restrictive, as we will see here. The next step, while still working with a uniform grid, was to extend the multiresolution concept without the two-scale relation. This led to the construction of *nonstationary* wavelets where the basis functions at distinct scales are no longer dilates of each other

Manuscript received April 5, 2001; revised November 11, 2001. 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 DMT/IOA, Swiss Federal Institute of Technology, Lausanne, Switzerland (e-mail: thierry.blu@epfl.ch; michael.unser@epfl.ch).

Publisher Item Identifier S 1053-587X(02)01351-X.

1053–587X/02$17.00 © 2002 IEEE

[18], [19]. Finally, there have also been extensions of wavelets to nonuniform grids based again on radial basis functions [20], [21]. A univariate example of special interest that falls into this last category are the nonuniform spline wavelets [22], [23].

Our purpose in this paper is to strengthen the connection between radial basis functions and wavelets even further. We will mainly concentrate on the standard wavelet setting (approximation of univariate functions on a uniform grid) and approach the problem from its two different sides. First, in Section III, we will show how to construct scaling functions (or wavelets) starting from some radial basis function $\rho(r)$, extending some of the earlier results of Buhmann and Utreras. In particular, we will identify a self-similarity condition for $\rho(r)$ that justifies the use of the word *fractal* in the title. We will then characterize the whole class of these "fractal" functions and show how these can be localized to yield valid scaling functions. We will illustrate our theoretical results with the construction of fractional B-splines and other related functions. In Section IV, we will look at the converse side of the problem and prove that any standard multiresolution analysis of $\mathbf{L}^2(\mathbb{R})$ can be expressed in terms of central basis functions, which are a slight extension of the radial ones. In other words, we will uncover the central basis function that lies hidden behind any scaling function or wavelet. Finally, by combining our results, we will present a general parametrization of all compactly supported wavelets that will provide us with a better understanding of their fractal nature.

The next section is meant as an introduction. Rather than starting with mathematics right away, we will consider a concrete example that contains most of the flavor of what will be investigated in more generality in the later sections of the paper.

II. Motivation: The Example of Linear Splines

Here, we use piecewise linear functions to build a multiresolution analysis of $\mathbf{L}^2(\mathbb{R})$, but we proceed in a nonstandard fashion.

A. Splines and One-Sided Power Functions

The simplest conceivable linear spline is the one-sided ramp (or power) function

$$x_+ = \begin{cases} x, & x \geq 0 \\ 0, & \text{elsewhere} \end{cases} \tag{3}$$

which has exactly one knot (or discontinuity) at the origin [cf. Fig. 1(a)]. By simply shifting and adding up such building blocks, we can generate functions of the form

$$s(x) = \sum_k a_k (x - x_k)_+ \tag{4}$$

where the a_k are some (arbitrary) weights and the x_k an increasing sequence of knots. Since $s(x)$ in (4) is a sum of splines, it is a spline as well that inherits the discontinuities of its individual constituants. Hence, it is a piecewise linear function with knots at x_k. Thus, by placing the knots at the integers ($x_k = k$), we can specify the basic space[1] of uniform splines of degree 1

$$\mathcal{V}_0 = \text{span}_{k\in\mathbb{Z}}\{\rho_+(x-k)\} \tag{5}$$

[1]In this paper, $\text{span}_{k\in\mathbb{Z}}\{\rho(x-k)\}$ stands for the function space of all the linear combinations of integer shifts of $\rho(x)$ that are *locally square integrable*.

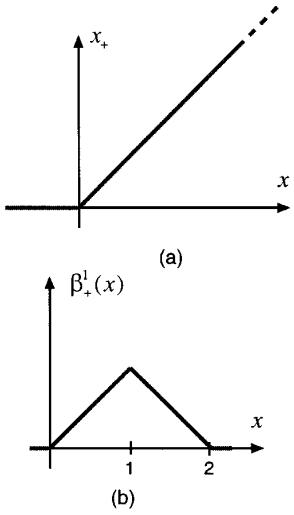

Fig. 1. Example of linear splines. (a) One-sided ramp x_+, which exhibits a discontinuity at the origin. (b) B-spline of degree 1. The function in (b) is localized, whereas the one in (a) is not.

where we have set $\rho_+(x) = x_+$.

By taking the second forward finite differences of $\rho_+(x)$—a specific instance of (4) with $a_0 = 1, a_1 = -2$, and $a_2 = 1$—we obtain the hat function (or causal B-spline of degree 1)

$$\beta_+^1(x) = \Delta_+^2 x_+ = x_+ - 2(x-1)_+ + (x-2)_+ \tag{6}$$

which is the more standard, compactly supported basis function for the linear splines [cf. Fig. 1(b)]. The remarkable fact with (6) is that the one-sided functions cancel out for $x \geq 2$; in other words, we are able to *localize* ρ_+ by taking a suitable linear combination of its shifted versions. Moreover, we can invert (6) and express the one-sided power function as a weighted sum of B-spline basis functions

$$x_+ = \sum_{k\geq 0} (k+1)\beta_+^1(x-k). \tag{7}$$

This shows that our definition (5) of the basic spline space is equivalent to the standard one that involves linear combinations of B-splines. Thus, $\{\rho_+(x-k)\}_{k\in\mathbb{Z}}$ is a valid basis for $\mathcal{V}_0$, albeit not a Riesz basis; the main difficulty is that the a_k in (4) are not necessarily in ℓ^2, which also means that this basis is not well conditioned. Thus, we will need to exert special care while manipulating expansions such as (4) both numerically and mathematically.

B. How Multiresolution Becomes Trivial

The advantage of the present formulation is that it makes the multiresolution structure of splines stand out quite naturally (cf. Fig. 2). Consider the coarse-to-fine sequence of subspaces $\cdots \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \cdots \subset \mathcal{V}_i \cdots$, where $\mathcal{V}_i$ represents the space of linear splines with knots at $x_k = k2^{-i}$, $k \in \mathbb{Z}$. These splines are generated simply by dropping all the basis functions in (5) that are not positioned at the desired knots. Thus, we define our uniform spline space with scale $T_i = 2^{-i}$ as

$$\mathcal{V}_i = \underset{k\in\mathbb{Z}}{\text{span}}\{\rho_+(x-k2^{-i})\}. \tag{8}$$

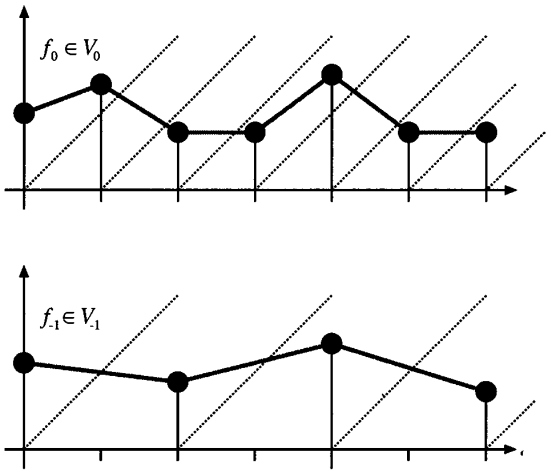

Fig. 2. Multiresolution spaces using one-sided power functions.

Clearly, the basis functions for $\mathcal{V}_i$ are a subset of those of $\mathcal{V}_{i'}$ for $i' > i$, which implies that $\mathcal{V}_i \subset \mathcal{V}_{i'}$, for all $i' > i$, which is a multiresolution property. Since each $\mathcal{V}_i$ also has a B-spline Riesz basis $\{2^{i/2}\beta_+^1(2^i x - k)\}_{k\in\mathbb{Z}}$, the whole ladder of spline subspaces for $i \in \mathbb{Z}$ generates a multiresolution of $\mathbf{L}^2(\mathbb{R})$ in the sense specified by Mallat [24]. Hence, it is possible to construct a whole variety of corresponding wavelet bases using any of the standard design techniques. Specific examples of linear spline wavelets that are orthogonal [24], semi-orthogonal [25], biorthogonal [26], or even shift-orthogonal [27] have been described in the literature.

C. Nonuniform Linear Splines

The power of the present formulation really becomes apparent if we move one step further and consider a given nonuniform sequence of knots $\cdots < x_k < x_{k+1} < \cdots$ with $k \in \mathbb{Z}$. We then define a corresponding embedded sequence of nonuniform spline spaces

$$\mathcal{V}_{-i} = \operatorname*{span}_{k\in\mathbb{Z}}\{\rho_+(x - x_{k2^i})\} \tag{9}$$

which share the same inclusion properties as before: $\cdots \subset \mathcal{V}_{-i} \subset \ldots \mathcal{V}_{-1} \subset \mathcal{V}_0$. Here, too, we are able to produce compactly supported basis functions (nonuniform B-splines), except that they lose the convenient shift-invariant structure that is inherent to standard (uniform) multiresolution analysis. Specifically, the kth (nonuniform) B-spline at resolution $-i$ is

$$\left(\frac{(x - x_{k2^i})_+ - (x - x_{(k+1)2^i})_+}{x_{(k+1)2^i} - x_{k2^i}}\right) - \left(\frac{(x - x_{(k+1)2^i})_+ - (x - x_{(k+2)2^i})_+}{x_{(k+2)2^i} - x_{(k+1)2^i}}\right). \tag{10}$$

It is a triangular function that takes the value one at $x = x_{(k+1)2^i}$ and vanishes for $x \le x_{k2^i}$ and $x \ge x_{(k+2)2^i}$. Note that this localization process involves divided differences of $\rho_+(x)$ rather than finite differences, as in (6). This nonuniform setting is also suitable for constructing wavelet bases that span the orthogonal complement of $\mathcal{V}_i$ in $\mathcal{V}_{i+1}$. Such nonuniform spline wavelets were first described by Lyche as early as in 1992 [22].

D. Central Basis Functions

Until now, we have worked with the one-sided function $\rho_+(x) = x_+$. We could have also used its reversed version $\rho_-(x) = \rho_+(-x)$, which can be localized using backward finite (or divided) differences. Another option would have been to use the symmetric function $\rho_*(x) = |x|$, which is the prototypical example of a radial basis function. In the sequel, we will consistently use the subscripts $*$, $+$, and $-$ to denote symmetric (or radial), causal, and anti-causal generating functions, respectively. These are all special cases of what we call *central basis functions* and which we denote by the generic symbol $\rho(x)$. The one-sided functions are typically simpler to work with in the univariate case. The symmetric ones, on the other hand, are usually preferred in a multivariate setting, for they fall within the framework the well-developed theory of radial basis functions.

E. Extension to Higher Dimensions: Variational Splines

Let us now briefly show how to extend this type of construction to a nonuniform grid in p dimensions. Here, it becomes preferable to work with the radial basis function $\rho_*(x) = |x|$. We also need to specify our nonuniform grid at the finest scale available. It is given in the form of an ordered set of (distinct) points $\mathbf{x}_k \in \mathbb{R}^p$, with $k \in \mathbb{Z}$. These may, for example, correspond to the discrete locations at which the samples of an input function are given. We then specify our basic radial spline space

$$\mathcal{V}_0 = \operatorname*{span}_{k\in\mathbb{Z}}\{\rho_*(\|\mathbf{x} - \mathbf{x}_k\|)\}. \tag{11}$$

For $p > 1$, the functions in $\mathcal{V}_0$ are no longer piecewise linear. Yet, they are compatible with the variational formulation of linear splines. The problem is the following: Given the samples of a function $f(\mathbf{x}_k) = f_k$, $k \in \mathbb{Z}$, one wishes to find the best interpolant in $\mathbf{L}^2(\mathbb{R}^p)$ such that an adequate Duchon semi-norm is minimized [2]. It can be shown that the optimum solution is necessarily included in $\mathcal{V}_0$; specifically, it is the unique function $s(\mathbf{x}) \in \mathcal{V}_0$ that satisfies the interpolation constraint $s(\mathbf{x}_k) = f_k$, $k \in \mathbb{Z}$. The functional $\int \|\nabla f(\mathbf{x})\|^2 \, d\mathbf{x}$ may be thought of as the bending energy of a thin membrane; hence, we have the term *membrane spline*. Finding the expansion coefficients for the basis $\rho_*(\|\mathbf{x} - \mathbf{x}_k\|)$ is precisely what the radial basis functions interpolation problem is all about.

F. Road Map to the Paper

The manipulations that we have made so far using linear splines are also valid for higher order splines. This is all well known in approximation theory. In the remainder of the paper, we will generalize these ideas further and show that the first part of the process described above (Sections II-A and II-B, in particular) is applicable in a much wider setting as had been realized so far. First, we will identify the whole class of central basis functions $\rho(x)$ (one-sided or radial) that can be used to specify a multiresolution analysis of $\mathbf{L}^2(\mathbb{R})$ in the sense defined by Mallat. We will see that an admissible central basis function must satisfy a self-similarity relation. This, together with

a suitable localization process [similar to (6)], will allow us to specify new classes of scaling functions much more explicitly than is done usually by means of an infinite recursion (two-scale relation). Second, we will look at the converse problem and prove that a central basis function $\rho(x)$ can be found for any compactly supported—and more generally, one-sided—scaling function $\varphi(x)$. The next logical step, which will not be pursued here, is to use these results to extend the standard wavelet constructions to the nonuniform case similar to what we have done with the linear splines in the example above.

III. From Central Basis Functions to Wavelets

We will now show how it is possible to construct wavelets starting from a radial (or central) basis function $\rho(x)$. We will not consider wavelets literally but, rather, their associated scaling functions, which are the key to the multiresolution structure of the wavelet transform. Once the scaling function has been specified, it is comparatively easy to construct a corresponding wavelet basis using standard techniques [12], [13].

A. Notation

We will need to raise complex numbers to some complex power. Euler showed that this is a well-defined operation, as long as one specifies the argument in a unique, unambiguous fashion, e.g., $z = |z|e^{j\arg(z)}$ with $\arg(z) \in [-\pi, \pi[$. Then, we specify

$$z^{z'} = e^{z'(\log|z|+j\arg z)}.$$

For instance, we have $|z^{z'}| = |z|^{\mathcal{R}(z')}e^{-\mathcal{J}(z')\arg z}$.

Euler's Gamma function is defined for $x > 0$ by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt.$$

This is a C^∞ interpolation of the factorial since $n! = \Gamma(n+1)$ if $n \in \mathbb{N}$. It is further possible to extend this function by analytic continuation to complex values of x, except at the negative integers, where $\Gamma(x)$ is unbounded (see [28]).

We denote by $\hat{f}(\omega) = \int_{\mathbb{R}} f(x)e^{-j\omega x}\,dx$ the Fourier transform of the absolutely integrable function $f(x)$. For functions such as $\rho(x)$ that are not bounded at ∞ but do not increase faster than a polynomial [i.e., $\rho(x) \in \mathcal{S}'$ where $\mathcal{S}'$ is Schwartz' class of tempered distributions], we will use a weaker definition. Specifically, $\hat{\rho}(\omega)$ is the Fourier transform of $\rho(x)$ in the sense of distributions if and only if $\langle\hat{\rho}, \psi\rangle = \langle\rho, \hat{\psi}\rangle$ for any indefinitely differentiable fastly decreasing function $\psi \in \mathcal{S}$ ("test function") [29].

B. Scaling Functions

Often, a scaling function is defined indirectly through its refinement filter G [cf. (12) below]. One then has to worry about the delicate issues of the convergence of the iterated filterbank and the $\mathbf{L}^2(\mathbb{R})$-completeness of the wavelet expansion. Here, we propose a more explicit definition that avoids these problems at the outset.

Definition 1: $\varphi(x)$ is a valid scaling function if and only if it satifies the following three conditions.

i) *Two-Scale Relation*: The function φ at a coarser resolution can be expressed using its shifts

$$\varphi(x/2) = \sum_{k\in\mathbb{Z}} g_k\varphi(x-k). \tag{12}$$

ii) $\mathbf{L}^2(\mathbb{R})$ *Stability*: $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ is a Riesz basis of $\mathcal{V}_0$, which is equivalent to require that there exist two positive constants A and B such that

$$A \le \sum_{k\in\mathbb{Z}} |\hat{\varphi}(\omega+2\pi k)|^2 \le B \tag{13}$$

almost everywhere.

iii) *Partition of Unity*:

$$\sum_{k\in\mathbb{Z}} \varphi(x-k) = 1 \Leftrightarrow \hat{\varphi}(2\pi k) = \delta_k. \tag{14}$$

The two first conditions are implicit to Mallat's axiomatic definition of a multiresolution analysis of $\mathbf{L}^2(\mathbb{R})$ [24]. Moreover, it can be shown that Mallat's completeness requirement [denseness in $\mathbf{L}^2(\mathbb{R})$] is equivalent the partition of unity condition, which has the merit of being explicit (cf. [30]). It follows that the three conditions are sufficient to build a multiresolution analysis of $\mathbf{L}^2(\mathbb{R})$.

C. Admissible Central and Radial Basis Functions

A central (either one-sided or radial) basis function $\rho(x)$ has properties that are quite distinct from those of a scaling function φ because of fundamental differences in the replication mechanism. With radial basis functions, the multiresolution bases (e.g., $\{\rho(x-2^ik)\}_{k\in\mathbb{Z}}$) are generated by simple translation of $\rho(x)$, as opposed to dilation plus translation (e.g., $\{\varphi(x/2^i - k)\}_{k\in\mathbb{Z}}$), as is traditionally the case with the wavelet transform. The most evident difference between the two species of functions is that the central basis functions are not well localized. To avoid gaps, they must be infinitely supported. In general, they are not even integrable, and yet, it can be shown that this apparent disadvantage brings approximation order (i.e., the ability to reproduce polynomials) to the generated spaces.

Definition 2: A central basis function $\rho(x) \in \mathcal{S}'$ is admissible if and only if the series of subspaces $\mathcal{V}_i = \text{span}_{k\in\mathbb{Z}}\{\rho(x-k2^{-i})\}$ for $i \in \mathbb{Z}$ are such that we have the following.

i) $\mathcal{V}_i$ is localizable, i.e., $\mathcal{V}_i \cap \mathbf{L}^2(\mathbb{R}) \neq \{0\}$.

ii) There exists $\varphi_i \in \mathcal{V}_i \cap \mathbf{L}^2(\mathbb{R})$ such that
   — $\mathbf{L}^2(\mathbb{R})$ stability: $\{\varphi_i(x-k2^{-i})\}_{k\in\mathbb{Z}}$ is a Riesz basis of $\mathcal{V}_i \cap \mathbf{L}^2(\mathbb{R})$
   — Partition of unity: $\sum_k \varphi_i(x-k2^{-i}) = 1$.

The existence of $\varphi_i \in \mathcal{V}_i \cap \mathbf{L}^2(\mathbb{R})$ means that $\rho(x)$ can be localized by taking a suitable linear combination of its shifts. The implication of the partition of unity is more intriguing. This means that it is possible to represent the constant as a linear combination of $\rho(x-kT_i)$, irrespective of the scale $T_i = 2^{-i}$. This nonunique way of writing the constant implies that $\rho(x)$ itself cannot generate a Riesz basis, which points out another fundamental difference between $\rho(x)$ and $\varphi(x)$.

By construction, the spaces $\mathcal{V}_i$ are embedded

$$\cdots \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \cdots \mathcal{V}_i \subset \mathcal{V}_{i+1} \subset \cdots$$

but this does not mean that the φ_i in Definition 2 are linked to each other through a two-scale relation.

Some good examples of admissible central basis functions are the one-sided power functions $x_+^n/n!$ with n integer. These can be localized using the backward $(n+1)$-order difference operator to yield the causal B-splines of degree n (cf. [31], [32]):

$$\beta_+^n(x) = \Delta_+^{n+1} \frac{x_+^n}{n!} \quad \stackrel{\text{Fourier}}{\longleftrightarrow} \quad \left(\frac{1-e^{-j\omega}}{j\omega}\right)^{n+1}. \tag{15}$$

The B-splines satisfy the three conditions in Definition 1; they constitute one of the earliest example of valid scaling functions [24].

Localizing less simple central basis functions turns out to be more intricate. It is theoretically possible in the good cases (see Buhmann in [33]) to localize them by building the corresponding interpolator, although it might not be an easy task to get the coefficients of the interpolating filter, especially when the Fourier transform of ρ is not a true function. Moreover, the interpolator is not always stable; see the classical example of the quadratic splines for which $\rho_+(x) = x_+^2$. We will give our alternative localization technique in Section III-F.

D. Self-Similar Central Basis Functions

We will say that a function ρ is *self-similar* (for dyadic zooming factors) iff it satisfies

$$\rho(x) = \lambda \rho\left(\frac{x}{2}\right) \tag{16}$$

for all $x \in \mathbb{R}$.

Theorem 1: Every self-similar admissible central basis function generates a valid multiresolution of $\mathbf{L}^2(\mathbb{R})$.

In Section IV, we will see that the converse is also true, provided that the scaling function φ is one-sided or compactly supported.

Proof: Let ρ be an admissible central basis function satisfying (16). Then, there exists $\varphi \in \mathcal{V}_0$ which is a Riesz basis for $\mathcal{V}_0 \cap \mathbf{L}^2(\mathbb{R})$. Since φ is in the span of $\rho(x-k)$, $\varphi(x/2)$ is in the span of $\rho(x/2-k)$, that is, in the span of $\rho(x-2n)$ because of the self-similarity of ρ. This also implies that $\varphi(x/2)$ is in the span of $\rho(x-n)$. Thus, $\varphi(x/2) \in \mathcal{V}_0 \cap \mathbf{L}^2(\mathbb{R})$. Since φ is a basis of $\mathcal{V}_0 \cap \mathbf{L}^2(\mathbb{R})$, $\varphi(x/2)$ is itself in the span of $\varphi(x-k)$, which proves that φ satisfies the scale relation (12). All the other required properties are provided by the definition 2 of an admissible central basis function. ■

Because of the self-similarity relation, the graph of such central basis functions are fractals.

Remarks:

- Equation (16) implies that $\rho(x)$ is either 0 or tends to infinity at one end ($x = 0$ or $x \to \infty$). If $|\lambda| > 1$, then $\rho(x)$ tends to grow globally from 0 to $+\infty$; otherwise, when $|\lambda| < 1$, it globally tends to decrease from $+\infty$ to 0.
- In order for ρ to be square integrable in $[0, 1]$, it is necessary that $|\lambda| > 1/\sqrt{2}$. This is because

$$\int_0^1 \rho(x)^2\,dx = \frac{1}{\lambda^2}\int_0^1 \rho(2x)^2\,dx = \frac{1}{2\lambda^2}\int_0^2 \rho(x)^2\,dx > \frac{1}{2\lambda^2}\int_0^1 \rho(x)^2\,dx$$

- A self-similar central basis functions is entirely and uniquely specified by its definition within the interval $[1/2, 1]$ and its eigenvalue λ. More precisely, if we denote by $\rho_{[1/2,\,1]}$ the restriction of ρ to $[1/2, 1]$, we have

$$\rho(x) = \sum_{i\in\mathbb{Z}} \lambda^i \rho_{[1/2,\,1]}\left(\frac{x}{2^i}\right).$$

  Conversely, this expression defines a self-similar central basis function with eigenvalue λ.
- The product and the convolution of two self-similar functions are self-similar as well. Admissibility, however, does not carry over so easily, e.g., the product of the admissible central basis functions $|x| \times |x|$ yields x^2, which is not admissible.

E. General Parametrization

A nice feature of one-sided self-similar central basis functions is that they can be expressed as a (countable) sum of one-sided power functions.

Theorem 2: Consider a self-similar one-sided central basis function ρ_+ with eigenvalue λ. Then, ρ is square integrable in $[1, 2]$ if and only if it can be expressed as

$$\rho_+(x) = \sum_{n\in\mathbb{Z}} \gamma_n x_+^{\log_2 \lambda + j(2n\pi/\log 2)} \quad \text{a.e.} \tag{17}$$

where

$$\gamma_n = \frac{1}{\log 2}\int_1^2 \rho_+(\xi)\xi^{-\log_2\lambda - 1 - j(2n\pi/\log 2)}\,d\xi \tag{18}$$

is a square summable sequence.

Likewise, under the same hypothesis, a self-similar *symmetric* central basis function ρ_* with eigenvalue λ can be expressed as a sum of symmetric monomials

$$\rho_*(x) = \sum_{n\in\mathbb{Z}} \gamma_n |x|^{\log_2\lambda + j(2n\pi/\log 2)} \quad \text{a.e.} \tag{19}$$

where $\{\gamma_n\}_{n\in\mathbb{Z}} \in \ell^2$.

For completeness, we have derived in Appendix A the Fourier transform of (17), which is defined in the sense of distributions.

Proof: As we have seen, a one-sided dyadic central basis function is completely defined by its restriction to the interval $[1, 2]$.

We define $\rho_0(x) = x^{-(\log\lambda/\log 2)}\rho_+(x)$ and observe that $\rho_0(2^x)$ is a 1-periodic function, i.e., $\rho_0(2^{x+1}) = \rho_0(2^x)$. Thus, we can apply Fourier's theorem to $\rho_0(2^x)$. If $\rho_0(x)$ is square integrable in $[1, 2]$ (which is the case when $\rho(x)$ is square integrable in $[1, 2]$), so is $\rho_0(2^x)$, and then, $\rho_0(2^x) = \sum_{n\in\mathbb{Z}} \gamma_n e^{j2n\pi x}$ almost everywhere, with $\gamma_n = \int_0^1 \rho_0(2^\xi)e^{-j2n\pi\xi}\,d\xi$. A change of variables then yields the expansion formula (17).

When the central basis function is symmetric, it can be expressed as $\rho_*(x) = \rho_+(x) + \rho_+(-x)$, where ρ_+ is the restriction of ρ to $[0, +\infty[$. Applying (17) to ρ_+ and noticing that $|x|^\alpha = x_+^\alpha + (-x)_+^\alpha$ immediately yields (19). ■

Of course, we may also consider more general self-similar central basis functions and express them as a sum of two self-similar one-sided central basis functions ρ_1 and ρ_2:

$\rho(x) = \rho_1(x) + \rho_2(-x)$, each of which can be decomposed according to (17).

F. Localization

Theorem 3: Let ρ_+ be a one-sided self-similar central basis function with eigenvalue λ, $|\lambda| > 1/\sqrt{2}$. If its representation in (17) is a finite sum, then the filter P defined through its z-transform $P(z) = \sum_{k\in\mathbb{Z}} p_k z^{-k}$ as

$$P(z)^{-1} = \int_0^\infty \rho_+(\xi) e^{-(1-z^{-1})\xi}\, d\xi = \sum_{n\geq 0} z^{-n} \int_0^\infty \rho_+(\xi) \frac{\xi^n}{n!} e^{-\xi}\, d\xi \qquad (20)$$

localizes the central basis function ρ_+, i.e., $\varphi(x) = \sum_{k\in\mathbb{Z}} p_k \rho_+(x-k)$ is in $\mathbf{L}^2(\mathbb{R})$.

Similarly, if ρ is a symmetric self-similar central basis function and if its representation in (19) is a finite sum, then the symmetric filter P defined through its z-transform as

$$P(z)^{-1} = \int_0^\infty \rho_*(\xi) \left(e^{-(1-z^{-1})\xi} + e^{-(1-z)\xi}\right) d\xi = \sum_{n\geq 0} (z^{-n} + z^n) \int_0^\infty \rho_*(\xi) \frac{\xi^n}{n!} e^{-\xi}\, d\xi \qquad (21)$$

localizes the symmetric central basis function ρ_*.

The proof of this result is given in Appendix B. The nonlocal character of $\rho(x)$ is entirely due to the $O(1/\omega^{1+\log_2\lambda})$ singularity of its Fourier transform at the origin (cf. Appendix A). Thus, the essence of the localization process is to cancel out these singularities. The intuitive explanation of Theorem 3 is as follows. Since $1 - e^{-j\omega} \approx j\omega$ as $\omega \to 0$, the integral on the first line of (20) for $z = e^{j\omega}$ and ω close to zero is essentially equivalent to the Fourier transform of $\rho(x)$. This means that the localization filter $P(e^{j\omega})$ will have multiple zeros at $\omega = 0$, which will precisely cancel all the singularities of $\hat{\rho}(\omega)$. An ideal localization [i.e., $\varphi(x) \to \delta(x)$] would be achieved if we could design a filter such that $\hat{\varphi}(\omega) = \hat{\rho}(\omega)P(e^{j\omega}) \approx 1$ over the whole frequency axis. While we can obviously not enforce this constraint everywhere because $P(e^{j\omega})$ is 2π-periodic and $\hat{\rho}(\omega)$ is not, we make sure that it is satisfied around the origin. In effect, our localization method is akin to using the bilinear transform to convert a continuous-time filter into a discrete one.

We conjecture that Theorem 3 still holds true when (17) is not finite, at least under wide conditions on $\rho_+(x)$. Our proof, however, can hardly be extended since, in that case, $\hat{\rho}_+(\omega)$ is no longer a function, which makes the product with $P(e^{j\omega})$ highly questionable.

Note that by construction, $\hat{\varphi}(\omega)$ is boundedly differentiable at 0 and has singularities of order $1 + \log_2\lambda$ at $\omega = 2n\pi$ for $n \in \mathbb{Z}\backslash\{0\}$. This implies that if $\lambda > 1$, $\varphi(x)$ decreases at least like $|x|^{-2}$ when $x \to \infty$ (an elaborate proof of a similar claim can be found in [34, App. B]).

The multiple zeros of $P(e^{j\omega})$ at the origin [or, equivalently, the singularities of $\hat{\rho}(\omega)$] have another desirable effect. They will endow the scaling function $\varphi(x)$ with a corresponding order of approximation $1 + \log_2\lambda$, which may even be noninteger [34]. This also means that the scaling function will reproduce all polynomials of degree $N \leq \lceil \log_2\lambda \rceil$ and that the corresponding wavelet will have $\lceil \log_2\lambda \rceil + 1$ vanishing moments [35].

Contrary to its inverse, the filter $P(z)$ in Theorem 3 is not necessarily causal; this means that the localized function might not be one-sided. However, when $P(z)$ has only a finite number of poles $\{\zeta_k\}_{k=1\ldots N}$ outside the unit circle, it is possible to choose $P_0(z) = P(z)\prod_{k=1}^N (1 - \zeta_k z^{-1})$ as an alternative localization filter; the advantage is that P_0 is causal and, thus, that $\varphi(x)$ is one sided.

G. Refinement Filter and Perfect Reconstruction Filterbank

Once the localization filter $P(e^{j\omega})$ has been specified, we may obtain the Fourier transform of refinement filter g in (12) by evaluating the ratio between $2\hat{\varphi}(2\omega)$ and $\hat{\varphi}(\omega)$. With the hypotheses of Theorem 3 [finite sum representation in (17)], this simplifies to

$$G(e^{j\omega}) = \frac{1}{\lambda} \frac{P(e^{j2\omega})}{P(e^{j\omega})}. \qquad (22)$$

A corresponding (highpass) wavelet filter $H(e^{j\omega})$ can then be determined by using any of the standard design techniques. For a complete specification of the wavelet transform algorithm, we also need the complementary analysis filters $\tilde{G}(e^{j\omega})$ and $\tilde{H}(e^{j\omega})$. These are determined by imposing the perfect reconstruction conditions in the Fourier domain.

Since the underlying filters are not necessarily compactly supported, the best way to implement this kind of wavelet transform is to perform the computation in the frequency domain. For this purpose, we may use the fast FFT-based implementation[2] of the wavelet transform described in [35] and [36].

H. Example 1: Fractional Splines

As a first example within our general class of self-similar functions, we consider the first term ($n = 0$) of (17): x_+^α with $\alpha = \log_2\lambda$ (cf. Fig. 3). The Fourier transform of this one-sided power function is given by the first term in (33). The application of Theorem 3 [cf. (35)] leads to the localization filter

$$P(e^{j\omega}) = \frac{(1 - e^{-j\omega})^{\alpha+1}}{\Gamma(\alpha+1)}$$

which is essentially a fractional iterate of the finite difference operator $\Delta \longleftrightarrow (1 - e^{-j\omega})$. Interestingly, this localization process yields the fractional B-splines of degree α

$$\beta_+^\alpha(x) = \Delta_+^{\alpha+1} \frac{x_+^\alpha}{\Gamma(\alpha+1)} \quad \overset{\text{Fourier}}{\longleftrightarrow} \quad \left(\frac{1 - e^{-j\omega}}{j\omega}\right)^{\alpha+1}$$

which is a recently proposed extension of the polynomial B-splines for noninteger degrees [34]. As can be seen in Fig. 4, these functions provide a smooth transition between the conventional B-splines. Each of them generates a valid Riesz basis and satisfies the partition of unity, provided that $\alpha > -(1/2)$ (cf. [34]) or, equivalently, $\lambda > 1/\sqrt{2}$, which is a necessary condition for Theorem 2.

[2]MATLAB and Java sources are available at http://bigwww.epfl.ch/blu/fract-splinewavelets/.

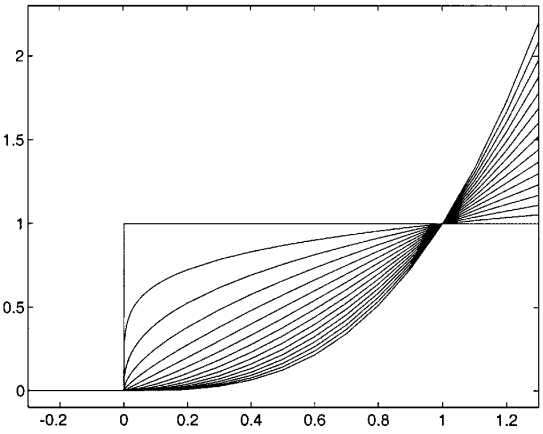

Fig. 3. Central basis functions $\rho_+(x) = x_+^\alpha$ of the fractional B-splines (see [37]) shown in Fig. 4.

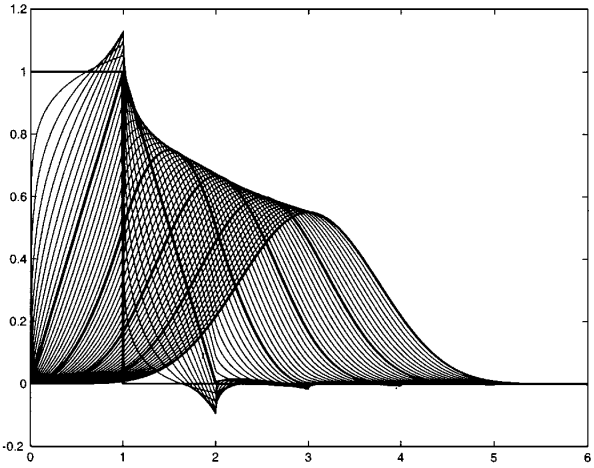

Fig. 4. Fractional B-splines $\beta_+^\alpha(x)$ with $\alpha \geq 0$. These functions interpolate the conventional B-splines that are represented using a thicker line. The noninteger degree splines are not compactly supported, although they decrease like $|x|^{-\alpha-2}$.

Next, we apply (22) to determine the corresponding refinement filter

$$G(e^{j\omega}) = \frac{1}{\lambda}\frac{(1-e^{-j2\omega})^{\alpha+1}}{(1-e^{-j\omega})^{\alpha+1}} = \frac{(1+e^{-j\omega})^{\alpha+1}}{2^\alpha} \tag{23}$$

which happens to be the generalized version of the celebrated binomial filter. This clearly shows that the fractional B-splines are valid scaling functions. They have led to interesting generalizations of the traditional spline wavelets—including the Battle–Lemarie orthogonal splines [35], [36]. The remarkable feature here is that these wavelets are indexed by a continuously-varying order parameter $(\alpha+1)$, as opposed to an integer, as is usually the case in wavelet theory.

I. Example 2: Harmonic Splines

The second type of elementary component in (17) is of the form

$$x_+^{\log_2 \lambda + j(2n\pi/\log 2)} = x_+^{\log_2 \lambda}(\cos(2\pi n \log_2 x_+) + j\sin(2\pi n \log_2 x_+)).$$

These are essentially modulated versions of the power functions with the period of the harmonic component increasing logarithmically. We can combine two complex conjuguate components with index $-n$ and n to obtain a real-valued central basis function. These are the generating functions of what we define as *harmonic splines*.

Fig. 5 illustrates the result of the localization of the function

$$\rho_+(x) = \frac{x_+^{3-j(2\pi/\log 2)}}{\Gamma\left(4-j\frac{2\pi}{\log 2}\right)} + \frac{x_+^{3+j(2\pi/\log 2)}}{\Gamma\left(4+j\frac{2\pi}{\log 2}\right)}.$$

The Fourier transform of the corresponding localization filter, as given by Theorem 3, is

$$P(e^{j\omega}) = \left(\frac{1}{(1-e^{-j\omega})^{4-j(2\pi/\log 2)}} + \frac{1}{(1-e^{-j\omega})^{4+j(2\pi/\log 2)}}\right)^{-1}.$$

Note that by construction, $\rho_+(x) = 8\rho_+(x/2)$. Obviously, even if its integral does not vanish [in fact, $\int \varphi(x)\,dx = 1$], this scaling function has a bandpass character that is rather unusual for a scaling function. It can also yield wavelet bases. In fact, we used the reconstruction part of a wavelet transform algorithm to graph the function in Fig. 5.

IV. From Wavelets to Central Basis Functions

We will now show that one can also follow the reverse path and uncover the central basis function(s) that lies hidden behind any scaling function $\varphi(x)$. To this end, we need to assume that φ is one sided. This includes all compactly supported scaling functions but excludes those that are supported on the whole real line.

A. Finding the Central Basis Function

Theorem 4: Let $\varphi(x)$ be an admissible scaling function with corresponding causal refinement filter $G(z) = \sum_{k\geq 0} g_k z^{-k}$ with $g_0 \neq 0$. Then, there exists a unique self-similar one-sided (i.e., supported in $[0, +\infty[$) central basis function $\rho_+(x)$ that generates the same multiresolution analysis. The self-similarity parameter λ defined by (16) is simply g_0^{-1}.

This central basis function $\rho_+(x)$ can be expressed as

$$\rho_+(x) = \sum_{n\geq 0} r_n \varphi(x-n) \tag{24}$$

where the coefficient sequence r_n satisfies

$$r_0 = 1$$
$$r_n = \frac{1}{g_0}\sum_{k\geq 0} g_{n-2k} r_k, \qquad \forall n > 0. \tag{25}$$

Proof: Because G is a causal filter, φ is one-sided, i.e., its support is contained in $[0, +\infty[$. We now dilate the function

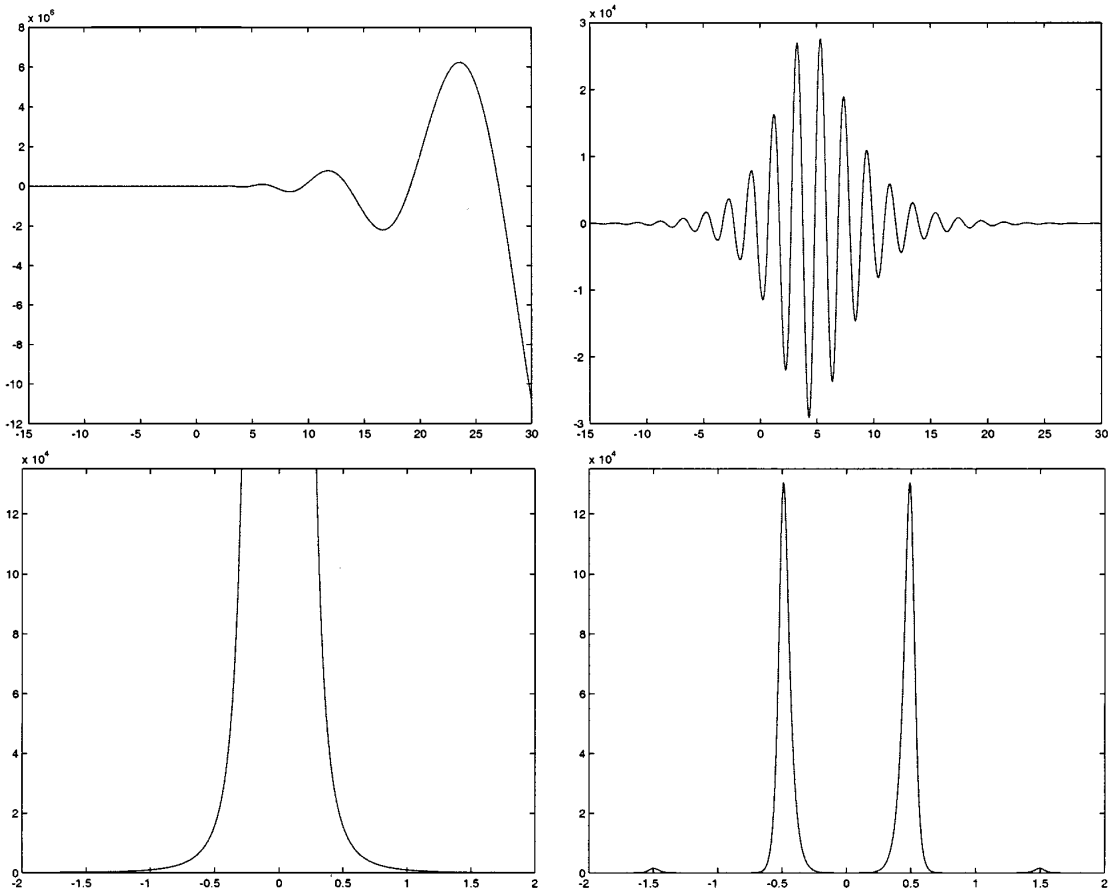

Fig. 5. Top: Central basis function $\rho_+(x)$ and its localized version $\varphi(x)$ (harmonic spline). Bottom: Fourier transforms $\hat{\rho}_+(\omega)$ and $\hat{\varphi}(\omega)$. See Section III-I for more details.

ρ, as defined by (24), and apply the two-scale equation (12) to show that it is self-similar.

$$\begin{aligned}\rho(x/2) &= \sum_n r_n \varphi(x/2 - n) \\ &= \sum_{n,k} g_{k-2n} r_n \varphi(x - k) \quad \text{[using (12)]} \\ &= \sum_k g_0 r_k \varphi(x - k) \quad \text{[using (25)]} \\ &= g_0 \rho(x).\end{aligned}$$

Note that we need not have technical convergence conditions for interverting sum signs here; this is because the summation in (24) is finite for fixed x. Thus, we have span$\{\rho(x-n)\} \subset$ span$\{\varphi(x-n)\}$.

Conversely, as shown later, (24) can be inverted exactly to yield $\varphi(x)$ as a linear combination of shifted versions of ρ_+. This shows that span$\{\varphi(x-n)\} \subset$ span$\{\rho_+(x-n)\}$. In fact, we even have $\varphi(x) = \rho_+(x)$ for $x \in [0, 1[$, which shows that ρ_+ is unique since a self-similar function is uniquely defined by its value in $[1/2, 1[$. ■

Corollary 1: Let $\varphi(x)$ be a symmetric compactly supported scaling function with corresponding refinement filter $G(z)$. Then, there exist a radial basis function $\rho_*(x)$ that generates the same multiresolution analysis

$$\rho_*(x) = \rho_+(x) + \rho_-(x) \tag{26}$$

where $\rho_+(x)$ is given by (24).

B. One-Sided Localization Filter

In the present case, φ is given *a priori*, which means that the localization filter p is specified implicitly. To get its explicit form, we consider the matrix form of (24)

$$\begin{bmatrix} \rho_+(x) \\ \rho_+(x-1) \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} r_0 & r_1 & \cdots & \\ 0 & r_0 & r_1 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & & \end{bmatrix}}_{\mathbf{R}} \begin{bmatrix} \varphi(x) \\ \varphi(x-1) \\ \vdots \end{bmatrix} \tag{27}$$

where the infinite matrix $\mathbf{R}$ is Toeplitz upper triangular. To determine p such that $\varphi(x) = \sum_{n \geq 0} p_n \rho(x-n)$, we simply need to evaluate the inverse of $\mathbf{R}$, which turns out to be Toeplitz upper

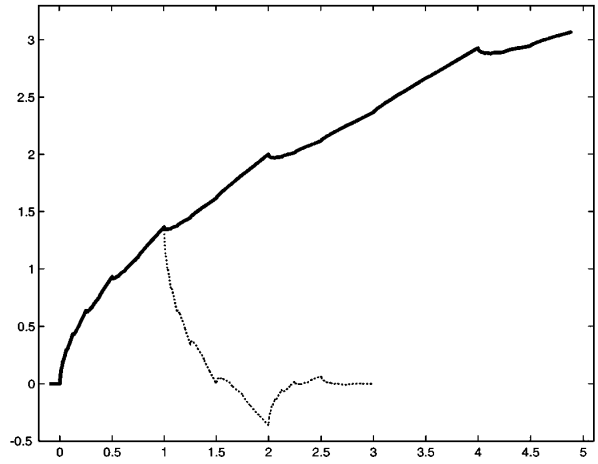

Fig. 6. Scaling function φ (dotted line) and self-similar central basis function ρ_+ (solid line) corresponding to Daubechies minimum phase wavelets of order 2. Note that ρ_+ and φ coincide over $[0,\ 1]$.

triangular as well. By solving this triangular system of equations, we find that

$$p_k = \frac{-1}{r_0} \sum_{n=1}^{k} r_n p_{k-n}. \tag{28}$$

Note that this localization is not necessarily the same as the one described in Theorem 3, which may yield another scaling function in $\mathcal{V}_0$.

C. Example 1: Daubechies

Daubechies scaling functions [38] are compactly supported and can thus be expressed using a unique self-similar one-sided central basis function. To illustrate our theory, we have evaluated the sequence r_n using (25) for a Daubechies filter of length 4, which corresponds to a wavelet of order 2. The corresponding functions are plotted in Fig. 6. Note that ρ_+, which is clearly fractal, also corresponds to the self-similar extrapolation of $\varphi_{[1/2,\,1]}$, i.e., the restriction of $\varphi(x)$ to $x \in [1/2,\ 1]$, with $\lambda = 4/(1+\sqrt{3}) \approx 1.46$.

D. Example 2: Polynomial Splines

The same kind of exercise can be performed for the causal B-splines (15) to recover the one-sided power functions x_+^n. Here, instead of using (25), we will determine the sequence r as the inverse of the localization operator p to illustrate the link in Section II-A.

The z-transform of the localization operator for the B-spline of degree n is

$$P^n(z) = \frac{1}{n!}\,(1 - z^{-1})^{n+1}. \tag{29}$$

Formally, $R^n(z)$ is obtained by taking the convolution inverse of $P^n(z)$. Using the Taylor development of $(1-z^{-1})^{-n-1}$ (which is given, e.g., in [28]), this yields

$$R^n(z) = \frac{1}{P^n(z)} = n! \sum_{k\geq 0} \binom{n+k}{n} z^{-k}. \tag{30}$$

One can check that the results for $n = 1$ are compatible with (7); therefore, we have gone full circle.

E. Time-Domain Parametrization of Scaling Functions

Theorem 4 allows us to state the utmost result of this paper, i.e., a full parametrization of one-sided dyadic scaling functions.

Corollary 2: Let φ be an admissible one-sided or compactly supported scaling function. Then, there exist a real number $\alpha > -0.5$, a complex sequence $\{\gamma_n\}_{n\in\mathbb{Z}} \in \ell^2$, and a real sequence $\{p_n\}_{n\in\mathbb{N}}$ such that

$$\varphi(x) = \sum_{k\in\mathbb{N},\, l\in\mathbb{Z}} p_k \gamma_l (x-k)_+^{\alpha + j(2l\pi/\log 2)}. \tag{31}$$

Conversely, any function described by (31) satisfies a two-scale difference equation but is not necessarily compactly supported.

As far as we know, this is the first time that an explicit time domain formula is given for scaling functions. Equation (31) is interesting because it explains the fractal nature of scaling functions and wavelets. The expansion can also be interpreted as an infinite sum of harmonic splines. This points to a new fundamental aspect of splines as elementary constituents of scaling functions.

Note that the exponent α in (31) has a very special role. Whenever the sum over γ_l is finite, it is tightly linked to four important wavelet properties:

i) the Hölder regularity, which is α;
ii) the Solobev regularity, which is $\alpha + (1/2)$;
iii) the order of approximation, which is $\alpha + 1$;
iv) the polynomial degree reconstruction, which is $\lceil \alpha \rceil$.

We remark that except for the case of splines, there is a whole range of such wavelets that has not yet been explored.

We should mention, however, that the meaning of $\alpha = -\log_2 g_0$ is not as transparent for the traditional wavelets such as Daubechies whose harmonic decomposition involves an infinite number of terms. This leaves us with the following interrogation: The Hölder regularity of the minimum phase Daubechies wavelets turns out to be $\alpha = -\log_2 g_0$ (cf. [39]–[41]). Is this fact coincidental or not?

F. Borderline Cases

We have established a full equivalence between self-similar central functions and scaling functions only for the one-sided case. It is likely that a similar equivalence holds for most scaling functions that are not one-sided as well, but we also have evidence that this will not always be the case.

An interesting counterexample is provided by the symmetric fractional splines whose generating function is $\rho_*(x) = x^{2n}\log|x|$, where n is some integer. We have shown recently that this classical radial basis function could be localized to yield the symmetric fractional B-spline β_*^{2n}, which is a perfectly valid scaling function [34]. Yet, the corresponding $\rho_*(x)$ is not strictly self-similar; it is only self-similar up to a polynomial term: $\rho_*(x/2) = (1/4^n)(\rho_*(x) - x^{2n}\log 2)$. However, the span of these radial basis functions is still a multiresolution space because x^{2n} belongs to it.

V. Conclusion

We have presented new results that make the connection between radial basis functions and wavelets very explicit.

The fact that we can move from radial basis functions to wavelets enables us to control some of their key mathematical properties: regularity and order of approximation. It also yields wavelets that have an explicit analytical form, the fractional splines being a notable example.

The existence of a link in the reverse direction—from wavelets to central basis functions—is especially interesting conceptually. It leads to an alternative interpretation of multiresolution. Basis functions are simply removed (resp. added) instead of being dilated (resp. contracted), as is usually the case. This opens up the door to many possible extensions, such as wavelets on nonuniform grids.

This study has also brought out the fundamental role of splines and their harmonic extensions that are the elementary constituents of wavelets. The decomposition in terms of harmonic splines leads to an explicit formula for scaling functions and wavelets in the continuous-time domain. It also provides a clear understanding of their fractal properties.

Appendix A
Fourier Transform of a One-Sided Self-Similar Central Basis Function

It turns out that the Fourier transform of the tempered distribution $u_z(x) = x_+^z$ is known in the general case where $z \in \mathbb{C}$. It is given by (see [42])

$$\hat{u}(\omega) = \frac{\Gamma(z+1)}{(j\omega)^{z+1}} \qquad \text{if } z \notin \mathbb{N}$$

$$\hat{u}(\omega) = \frac{n!}{(j\omega)^{n+1}} + \pi j^n \delta^{(n)}(\omega) \qquad \text{if } n \in \mathbb{N}.$$

Thus, in the sense of distributions, we have

$$\hat{\rho}(\omega) = \sum_{n\in\mathbb{Z}} \gamma_n \frac{\Gamma\left(\alpha+1+j\dfrac{2n\pi}{\log 2}\right)}{(j\omega)^{\alpha+1+j(2n\pi/\log 2}} \qquad \text{if } \alpha = \log_2 \lambda \notin \mathbb{N}$$

$$\hat{\rho}(\omega) = \sum_{n\in\mathbb{Z}} \gamma_n \frac{\Gamma\left(\alpha+1+j\dfrac{2n\pi}{\log 2}\right)}{(j\omega)^{\alpha+1+j(2n\pi/\log 2)}} + \gamma_0 \pi j^\alpha \delta^{(\alpha)}(\omega) \quad \text{if } \alpha = \log_2 \lambda \in \mathbb{N}. \tag{32}$$

Appendix B
Proof of Theorem 3

Proof: Here, we assume that ρ can be expressed as in (17) with a finite sum. From Appendix A, we also have its Fourier transform in the sense of distributions:

$$\hat{\rho}(\omega) = \sum_{\text{finite } n} \gamma_n \frac{\Gamma\left(1+\log_2\lambda + j\dfrac{2n\pi}{\log 2}\right)}{(j\omega)^{1+\log_2\lambda+j(2n\pi/\log 2)}} \tag{33}$$

if λ is not an integer power of two. In the case where $\lambda = 2^k$ for some $k \in \mathbb{N}$, $\gamma_0 j^k \pi \delta^{(k)}(\omega)$ has to be added to this expression.

We now use (20) to derive the filter contribution associated with a single component of the form x_+^α, making the complex change of variable[3] $u = (1 - z^{-1})x$

$$\begin{aligned}\int_0^\infty x^\alpha e^{-(1-z^{-1})x}\,dx &= \int_0^\infty \left(\frac{u}{1-z^{-1}}\right)^\alpha e^{-u}\frac{du}{1-z^{-1}} \\ &= \frac{1}{(1-z^{-1})^{\alpha+1}}\int_0^\infty u^\alpha e^{-u}\,du \\ &= \frac{\Gamma(1+\alpha)}{(1-z^{-1})^{\alpha+1}}. \end{aligned} \tag{34}$$

By summing over all components, we get the explicit form of the inverse of the localization filter

$$P(e^{j\omega})^{-1} = \sum_{\text{finite } n} \gamma_n \frac{\Gamma\left(1+\log_2\lambda + j\dfrac{2n\pi}{\log 2}\right)}{(1-e^{-j\omega})^{1+\log_2\lambda+j(2n\pi/\log 2)}}. \tag{35}$$

Because $1 - e^{-j\omega} = j\omega + O(\omega^2)$, it is clear that $P(e^{j\omega})^{-1}$ is equivalent to $\hat{\rho}(\omega)$ in the neighborhood of $\omega = 0$ [except for the potential $\delta^{(k)}(\omega)$ correction in the few cases where $\lambda = 2^k$]. Clearly, the localization filter $P(e^{j\omega})$ will be well defined and bounded provided that $P(e^{j\omega})^{-1} \neq 0$ over $[0, 2\pi[$. Under those conditions, we can state that $\hat{\varphi}(\omega) = P(e^{j\omega})\hat{\rho}(\omega)$ is bounded over $\mathbb{R}$ due to the fact that $\hat{\rho}(\omega)$ is bounded away from $\omega = 0$. The same argument holds for $\lambda = 2^k$ because then, $P(e^{j\omega}) = O(\omega^{k+1})$ and because $\omega^{k+1}\delta^{(k)}(\omega) = 0$. Now, we claim that $\hat{\varphi}$ is in $\mathbf{L}^2(\mathbb{R})$. This is because

$$\begin{aligned}\left|\frac{1}{(j\omega)^{1+\log_2\lambda+j(2n\pi/\log 2)}}\right| &= \frac{e^{n\pi^2/\log 2}\text{sign}(\omega)}{|\omega|^{1+\mathcal{R}(\log_2\lambda)}} \\ &\le \frac{\text{Constant}}{|\omega|^{1+\mathcal{R}(\log_2\lambda)}}\end{aligned}$$

which implies that $|\hat{\varphi}(\omega)|^2 \le \text{Constant} \times |\omega|^{-2-2\mathcal{R}(\log_2\lambda)}$ away from $\omega = 0$. Thus, $|\hat{\varphi}|^2$ is integrable over $]-\infty, -1] \cup [1, \infty[$, provided that $|\lambda| \ge 1/\sqrt{2}$. Since $|\hat{\varphi}|^2$ is also bounded, it is obviously integrable over $[-1, 1]$. By Plancherel identity, we conclude that $\varphi \in \mathbf{L}^2(\mathbb{R})$.

If ρ is symmetric, we decompose it as $\rho_+(x) + \rho_+(-x)$ and proceed similarly to show that $\hat{\varphi}(\omega) = P(e^{j\omega})\hat{\rho}(\omega)$ is bounded and square integrable, which proves that φ is square integrable. ■

References

[1] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," *Zeitschrift Angewandte Math. Phys.*, vol. 30, pp. 292–304, 1979.

[2] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller, Eds. Berlin, Germany: Springer-Verlag, 1977, pp. 85–100.

[3] F. L. Bookstein, *Morphometric Tools for Landmark Data*. New York: Cambridge Univ. Press, 1991.

[4] J. C. Carr, W. R. Fright, and R. K. Beatson, "Surface interpolation with radial basis functions for medical imaging," *IEEE Trans. Med. Imag.*, vol. 16, pp. 96–107, Jan. 1997.

[5] M. Gabrani and O. J. Tretiak, "Surface-based matching using elastic transformations," *Pattern Recognit.*, vol. 32, no. 1, pp. 87–97, Jan. 1999.

[3] This manipulation is allowed for $|z| < 1$ because $x^\alpha e^{-(1-z^{-1})x}$ is analytic in x in the cone bounded by the half lines $]0, +\infty[$ and $(1 - z^{-1})^{-1}]0, +\infty[$ and that it is fastly decreasing at its infinite boundary.

[6] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, vol. 78, pp. 1481–1497, Sept. 1990.
[7] P. Niyogi, F. Girosi, and T. Poggio, "Incorporating prior information in machine learning by creating virtual examples," *Proc. IEEE*, vol. 86, pp. 2196–2207, Nov. 1998.
[8] C. M. Bishop, "Training with noise is equivalent to Tikhonov regularization," *Neural Comput.*, vol. 7, pp. 108–116, 1995.
[9] F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures," *Neural Comput.*, vol. 7, pp. 219–269, 1995.
[10] T. K. Leen, "From data distributions to regularization in invariant learning," *Neural Comput.*, vol. 7, no. 5, pp. 974–981, Sept. 1995.
[11] C. A. Micchelli, "Interpolation of scattered data: distance matrices and conditionally positive definite functions," *Constr. Approx.*, vol. 2, pp. 11–22, 1986.
[12] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1998.
[13] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.
[14] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
[15] M. D. Buhmann, "Pre-wavelets on scattered knots and from radial function spaces: A review," in *The Mathematics of Surfaces VI*. New York: Oxford Univ. Press, 1996, pp. 309–324.
[16] ——, "Discrete least squares approximation and prewavelets from radial function spaces," *Math. Proc. Cambridge Philo. Soc.*, vol. 114, no. 3, pp. 533–558, 1993.
[17] F. I. Utreras, "Multiresolution and pre-wavelets using radial basis functions," in *Multivariate Approximation: From CAGD to Wavelets*, K. Jetter and F. I. Utreras, Eds. Singapore: World Scientific, 1993, pp. 321–333.
[18] J. Stöckler, *Non-Stationary Wavelets, Multivariate Approximation: From CAGD to Wavelets*. Singapore: World Scientific, 1993.
[19] C. K. Chui, J. Stöckler, and J. D. Ward, "Analytic wavelets generated by radial functions," *Adv. Comput. Math.*, vol. 5, pp. 95–123, 1996.
[20] M. D. Buhmann, "Multiquadric prewavelets on nonequally spaced knots in one dimension," *Math. Comput.*, vol. 64, no. 212, pp. 1611–1625, 1995.
[21] C. K. Chui, J. D. Ward, K. Jetter, and J. Stöckler, "Wavelets for analyzing scattered data: An unbounded operator approach," *Appl. Comput. Harmon. Anal.*, vol. 3, pp. 254–267, 1996.
[22] T. Lyche and K. Mørken, "Spline-wavelets of minimal support," in *Numerical Methods of Approximation Theory IX*, D. Braess and L. L. Shumaker, Eds. Basel, Switzerland: Birkhäuser, 1992, vol. 105, Int. Ser. Numer. Math., pp. 177–194.
[23] M. D. Buhmann and C. A. Micchelli, "Spline prewavelets for nonuniform knots," *Numer. Math.*, vol. 61, pp. 455–474, 1992.
[24] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet decomposition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.
[25] M. Unser, A. Aldroubi, and M. Eden, "A family of polynomial spline wavelet transforms," *Signal Process.*, vol. 30, pp. 141–162, 1993.
[26] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, no. 5, pp. 485–560, 1992.
[27] M. Unser, P. Thévenaz, and A. Aldroubi, "Shift-orthogonal wavelet bases using splines," *IEEE Signal Processing Lett.*, vol. 3, pp. 85–88, Mar. 1996.
[28] M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions*. Boulder, CO: Nat. Bur. Stand., 1972.
[29] L. Schwartz, *Théorie des Distributions* (in French). Paris, France: Hermann, 1966.
[30] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, pp. 569–587, Apr. 2000.
[31] I. J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, vol. 4, pp. 45–99 and 112–141, 1946.
[32] ——, *Cardinal Spline Interpolation*. Philadelphia, PA: SIAM, 1973.
[33] M. D. Buhmann, "Multivariate cardinal interpolation with radial-basis functions," *Constr. Approx.*, vol. 6, pp. 225–255, 1990.
[34] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Rev.*, vol. 42, no. 1, pp. 43–67, Jan. 2000.
[35] T. Blu and M. Unser, "Fractional wavelet bases," *IEEE Trans. Signal Processing*, 2001, submitted for publication.
[36] ——, "The fractional spline wavelet transform: Definition and implementation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. I, Istanbul, Turkey, June 2000, pp. 512–515.
[37] ——, "Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors," *IEEE Trans. Signal Processing*, vol. 47, pp. 2783–2795, Oct. 1999.
[38] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. XLI, pp. 909–996, Nov. 1988.
[39] I. Daubechies and J. Lagarias, "Two-scale difference equations I. Existence and global regularity of solutions," *SIAM J. Math. Anal.*, vol. 22, no. 5, pp. 1388–1410, Sept. 1991.
[40] ——, "Two-scale difference equations II. Local regularity, infinite products of matrices and fractals," *SIAM J. Math. Anal.*, vol. 23, no. 4, pp. 1031–1079, July 1992.
[41] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544–1576, Nov. 1992.
[42] D. C. Champeney, *A Handbook of Fourier Theorems*. Cambridge, U.K.: Cambridge Univ. Press, 1987.

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

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

Michael Unser (F'99) 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 (EPFL), Lausanne.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, where he headed the Image Processing Group. 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 over 90 published journal papers in these areas. He is on the editorial boards of *Signal Processing, the Journal of Visual Communication and Image Representation, Sampling Theory in Signal and Image Processing*, and *Pattern Recognition*.

Dr. Unser is an Associate Editor for the IEEE TRANSACTIONS ON MEDICAL IMAGING and is a Guest Editor for its special issue on Wavelets in Medical Imaging. He was a member of the Image and Multidimensional Signal Processing Committee of the IEEE Signal Processing Society from 1993 to 2000 and was former Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1992 to 1995 and of the IEEE SIGNAL PROCESSING LETTERS from 1994 to 1998. He co-organized the 1994 IEEE-EMBS Workshop on Wavelets in Medicine and Biology and serves as regular conference chair for SPIE's Wavelet Applications in Signal and Image Processing, which has been held annually since 1993. He is general co-chair for the IEEE International Symposium on Biomedical Imaging (ISBI'2002), which is a new conference to be held in Washington, DC, from July 7 to 10, 2002. He received the Dommer prize for excellence from EPFL in 1981, the research prize of the Brown-Boveri Corporation (Switzerland) for his thesis in 1984, the IEEE Signal Processing Society's 1995 Best Paper Award, and IEEE Signal Processing Society's 2000 Magazine Award. He is a member of IEEE, EURASIP, SPIE, and SIAM.