

**Chapter 3: Fractal and Multifractal Theory, Wavelet Algorithm, Wavelet
in Numerical Analysis**

Wavelet Leaders in Multifractal Analysis <i>S. Jaffard, B. Lashermes and P. Abry</i>	201
Application of Fast Wavelet Transformation in Parametric System Identification <i>K. Markwardt</i>	247
Image Denoising by a Novel Digital Curvelet Reconstruction Algorithm <i>J. Bai and X.-C. Feng</i>	255
Condition Number for Under-Determined Toeplitz Systems <i>H. Diao and Y. Wei</i>	263
Powell–Sabin Spline Prewavelets on the Hexagonal Lattice <i>J. Maes and A. Bultheel</i>	273

Wavelet Leaders in Multifractal Analysis

Stéphane Jaffard, Bruno Lashermes and Patrice Abry

Abstract. The properties of several multifractal formalisms based on wavelet coefficients are compared from both mathematical and numerical points of view. When it is based directly on wavelet coefficients, the multifractal formalism is shown to yield, at best, the increasing part of the weak scaling exponent spectrum. The formalism has to be based on new multiresolution quantities, the wavelet leaders, in order to yield the entire and correct spectrum of Hölder singularities. The properties of this new multifractal formalism and of the alternative weak scaling exponent multifractal formalism are investigated. Examples based on known synthetic multifractal processes are illustrating its numerical implementation and abilities.

1. Introduction

The purpose of multifractal analysis is to study functions or signals whose pointwise Hölder regularity may change widely from point to point. In such situations, the determination of the pointwise regularity at each point is numerically unstable; usually, it is quite meaningless since the exact regularity at a particular point usually does not carry a useful information. Therefore, one rather wishes to derive some information concerning the size of the sets of points where the pointwise regularity exponent takes a given value H . This “size” is mathematically formalized as the Hausdorff dimension. These dimensions define a function of the exponents H referred to as the spectrum of singularities (or multifractal spectrum) of f and denoted $d_f(H)$. Therefore, **performing the multifractal analysis of a function (or of a signal) f means to determine (or to estimate) its spectrum of singularities $d_f(H)$.** When working on real-life signals, the spectrum $d_f(H)$ cannot be computed by first determining the regularity exponent at each point, since it was precisely introduced as a substitute for this quantity; hence the necessity to introduce a method that yields this spectrum from numerically computable quantities derived from the signal. This is precisely the goal of “multifractal formalisms” and the purpose of the present work is to introduce a new multifractal formalism, based on new multiresolution quantities, the wavelet leaders.

However before introducing wavelet techniques, we will first examine how such formalisms were introduced in the setting of measures and functions; indeed a careful inspection of these more simple settings will be the key to a good understanding of what the alternative wavelet extensions yield.

In this paper, we first provide the reader with a description of the different multifractal formalisms which have been introduced in the setting of measures (Section 2) and in the setting of functions (Section 3). The first wavelet-based formulas are presented in Section 4; we will discuss the pertinence of these wavelet-based formulas and show that they lead to numerically unstable computations. The way to overcome these problems is to give up basing the multifractal formalism directly on wavelet coefficients but rather on wavelet leaders. This is developed in Section 5. Finally, In Section 6 we show that the previous multifractal formalism based on wavelet coefficients can be (partly) interpreted as yielding the spectrum of singularities based on another pointwise regularity exponent: The weak scaling exponent.

This paper is partly a review paper and partly a research paper. Its main novelty is twofold:

- We show that, both numerically **and** theoretically, a wavelet-based multifractal formalism yields more accurate results if it is built on wavelet leaders rather than directly on wavelet coefficients. The particular examples supplied by Brownian motion and fractional Brownian motions are investigated in details and the performances of the different multifractal formalisms are compared both theoretically and numerically on these examples.
- We show that a multifractal formalism based on wavelet coefficients can only be expected to yield the *weak scaling spectrum* (see Definition 19), and therefore it can yield the spectrum of singularities only in the particular cases where the two spectra coincide.

The numerical data shown in this paper only involve synthetic signals (Fractional Brownian Motions and multiplicative cascades) whose spectra are known exactly, since they thus supply reliable benchmarks in order to compare the different methods under investigation. Let us mention, however, that multifractal analysis is now successfully used in many fields of science (turbulence, clouds modelling, physiological signals and images, traffic data, rough interfaces...), see [1, 4, 13, 23, 34, 48, 58] and references therein. Inside mathematics, multifractal measures or functions were also shown to be relevant in many different areas, such as analytic number theory, Diophantine approximation, Peano-type functions, dynamical systems, stochastic processes,..., see [29, 30, 34] and references therein.

2. Multifractal Analysis of Measures

2.1. Mathematical Notions

We start by introducing the mathematical tools that are needed in the multifractal analysis of measures. The first one is the definition of Hausdorff dimension (see e.g., [20]).

Definition 1. Hausdorff dimension: Let $A \subset \mathbb{R}^d$. If $\varepsilon > 0$ and $\delta \in [0, d]$, we denote

$$M_\varepsilon^\delta = \inf_R \left(\sum_i |A_i|^\delta \right),$$

where R is an ε -covering of A , i.e. a covering of A by a countable collection of bounded sets $\{A_i\}_{i \in \mathbb{N}}$ of diameters $|A_i| \leq \varepsilon$. The infimum is therefore taken on all ε -coverings. For any $\delta \in [0, d]$, the δ -dimensional Hausdorff measure of A is

$$mes_\delta(A) = \lim_{\varepsilon \rightarrow 0} M_\varepsilon^\delta;$$

note that the limit exists (it can take the value $+\infty$) since M_ε^δ is a decreasing function of ε . There exists $\delta_0 \in [0, d]$ such that

$$\forall \delta < \delta_0, \quad mes_\delta(A) = +\infty; \quad \text{and} \quad \forall \delta > \delta_0, \quad mes_\delta(A) = 0.$$

This critical δ_0 is called the Hausdorff dimension of A .

Multifractal analysis is relevant for measures whose regularity changes from point to point. Therefore we need to introduce the following notion of pointwise regularity of measures.

Definition 2. Hölder exponent: Let $x_0 \in \mathbb{R}^d$ and let $\alpha \geq 0$. A nonnegative measure μ defined on \mathbb{R}^d belongs to $C^\alpha(x_0)$ if there exists a constant $C > 0$ such that, in a neighbourhood of x_0 ,

$$\mu(B(x_0, r)) \leq Cr^\alpha,$$

where $B(x_0, r)$ denotes the open ball of center x_0 and radius r . Let x_0 belong to the support of μ ; then the Hölder exponent of μ at x_0 is

$$h_\mu(x_0) = \sup\{\alpha : \mu \in C^\alpha(x_0)\}.$$

Definition 3. Singularity (or multifractal) spectrum: Let $E_\mu(H)$ denote the set of points where the Hölder exponent of μ takes the value H . (Note that $E_\mu(H)$ is included in the support of μ .) The spectrum of singularities of μ (denoted by $d_\mu(H)$) is the Hausdorff dimension of $E_\mu(H)$.

Remarks: In the previous definition, when $E_\mu(H) = \emptyset$, then its dimension is $-\infty$. This is actually more than a simple convention. Indeed, the multifractal formalism that is studied below is expected to yield $-\infty$ for the values of H for which $E_\mu(H) = \emptyset$.

The Hölder exponent of a measure is called the “local dimension” by some authors.

We will need to be able to deduce the Hölder exponent at every point from a “discretized version” of μ , i.e. from the values of μ on a countable collection of sets. A possible choice for this collection of sets is supplied by the dyadic cubes which are defined as follows.

Definition 4. Dyadic cube: A dyadic cube of scale j is a cube of the form

$$\lambda = \left[\frac{k_1}{2^j}, \frac{k_1+1}{2^j} \right) \times \cdots \times \left[\frac{k_d}{2^j}, \frac{k_d+1}{2^j} \right),$$

where $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$.

Each point $x_0 \in \mathbb{R}^d$ is contained in a unique dyadic cube of scale j , denoted by $\lambda_j(x_0)$.

The cube $3\lambda_j(x_0)$ is the cube of same center as $\lambda_j(x_0)$ and three times wider; i.e. it is the cube

$$\lambda = \left[\frac{k_1-1}{2^j}, \frac{k_1+2}{2^j} \right) \times \cdots \times \left[\frac{k_d-1}{2^j}, \frac{k_d+2}{2^j} \right).$$

The following lemma is a key ingredient in the derivation of the multifractal formalism for measures.

Lemma 2.1. *Let μ be a nonnegative measure defined on \mathbb{R}^d . Then*

$$h_\mu(x_0) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(\mu[3\lambda_j(x_0)])}{\log(2^{-j})} \right). \quad (2.1)$$

Proof. By definition of the Hölder exponent,

$$\forall \epsilon > 0, \exists r > 0, \forall r \leq R, \quad \mu(B(x, r)) \leq r^{H-\epsilon};$$

but $3\lambda_j(x_0) \subset B(x_0, 3\sqrt{d}2^{-j})$, so that

$$\mu(3\lambda_j(x_0)) \leq (3\sqrt{d})^{H-\epsilon} 2^{-j(H-\epsilon)},$$

and it follows that

$$h_\mu(x_0) \leq \liminf_{j \rightarrow +\infty} \left(\frac{\log(\mu(3\lambda_j(x_0)))}{\log(2^{-j})} \right).$$

On the other hand, if $h_\mu(x_0) = H$, then there exists a sequence of balls $B_n = B(x_0, r_n)$ and $\epsilon_n > 0$ such that $r_n \rightarrow 0$, $\epsilon_n \rightarrow 0$ and $r_n^{H+\epsilon_n} \leq \mu(B_n) \leq r_n^{H-\epsilon_n}$. Let j_n be such that $\frac{1}{2}2^{-j_n} < r_n \leq 2^{-j_n}$; then $B_n \subset 3\lambda_{j_n}(x_0)$ so that $\mu(B_n) \leq \mu(3\lambda_{j_n}(x_0))$, which implies the lower bound for the Hölder exponent. \square

Remark: This lemma relies heavily on the fact that the measure μ is nonnegative, and therefore is an increasing set function; indeed, if μ is no more assumed to be a nonnegative measure, then one easily checks that (2.1) is no more valid. This property will play a key-role in the following, therefore we introduce the following terminology.

Definition 5. Hierarchical set functions: A function μ defined on a collection of sets is called hierarchical if it is nonnegative and increasing, i.e. satisfies

$$A \subset B \implies \mu(A) \leq \mu(B);$$

The first example of multifractal measures studied were multiplicative cascades, which were introduced by B. Mandelbrot for modelling the distribution of energy in fully developed turbulence, see [45]; their mathematical properties were investigated by J.-P. Kahane and J. Peyrière in [36]. The purpose of the multifractal formalism is to derive the spectrum of singularities from global quantities which are effectively computable in practice. Such formulas were initially introduced by G. Parisi and U. Frisch in the context of fully developed turbulence in order to interpret the nonlinearity of the scaling function associated with the increments of the velocity field, see [55]; in the measure setting, and more precisely for invariant measures of dynamical systems they were introduced by T. Halsey, M. Jensen, L. Kadanoff, I. Procaccia and B. Shraiman in [23]. There exists several variants for the mathematical formulation of the multifractal formalism, see [12, 54] for instance, and we present the one given by R. Riedi in [56], because it presents a very good compromise between effective computability and numerical stability, as will be shown below.

2.2. Derivation of the Multifractal Formalism

Since Lemma 2.1 shows that the pointwise Hölder exponents can be derived from the quantities $\mu(3\lambda)$, it is natural to base a multifractal formalism on these quantities. We now assume that μ is compactly supported.

Definition 6. Measure (or box-aggregated) structure functions and scaling functions: Let Λ_j denote the collection of dyadic cubes of scale j . The structure function of the measure μ is

$$\Sigma_\mu(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j}^* \mu(3\lambda)^p, \quad (2.2)$$

where the notation Σ^* means that the sum is only taken on the cubes λ such that $\mu(\lambda) \neq 0$.

The scaling function of μ is defined for $p \in \mathbb{R}$ by

$$\eta_\mu(p) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(\Sigma_\mu(p, j))}{\log(2^{-j})} \right).$$

Let us now show why the spectrum of singularities is expected to be recovered from the scaling function. The definition of the scaling function roughly means that $\Sigma_\mu(p, j) \sim 2^{-\eta_\mu(p)j}$. Let us estimate the contribution to $\Sigma_\mu(p, j)$ of the cubes λ that cover the points of $E_\mu(H)$. Lemma 2.1 asserts that they satisfy $\mu(3\lambda) \sim 2^{-Hj}$; since we need about $2^{-d_\mu(H)j}$ such cubes to cover $E_\mu(H)$, the corresponding contribution roughly is

$$2^{-dj} 2^{d_\mu(H)j} 2^{-Hpj} = 2^{-(d-d_\mu(H)+Hp)j}.$$

When $j \rightarrow +\infty$, the dominant contribution comes from the smallest exponent, so that

$$\eta_\mu(p) = \inf_H (d - d_\mu(H) + Hp). \quad (2.3)$$

Proposition 1. For any compactly supported Borelian measure μ , the scaling function $\eta_\mu(p)$ is a concave function on \mathbb{R} .

Remark: We state this concavity result only for the first scaling function that we meet. However, the same proof applies to *all* scaling functions defined in the paper: All of them are concave.

Proposition 1 is a consequence of the following lemma of [24].

Lemma 2.2. *Let $(a_i)_{i=1,\dots,N}$ be a finite collection of positive real numbers. Then the function $\omega : \mathbb{R} \rightarrow \mathbb{R}$ defined by*

$$\omega(p) = \log \left(\sum_{i=1}^N a_i^p \right)$$

is a convex function on \mathbb{R} .

Proof of Lemma 2.2: The function $\omega(p)$ clearly is a continuous function defined on the whole \mathbb{R} . Thus, in order to prove that $\omega(p)$ is convex, it is sufficient to check that

$$\forall p, q \in \mathbb{R}, \quad \omega \left(\frac{1}{2}(p+q) \right) \leq \frac{1}{2}(\omega(p) + \omega(q)). \quad (2.4)$$

Consider the vectors in \mathbb{R}^N

$$A = (a_1^{p/2}, \dots, a_N^{p/2}) \quad \text{and} \quad B = (a_1^{q/2}, \dots, a_N^{q/2});$$

The Cauchy-Schwartz inequality applied to these vectors yields

$$\sum_{i=1}^N a_i^{(p+q)/2} \leq \left(\sum_{i=1}^N a_i^p \right)^{1/2} \left(\sum_{i=1}^N a_i^q \right)^{1/2}.$$

Taking the logarithm on both sides of this inequality yields exactly (2.4).

Proof of Proposition 1: For each j , we will apply Lemma 2.2 to the collection of $(\mu(3\lambda))_{\lambda \in \Lambda_j}$ such that $\mu(\lambda) \neq 0$ (and therefore $\mu(3\lambda) \neq 0$); this collection is finite, since μ is assumed to be compactly supported; it follows that, for any j , the function

$$p \rightarrow \log \left(\sum_{\lambda \in \Lambda_j}^* \mu(3\lambda)^p \right)$$

is convex; therefore, when divided by $\log(2^{-j})$, it is concave; Proposition 1 follows because concavity is preserved under taking infimums and pointwise limits, and therefore under taking liminfs.

Proposition 1 is in agreement with the fact that the right-hand side of (2.3) necessarily is a concave function (as an infimum of a family of linear functions) no matter whether $d_\mu(H)$ is concave or not. However, if the spectrum also is a concave function, then the Legendre transform in (2.3) can be inverted (as a consequence of general result on the duality of convex functions, see for instance Chapter 1.3 of [11]), which justifies the following definition.

Definition 7. A measure μ follows the multifractal formalism for measures if its spectrum of singularities satisfies

$$d_\mu(H) = \inf_{p \in \mathbb{R}} (d - \eta_\mu(p) + Hp). \quad (2.5)$$

Let us now explain the reason for the convention in the definition of Σ^* used in (2.2); indeed structure functions are often defined using $\mu(\lambda)$ instead of $\mu(3\lambda)$, and with the convention that the sum is taken only on the nonvanishing terms. One easily checks that this simpler way to define the structure function actually yields the same values of $\eta_\mu(p)$ for positive p ; however, it is no more the case if p is negative for the following reason: It may happen that the cube λ barely intersects the support of the measure; then $\mu(\lambda)$ does not vanish, but may be arbitrarily small and, when raised to a negative power, it will therefore lead to totally unstable computations; the convention for the Σ^* used above turns this drawback: When $\mu(\lambda) \neq 0$, the cube 3λ “widely” intersects the support of the measure.

The derivation exposed above is not a mathematical proof, and the determination of the range of validity of (2.5) (and of its variants) is one of the main mathematical problems concerning the multifractal analysis of measures. Nonetheless, let us stress the fact that the justification of this derivation relies heavily on (2.1), i.e. on the fact that the Hölder exponent of a measure can be estimated from the set of values that it takes on dyadic cubes. The formulation of the multifractal formalism given by (2.5) combines two advantages:

- It is based on quantities that are effectively computable in practice: By contrast with alternative formulas proposed by some mathematicians, the structure function is not based on the consideration of a non-countable collection of coverings of the support of μ .
- The scaling function has “good” mathematical properties, see [38, 56] (for instance it is invariant under bi-Lipschitz deformations of the measure, which is a natural requirement since the spectrum of singularities has this invariance property).

This last remark points the way towards the kind of criteria that we will use in order to select multifractal formalisms: In situations where the validity of several possible multifractal formalisms cannot be justified in all generality, a weaker benchmark in order to compare them will be to determine which ones satisfy invariance properties which are obvious for the spectrum of singularities. Such properties will be referred to as *robustness properties* in the following. For instance, if the scaling function is defined through wavelet coefficients, we will require that it is independent of the (smooth enough) wavelet basis chosen. Note also that, in several applications, it happens that the spectrum of singularities itself has no direct scientific interpretation and multifractal analysis is only used as a classification tool in order to discriminate between several types of signals; then, one is no more concerned with the validity of (2.5) but only with having its right-hand side defined in a meaningful way; therefore, in such cases, robustness

criteria are the only mathematical requirements which remain in order to compare the pertinence or several possible scaling functions.

3. Multifractal Analysis of Functions: Increments vs. Oscillations

Let us now consider the multifractal analysis of functions. We will start by recalling the corresponding relevant definitions in this context. Multifractal analysis is relevant for functions whose regularity changes from point to point. Therefore we introduce the following notion of pointwise regularity of functions, which is the most widely used. (Note however that in some specific settings, other pointwise regularity exponents of functions can be used: The weak scaling exponent, see Section 6 and the T_u^p exponent, see [32] and references therein.)

Definition 8. Hölder exponent: Let $x_0 \in \mathbb{R}^d$ and let $\alpha \geq 0$. A locally bounded function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ belongs to $C^\alpha(x_0)$ if there exists a constant $C > 0$ and a polynomial P satisfying $\deg(P) < \alpha$ and such that, in a neighbourhood of x_0 ,

$$|f(x) - P(x - x_0)| \leq C|x - x_0|^\alpha.$$

The Hölder exponent of f at x_0 is

$$h_f(x_0) = \sup\{\alpha : f \in C^\alpha(x_0)\}.$$

Definition 9. Singularity (or multifractal) spectrum: Let f be a locally bounded function, and let $E_f(H)$ denote the set of points where the Hölder exponent of f takes the value H . The spectrum of singularities of f (denoted by $d_f(H)$) is the Hausdorff dimension of $E_f(H)$.

Remarks: If $h_f(x_0) < 1$ (which is often the case in signal processing), then the polynomial $P(x - x_0)$ boils down to $f(x_0)$.

The function $h_f(x_0)$ may take the value $+\infty$.

If $0 < h_f(x_0) < 1$, then the Hölder exponent expresses how “spiky” the graph of f is at x_0 . For instance the Hölder exponent of $f(x) = |x - x_0|^\alpha$ is α at x_0 and $+\infty$ elsewhere (if α is not an even integer).

3.1. Comparison of Multifractal Formalisms

The numerical determination of the spectrum of singularities of a signal meets the same problem as for measures. The multifractal formalism in this context was introduced by G. Parisi and U. Frisch; they proposed to derive it from the estimation of the L^p norm of increments of the signal [55]: Let us assume that the function f considered is a one-variable function. A structure function based on increments is

$$\Sigma_f^1(p, j) = 2^{-j} \sum_k^* \left| f\left(\frac{k+1}{2^j}\right) - f\left(\frac{k}{2^j}\right) \right|^p \quad (3.1)$$

where the \sum^* means that the sum is taken only on non vanishing terms. The scaling function of f is defined for $p \in \mathbb{R}$ by

$$\eta_f^1(p) = \liminf_{j \rightarrow +\infty} \left(\frac{\log \left(\Sigma_f^1(p, j) \right)}{\log(2^{-j})} \right). \quad (3.2)$$

The same arguments as for the derivation of the multifractal formalism for measures lead to

$$d_f(H) = \inf_{p \in \mathbb{R}} (1 - \eta_f^1(p) + Hp) \quad (3.3)$$

(recall that we deal with functions of one variable here, so that d is replaced by 1 in this formula). Since (3.1) involves only first order differences, one expects (3.3) to yield the spectrum of singularities of f only if all Hölder exponents take values less than 1.

A first problem which is met here is that there is no formula corresponding to (2.1) and based on increments of f : For instance, if $0 < \alpha < 1$; the function

$$x^\alpha \sin \left(\frac{2\pi}{x} \right)$$

vanishes at the points 2^{-j} but its Hölder exponent at 0 is not $+\infty$ but α . A second problem is that this structure function does not clearly extend to the several dimensional setting. (Which increments should be preferred on a cube?)

Let us now describe an alternative point of view which solves these difficulties. The function f is defined on \mathbb{R}^d , and we assume for the sake of simplicity that $0 < h_f(x) < 1$; then the local quantity based on dyadic cubes which is considered is the *oscillation* of f .

Definition 10. Oscillations: The oscillation of a function f over a set K is

$$Osc_f(K) = \sup_{x \in K} f(x) - \inf_{x \in K} f(x).$$

The motivation for basing the study of pointwise Hölder regularity on the oscillation is that it is a hierarchical notion in the sense of Definition 5: Indeed, clearly,

$$\mu \subset \nu \implies Osc_f(\mu) \leq Osc_f(\nu)$$

and therefore the Hölder exponent at each point can be derived from the knowledge of the oscillation on the countable collection of dyadic cubes, as shown by the following lemma.

Lemma 3.1. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a locally bounded function satisfying $h_f(x_0) = H$, with $0 < H < 1$; then*

$$H = \liminf_{j \rightarrow +\infty} \left(\frac{\log(Osc_f(3\lambda_j(x_0)))}{\log(2^{-j})} \right). \quad (3.4)$$

This lemma corresponds to Lemma 2.1 in the context of functions; its proof is very similar, so that we leave it. Let us just insist on the fact that it holds because the oscillation is a hierarchical notion, when increments are not.

Following the same arguments as in the case of positive measures, one can base a multifractal formalism on this lemma by introducing the structure function

$$\Sigma_f^2(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j}^* (\text{Osc}_f(3\lambda))^p;$$

following in this function setting the idea of [56], the Σ^* means that the sum is restricted to the cubes λ for which $\text{Osc}_f(\lambda) \neq 0$. The corresponding *scaling function* of f is

$$\eta_f^2(p) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(\Sigma_f^2(p, j))}{\log(2^{-j})} \right). \quad (3.5)$$

The same arguments as above lead to the formula

$$d_f(H) = \inf_{p \in \mathbb{R}} (d - \eta_f^2(p) + Hp), \quad (3.6)$$

which we expect to hold only when the spectrum of singularities of f is supported inside the interval $(0, 1)$ (i.e. if there are no Hölder exponents larger than 1 in the signal).

3.2. Examples: Brownian Motion and Fractional Brownian Motions

We do not intend to investigate in details the properties of the multifractal formalisms supplied by (3.3) and (3.6), because our main motivation is to focus on wavelet-based formulas, which will be shown to possess better mathematical and numerical properties. However, we will only illustrate them by simple examples supplied by Brownian motion and by fractional Brownian motions (hereafter F.B.M.). Such examples provide us with theoretical and numerical benchmarks on which alternative formalisms can as well be tested.

Recall that Brownian motion is the only stochastic process (or random function) $(B_t)_{t \geq 0}$ with stationary independent increments (i.e. satisfying if $t > s$, $B_t - B_s$ is independent of B_s and has the same law as B_{t-s}) and with continuous sample paths. (Uniqueness is implied by the normalization $E(|B_t|^2) = t$.)

Fractional Brownian motion of index γ ($0 < \gamma < 1$) is the only Gaussian random process $(B_t^\gamma)_{t \geq 0}$ satisfying

$$E(|B_t^\gamma - B_s^\gamma|^2) = |t - s|^{2\gamma}.$$

One can show that Brownian motion is precisely $B_t^{1/2}$. The key role played by fractional Brownian motions in signal processing comes from the fact that they supply the most simple one parameter family of stochastic processes with stationary increments, and therefore are widely used in modelling. We will use the following important feature: F.B.M. of index γ can be deduced from Brownian motion by a sample path by sample path fractional integration of order $\gamma - 1/2$ if $\gamma > 1/2$, and by a sample path by sample path fractional derivation of order $(1/2) - \gamma$ if $\gamma < 1/2$.

Recall that, with probability 1, a sample path of Brownian motion has everywhere the Hölder exponent $1/2$, so that its spectrum of singularities is

$$d(H) = \begin{cases} 1 & \text{if } H = \frac{1}{2} \\ -\infty & \text{else,} \end{cases} \quad (3.7)$$

see for instance [1, 35].

The following theorem illustrates the superiority of the multifractal formalism based on oscillations (as opposed to increments).

Theorem 3.2. *Let B_t be a generic sample path of Brownian motion; then, with probability 1, the multifractal formalism based on increments (3.3) yields that a.s.*

$$\inf_{p \in \mathbb{R}} (d - \eta_B^1(p) + Hp) = \begin{cases} \frac{3}{2} - H & \text{if } H \in [\frac{1}{2}, \frac{3}{2}] \\ -\infty & \text{else,} \end{cases}$$

whereas the multifractal formalism based on oscillations (3.6) yields the correct spectrum given by (3.7).

The proof of Theorem 3.2 is given in the Appendix where, in particular, we will show that a.s.

$$\eta_B^1(p) = \begin{cases} p/2 & \text{if } p \geq -1 \\ 1 + 3p/2 & \text{else,} \end{cases} \quad (3.8)$$

and a.s. $\forall p \in \mathbb{R}$, the \liminf in (3.2) is a true limit, which clearly implies the first part of Theorem 3.2. The fact that this \liminf actually is a limit is important when one wants to double check numerically this result in simulation, since, in practice, only true limits can be estimated. Note that we will give another proof of the second part of Theorem 3.2 in the more general setting supplied by F.B.M.; however, we prefer to separate the case of Brownian motion which will be treated completely by elementary means, whereas the F.B.M. case requires the use of more sophisticated tools derived from the so-called *small ball estimates*.

Let us consider a generic sample path of Brownian motion B_t on $[0, 1]$ (by scaling invariance, the particular choice of interval is irrelevant). The increments $B_{(k+1)/2^j} - B_{k/2^j}$ are I.I.D. random variables of common law $2^{-j/2}\chi_{j,k}$, where the $\chi_{j,k}$ are standard Gaussians; thus, in order to estimate (3.1) for Brownian motion, we have to estimate the order of magnitude of

$$A(p, j) = \sum_{k=1}^{2^j} |\chi_{j,k}|^p, \quad (3.9)$$

and the structure function will be

$$\Sigma_B^1(p, j) = 2^{-j(1+p/2)} A(p, j).$$

(With probability one, a non-degenerate Gaussian random variable does not vanish so that, in all computations that will be performed in this section and in the following concerning Brownian motion or F.B.M., the \sum^* sums are just usual sums.)

Note at this point that the problem of estimation of the $A(p, j)$ is not the same as estimating moments of order p of a Gaussian variable, since the computations here are performed sample path by sample path, and not in expectation; indeed, if such a process models a given observed signal, then the “rule of the game” is that one sample path is observed, and not averages over a large number of realizations. This remark is particularly relevant for large negative values of p where the two approaches lead to different results (moments diverge if $p < -1$, whereas the order of magnitude of $A(p, j)$) can always be estimated for any negative value of p).

Let us now consider the fractional Brownian case. With probability 1, a sample path of F.B.M. of order β has everywhere the Hölder exponent β so that its spectrum of singularities is

$$\left. \begin{aligned} d(H) &= 1 && \text{if } H = \beta \\ &= -\infty && \text{else,} \end{aligned} \right\} \quad (3.10)$$

see for instance [1, 35].

Theorem 3.3. *Let $\beta \in (0, 1)$ and $B_\beta(t)$ be a generic sample path of F.B.M. of order β ; then, with probability 1,*

$$\forall p \in \mathbb{R}, \quad \eta_{B_\beta}^2(p) = \beta p \quad (3.11)$$

and the \liminf in (3.5) is a true limit.

The multifractal formalism based on oscillations (3.6) yields the correct spectrum (3.10) for the F.B.M..

Proof of Theorem 3.3: First we recall a well-known result concerning the uniform modulus of continuity of the sample paths of F.B.M. of order β [35]: With probability 1, there exists $C > 0$ such that

$$\sup_t \left(\sup_{h \leq 1} \frac{|B_\beta(t+h) - B_\beta(t)|}{|h|^\beta |\log h|} \right) \leq C$$

It follows that, with probability 1, all oscillations

$$Osc_{B_\beta}(I_{j,k}) = \sup_{s \in I_{j,k}} B_\beta(s) - \inf_{s \in I_{j,k}} B_\beta(s)$$

are bounded by $CN^{-\beta} \log N$ (where $N = 2^j$ is the number of intervals considered).

‘Small ball estimates’ for a random process X_t are concerned with the estimation of

$$\mathbb{P} \left(\sup_{0 \leq s \leq t} |X_s| \leq \epsilon \right)$$

Lower bounds for the oscillation are a consequence of the small ball estimates for the F.B.M.; indeed

$$\sup_{0 \leq s \leq t} B_\beta(s) - \inf_{0 \leq s \leq t} B_\beta(s) \leq 2 \sup_{0 \leq s \leq t} |B_\beta(s)|$$

and, by Theorem 2.1 of [53], if $\epsilon \leq t^\beta$,

$$\mathbb{P} \left(\sup_{0 \leq s \leq t} |B_\beta(s)| \leq \epsilon \right) \leq \exp \left(-Ct\epsilon^{-1/\beta} \right).$$

Since all oscillations have the same law it follows that, for a given N , all oscillations $Osc_{B_\beta}(I_{j,k})$ are larger than $2N^{-\beta}(\log N)^{-2\beta}$ with probability at least $1 - N \exp(-C(\log N)^2)$. As above, a direct application of the Borel-Cantelli Lemma shows that the multifractal formalism based on oscillations yields the correct spectrum (3.6) for a.e. sample path of the F.B.M.

As mentioned above, our purpose in this section was only to illustrate the two multifractal formalisms based on increments and oscillations on the particular examples provided by Brownian motion and F.B.M.. However, the proof of Theorem 3.3 clearly shows that results on the multifractal formalism based on oscillations immediately follow from small ball estimates for the process considered (such estimates for Gaussian processes can be found in [41, 42, 53] for instance), so that general results of validity of the multifractal formalism supplied by (3.6) could clearly be easily proved for general Gaussian processes.

4. Multifractal Analysis of Functions: Wavelet Based Formulas

Lemma 3.1 can be generalized to higher Hölder exponents by using higher order differences in the definition of the oscillation, see [29], however, it leads to rather complicated quantities for the computation of structure functions, and it presents strong instabilities under the presence of noise. Therefore, once wavelet techniques were available, alternative formulas were proposed; they were based either on the continuous wavelet transform of the signal (by Arneodo et al., see [4, 6] and references therein) or on its coefficients on an orthonormal wavelet basis, see [27, 29] and references therein. The starting point of all these methods is a wavelet characterization of the Hölder exponent. Let us start by recalling basic definitions concerning wavelet expansions. Though formulas based on the discrete wavelet coefficients were introduced later than those based on the continuous wavelet transform, we start by describing the discrete ones, since they are in spirit very close to the dyadic partitionings we introduced in the measure setting, and they pave the way to the wavelet leaders technique of Section 5.

4.1. Wavelet Bases

We now recall the definition of wavelet bases. Let $r \in \mathbb{N}$; an r -smooth wavelet basis of \mathbb{R}^d is composed of $2^d - 1$ wavelets $\psi^{(i)}$ which belong to C^r and satisfy the following properties:

- $\forall i, \forall \alpha$ such that $|\alpha| \leq r$, $\partial^\alpha \psi^{(i)}$ has fast decay,
- The set of functions $2^{dj/2} \psi^{(i)}(2^j x - k)$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^d$, $i \in \{1, \dots, 2^d - 1\}$ is an orthonormal basis of $L^2(\mathbb{R})^d$.

The wavelet basis is ∞ -smooth if it is r -smooth for any $r \in \mathbb{R}$, in which case all wavelets $\psi^{(i)}$ belong to the Schwartz class.

Thus any function f in $L^2(\mathbb{R}^d)$ can be written

$$f(x) = \sum c_{j,k}^{(i)} \psi^{(i)}(2^j x - k) \quad (4.1)$$

where

$$c_{j,k}^{(i)} = 2^{dj} \int f(x) \psi^{(i)}(2^j x - k) dx.$$

(Note that, in (4.1), wavelets are not normalized for the L^2 norm but for the L^∞ norm, which avoids an extra factor $2^{dj/2}$ in all mathematical results concerning Hölder regularity.) Let us note at this point that it is often relevant to use a slight generalization based on biorthogonal wavelets, the definition of which we now recall.

A *Riesz basis* of an Hilbert space H is a collection of vectors (e_n) such that the finite linear expansions $\sum_{n=1}^N a_n e_n$ are dense in H and

$$\exists C, C' > 0 : \forall N, \forall a_n, \quad C \sum_{n=1}^N |a_n|^2 \leq \left\| \sum_{n=1}^N a_n e_n \right\|_H^2 \leq C' \sum_{n=1}^N |a_n|^2.$$

Two collections of functions (e_n) and (f_n) form *biorthogonal bases* if each collection is a Riesz basis, and if $\langle e_n | f_m \rangle = \delta_{n,m}$. When such is the case, any element $f \in H$ can be written

$$f = \sum_{n=1}^{\infty} \langle f | f_n \rangle e_n.$$

Biorthogonal wavelet bases are couples of bases of the form $2^{dj/2} \tilde{\psi}^{(i)}(2^j x - k)$ and $2^{dj/2} \psi^{(i)}(2^j x - k)$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^d$, $i \in \{1, \dots, 2^d - 1\}$ which are biorthogonal (for the L^2 norm).

The relevance of biorthogonal wavelet bases is due to two reasons: On one hand their construction is more flexible and, for instance, allows for wavelets which have some symmetry properties, which is an important requirement in image processing, see [15]; on the other hand, for theoretical purposes, this setting is often more adapted to derive the properties of some random processes; we will see the example of Brownian motion and of F.B.M. in Sections 4.3 and 5.4 where a decomposition on well chosen biorthogonal wavelet bases allows to decorrelate the wavelet coefficients of these processes (the wavelet coefficients become independent random variables), and therefore greatly simplifies their analysis.

Wavelets will be indexed by dyadic cubes as follows: Since i takes $2^d - 1$ values, we can consider that i takes values among all dyadic subcubes λ_i of $[0, 1]^d$ of width $1/2$ except for $[0, 1/2]^d$; thus, the set of indices (i, j, k) can be relabelled using dyadic cubes as follows: λ denotes the cube $\{x : 2^j x - k \in \lambda_i\}$; we note $\psi_\lambda(x) = \psi^{(i)}(2^j x - k)$ (an L^∞ normalization is used), and $c_\lambda = 2^{dj} \int \psi_\lambda(x) f(x) dx$. We will use the notations $c_{j,k}^{(i)}$ or c_λ indifferently for wavelet coefficients. Note that the index λ gives an information on the localization and the scale of the corresponding wavelet; for instance, if the wavelets $\psi^{(i)}$ are compactly supported

then $\exists C : \text{supp}(\psi_\lambda) \subset C\lambda$ where $C\lambda$ denotes the cube of same center as λ and C times larger; thus the indexation by the dyadic cubes is more than a simple notation: The wavelet ψ_λ is “essentially” localized around the cube λ . Finally, Λ_j will denote the set of dyadic intervals λ of width 2^{-j} .

4.2. Hölder Regularity and Derivation of the Multifractal Formalism

The wavelet characterization of the Hölder exponent requires the following regularity hypothesis, which is slightly stronger than continuity.

Definition 11. Uniform Hölder function: A function f is a uniform Hölder function if there exists $\epsilon > 0$ such that $f \in C^\epsilon(\mathbb{R}^d)$, i.e.

$$\exists C > 0 \quad \text{such that} \quad \forall x, y \in \mathbb{R}, \quad |f(x) - f(y)| \leq C|x - y|^\epsilon.$$

The following proposition was proved in [25].

Proposition 2. Let $\alpha > 0$. If f is $C^\alpha(x_0)$, then there exists $C > 0$ such that the wavelet coefficients of f satisfy

$$\forall j \geq 0, \quad |c_{j,k}| \leq C2^{-\alpha j}(1 + |2^j x_0 - k|)^\alpha. \quad (4.2)$$

Conversely, if (4.2) holds and if f is uniform Hölder, then $\exists C > 0$ and a polynomial P satisfying $\deg(P) < \alpha$ and such that, in a neighbourhood of x_0 ,

$$|f(x) - P(x - x_0)| \leq C|x - x_0|^\alpha \log(1/|x - x_0|).$$

The *influence cone above x_0* is the set of dyadic cubes which are of the form $\lambda_j(x_0)$ and their $3^d - 1$ immediate neighbours at the same scale, i.e. the dyadic cubes λ of scale j such that $\text{dist}(\lambda, \lambda_j(x_0)) = 0$. Note that it is composed of the cubes of scale j included in $3\lambda_j(x_0)$. The regularity criterium supplied by Lemma 2 has often been loosely interpreted as stating that the wavelet coefficients decay like $2^{-\alpha j}$ in the influence cone; indeed, it is the case for *cusp-like* singularities which behave like

$$A + B|x - x_0|^\alpha$$

in the neighbourhood of x_0 ; such functions are characterized by the fact that there are no strong oscillations in the neighbourhood of x_0 . Let us assume for the moment that, indeed, the function considered exhibits only this type of pointwise singularities, and therefore, the Hölder exponent at x_0 is given by

$$h_f(x_0) = \lim_{j \rightarrow +\infty} \left(\frac{\log(|c_{\lambda_n}|)}{\log(2^{-j_n})} \right), \quad (4.3)$$

where the λ_n are dyadic cubes of scale j_n in the influence cone above x_0 . Following the same arguments as above, we introduce the structure function

$$W_f^1(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j}^* |c_\lambda|^p, \quad (4.4)$$

where the \sum^* means here that the sum is taken on the nonvanishing wavelet coefficients. The corresponding *scaling function* of f is

$$\zeta_f^1(p) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(W_f^1(p, j))}{\log(2^{-j})} \right).$$

One is therefore led to the following multifractal formalism

$$d_f(H) = \inf_{p \in \mathbb{R}} (d - \zeta_f^1(p) + Hp). \quad (4.5)$$

Several criticisms can be addressed to this multifractal formalism:

1. It implicitly assumes that the only singularities met are *cusplike* singularities. This is of course an assumption which is impossible to check on a signal. Let us briefly mention other types of singularities which can be met. On the opposite from cusp singularities are the *chirplike* singularities which display very strong oscillations in the neighbourhood of x_0 , such as

$$C_{\alpha, \beta}(x) = |x - x_0|^\alpha \sin \left(\frac{1}{|x - x_0|^\beta} \right), \quad (4.6)$$

where $\alpha > 0$ and $\beta > 0$. Such functions are counterexamples to (4.3); indeed their wavelet coefficients display a much stronger decay in the influence cone: They decay faster than 2^{-Nj} for any $N > 0$. They have indeed large wavelet coefficients which make (4.2) optimal for them too, but these large coefficients are situated far away from the influence cone: They correspond to indices (j, k) such that $|2^j x_0 - k| \sim 2^{-j/(1+\beta)}$, see [33] for precise statements. This is illustrated numerically in Fig. 1, top row.

2. The quantity

$$\liminf_{j \rightarrow +\infty} \left(\frac{\log(|c_{\lambda_j(x_0)}|)}{\log(2^{-j})} \right) \quad (4.7)$$

on which the corresponding exponent is based does not define a quantity which is independent of the wavelet basis chosen.

3. Wavelet coefficients can be extremely small by chance, so that we expect the structure function defined in (4.4) to be completely unstable for $p < 0$.
4. One can show that the scaling function $\zeta_f^1(p)$ is independent of the (smooth) wavelet basis chosen when $p > 0$ but it is not the case any longer if $p < 0$.

We will address these problems in a detailed way in the following: In Section 5 we will introduce a multifractal formalism based on alternative quantities that will have the required robustness properties; and in Section 6 we will show that the wavelet-based formula (4.5) actually is a multifractal formalism adapted (for $p > 0$) to another exponent, the *weak-scaling exponent*, and we will extend this multifractal formalism in a robust way for $p < 0$.

4.3. Examples: Brownian Motion and Fractional Brownian Motions

We now show that, even when the signal only displays cusp-like singularities, then (4.5) does not necessarily yield the right spectrum of singularities. This pathology already appears on the particularly striking examples supplied by the sample paths of Brownian motion, and of F.B.M. (we treat only the F.B.M. case since Brownian motion is the subcase corresponding to the Hurst exponent $\beta = 1/2$).

An important result of Paul Lévy states that, if (e_n) is an orthogonal basis of $L^2(\mathbb{R})$, and if f_n denotes a primitive of e_n , then Brownian motion can be decomposed on the f_n in the following particularly simple way

$$B_t = \sum \chi_n(f_n(t) - f_n(0))$$

where the χ_n are independent identically distributed (i.i.d.) standard Gaussians. Let us apply this result using an orthonormal wavelet basis for the e_n . A primitive of ψ is supplied by the function ψ_1 whose Fourier transform is given by $\hat{\psi}_1(\xi) = \hat{\psi}(\xi)/\xi$. The primitive of $2^{j/2}\psi(2^jx - k)$ is $2^{-j/2}\psi_1(2^jx - k)$, therefore

$$B_t = \sum_{j,k} \chi_{j,k}(\psi_{j,k}(t) - \psi_{j,k}(-k)).$$

The contributions of the terms corresponding to $j < 0$ and the constant terms belongs to C^∞ (if the wavelet used is C^∞), therefore one can write

$$B_t = \sum_{j \geq 0, k} \chi_{j,k} \psi_{j,k}(t) + R(t),$$

where $R(t)$ is a C^∞ process. We can apply the same argument in order to obtain a wavelet decomposition of F.B.M. since, as mentioned already, F.B.M. of index γ can be deduced from Brownian motion by a sample path by sample path fractional integration of order $\gamma - 1/2$ if $\gamma > 1/2$, or a fractional derivation of order $(1/2) - \gamma$ if $\gamma < 1/2$. (We refer the reader to [2, 51] where the wavelet decomposition of F.B.M. is investigated in details and, in particular, the remainders $R(t)$ and $R^\alpha(t)$ are given an explicit form which allows for accurate simulations of the long range dependence.) Let

$$\hat{\psi}_\alpha(\xi) = \frac{1}{|\xi|^\alpha} \hat{\psi}(\xi) \quad (4.8)$$

(ψ_α is the fractional integral of ψ of order α). If the wavelet ψ has enough vanishing moments, then ψ_α is a wavelet and the $2^{j/2}\psi_\alpha(2^jx - k)$ and the $2^{j/2}\psi_{-\alpha}(2^jx - k)$ form biorthogonal bases, see [29, 49]; the point of using these bases in order to analyze F.B.M. is that, as a consequence of the previous remarks, the coefficients of F.B.M. are decorrelated on it. More precisely, if $t \in [0, 1]$ then

$$B_\beta(t) = \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}} 2^{-\beta j} \xi_{j,k} \psi_{\beta+1/2}(2^j t - k) + R(t) \quad (4.9)$$

where $R(t)$ is a C^∞ random process, and the $\xi_{j,k}$ are I.I.D. standard centered Gaussians, see [2, 51]. Therefore

$$W_{B_\beta}^1(p, j) = 2^{-\beta p j} \sum_{k=1}^{2^j} |\chi_{j,k}|^p,$$

which, up to the factor $2^{-(1+\beta p)j}$ has exactly the same expression as $A(p, j)$ defined by (3.9). Therefore the computation performed in Section 3.2 yields the following result.

Proposition 3. Let $B_\beta(t)$ be a generic sample path of F.B.M. of order $\beta \in (0, 1)$, and assume that the wavelet used is C^2 . Then, with probability 1, the wavelet multifractal formalism (4.5) applied to $B_\beta(t)$ yields

$$\left. \begin{aligned} \inf_{p \in \mathbb{R}} (d - \zeta_{B_\beta}^1(p) + Hp) &= \beta + 1 - H & \text{if } H \in [\beta, \beta + 1] \\ &= -\infty & \text{else,} \end{aligned} \right\} \quad (4.10)$$

and the \liminf in the definition of the scaling function $\zeta_{B_\beta}^1(p)$ is a limit.

5. Wavelet Leaders

In this section, we exhibit quantities d_λ called the wavelet leaders which are based on the wavelet coefficients, and such that the formula corresponding to (4.7) yields an exponent which is independent of the wavelet basis chosen, and which, under a very mild uniform regularity assumption, actually is the Hölder exponent. We investigate the properties of the multifractal formalism based on these quantities and, in particular, the stability of the structure function for $p < 0$.

5.1. Pointwise Hölder Regularity Conditions

We saw that the Hölder exponent of a function f is not necessarily given by (4.7). Another indication that (4.7) is not the right quantity to consider in the derivation of the multifractal formalism is that the necessity to base a multifractal formalism on a quantity which is “hierarchical” (in the sense of Definition 5) was put into light several times in previous sections. A simple quantity which is larger than $|c_\lambda|$ and is hierarchical is supplied by the wavelet leaders, which are defined as follows.

Definition 12. Wavelet Leaders: Let f be a bounded function; the wavelet leaders of f are

$$d_\lambda = \sup_{\lambda' \subset 3\lambda} |c_{\lambda'}|. \quad (5.1)$$

If x_0 is a given point, then

$$d_j(x_0) = d_{\lambda_j(x_0)}.$$

Note that since $f \in L^\infty$,

$$|c_\lambda| \leq 2^{dj} \int |f(x)| |\psi_\lambda(x)| dx \leq C \sup |f(x)|,$$

so that $\forall \lambda, d_\lambda \leq C \|f\|_\infty$, and therefore the wavelet leaders are finite. We will usually assume in the following that the function studied is bounded, so that the wavelet leaders are finite. Note however that wavelet leaders are well defined under the weaker assumption that f belongs to the *Bloch space* which coincides with the Besov space $B_\infty^{0,\infty}$ and is characterized by the condition

$$\exists C > 0, \forall \lambda, |c_\lambda| \leq C$$

(see Chap. 6.8 of [49] and references therein for properties of this function space).

The following proposition allows to characterize the pointwise regularity by a decay condition of the $d_j(x_0)$ when $j \rightarrow +\infty$.

Proposition 4. Let $f \in L^\infty(\mathbb{R}^d)$ and $\alpha > 0$. The condition

$$\forall j \geq 0, \quad d_j(x_0) \leq C 2^{-\alpha j} \quad (5.2)$$

is equivalent to (4.2). (This is illustrated numerically in Fig. 1, bottom row.)

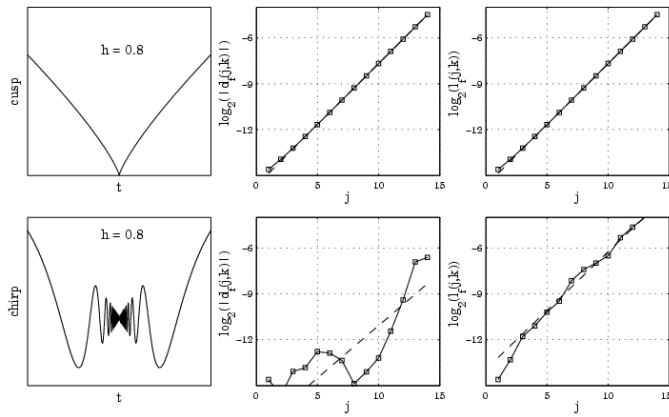


FIGURE 1. **Cusp vs Chirp singularity.** Left column, cusp singularity $|x - x_0|^h$ (top row) versus chirp singularity (bottom row) $|x - x_0|^h \sin(\frac{1}{|x - x_0|^\beta})$ with $\beta = 1$. Central column, wavelet coefficients, c_λ for λ such that $2^{-j}k = t_0$, right column, wavelet leaders, d_λ for λ such that $2^{-j}k = t_0$. One sees that while the decrease along scales j of the wavelet coefficients correctly characterizes the Hölder exponent cusp singularities, it does not for chirp-type ones. The decrease along scales j of the wavelet leaders do accurately characterize all type of singularities, as in Eq. 5.2.

Proof of Proposition 4: We first prove that (4.2) implies (5.2). Let $j \geq 0$ and assume that $\lambda' \subset 3\lambda_j(x_0)$. Since

$$|c_{\lambda'}| \leq C2^{-\alpha j'}(1 + |2^{j'}x_0 - k'|)^\alpha, \quad j' \geq j-1 \quad \text{and} \quad |k'2^{-j'} - x_0| \leq 4d2^{-j},$$

it follows that $|c_{\lambda'}| \leq C2^{-\alpha j}$, so that $d_j(x_0) \leq C2^{-\alpha j}$.

Let us now prove the converse result. If λ' is a cube of side $2^{-j'}$, denote by $\lambda (= \lambda(\lambda'))$ the dyadic cube defined by

- If $\lambda' \subset 3\lambda_{j'}(x_0)$, then $\lambda = \lambda_{j'}(x_0)$,
- else, if $j = \sup\{l : \lambda' \subset 3\lambda_l(x_0)\}$, then $\lambda = \lambda_j(x_0)$, and it follows that

$$2^{-j-1} \leq |k'2^{-j'} - x_0| \leq 4d2^{-j}.$$

In the first case, by hypothesis, $|c_{\lambda'}| \leq d_{j'}(x_0) \leq C.2^{-\alpha j'}$. In the second case,

$$|c_{\lambda'}| \leq d_j(x_0) \leq C2^{-\alpha j} \leq C|x_0 - k'2^{-j'}|^\alpha,$$

so that (4.2) holds in both cases.

Note that, as a consequence of Proposition 4 and Theorem 3 of [26], it follows that Condition (5.2) is independent of the wavelet basis which is chosen, if the wavelets are r -smooth with $r > \alpha$.

5.2. Multifractal Formalisms

The reader will have noticed the striking similarity between Lemma 2.1 and Proposition 4: Both provide a characterization of pointwise Hölder regularity by a condition on hierarchical quantities considered in the influence cone. Therefore one can derive the multifractal formalism for functions exactly in the same manner as was done previously for measures. It is therefore natural to use a structure function based on wavelet leaders, i.e. which is of the form

$$2^{-dj} \sum_{\lambda \in \Lambda_j}^* (d_\lambda)^p;$$

however, obtaining the correct definition for the \sum^* in this setting is much more delicate than in the measure setting of Section 2; the problem for measures was to find a way to keep the contribution of a cube in the structure function only if it did include an important amount of the support of the measure. When one tries to reproduce this feature in the wavelet setting one meets three problems:

- The size of the support of the wavelet changes with the wavelet used, so that formulas based on the consideration that the support of the function analyzed intersects “widely” the support of the wavelet cannot be universal, but have to be tailored to the particular wavelet basis used.
- Such considerations become irrelevant if the support of the wavelet is the whole \mathbb{R}^d , which is the case if the wavelet used belongs to the Schwartz class.
- If the wavelets used have a finite smoothness and a finite number of vanishing moments, then they cannot analyze smoother parts of the function. If such smoother parts occur on a set of dimension d , the multifractal formalism can yield incorrect results for the largest H s (which may be infinite); since

the multifractal formalism yields a concave function, this error can make the whole decreasing part of the spectrum wrong (which is the part obtained for $p < 0$ in the Legendre transform formula).

We are confronted with a deadlock: \sum^* formulas make sense only if the wavelet is compactly supported, hence has a finite smoothness, in which case, the $p < 0$ part of the scaling function may be completely unstable, since it can be changed by adding an arbitrarily small and smooth perturbation on the function.

Therefore, there is no universal formula without any drawback; However, this discussion shows that one may use the following “rule of thumb”: On one hand, it is reasonable to use a \sum^* formula based on compactly supported wavelets when analyzing compactly supported functions which are not arbitrarily smooth inside their support; on the other hand, one should rather use wavelets in the Schwartz class when analyzing functions with full support, in which case wavelet leaders are not expected to vanish (this could only happen for “toy examples”, i.e. for artificial functions which are defined through their wavelet coefficients on the precise wavelet basis which is used for the analysis); and, in that case, we do not need to eliminate vanishing wavelet leaders in the definition of the structure function. Therefore, we separate two cases depending on whether the wavelets are compactly supported or belong to the Schwartz class.

First case: Compactly supported wavelets

Definition 13. Leader based Multifractal Formalism 1: Let f be a uniform Hölder function and assume that the wavelets used are compactly supported. The extended wavelet leaders are

$$e_\lambda = \sup_{\text{supp}(\psi_{\lambda'}) \subset 3\text{supp}(\psi_\lambda)} |c_{\lambda'}|,$$

where $\text{supp}(\psi_\lambda)$ stands for the support of the wavelet ψ_λ , i.e., the closure of the set of points x such that $\psi_\lambda(x) \neq 0$.

The wavelet structure function $W_f^2(p, j)$ is defined for $p \in \mathbb{R}$ by

$$W_f^2(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j}^* (e_\lambda)^p,$$

where the \sum^* means that the sum is taken on all λ' such that

$$\sup_{\text{supp}(\psi_{\lambda'}) \subset \text{supp}(\psi_\lambda)} |c_{\lambda'}| \neq 0.$$

The scaling function of f is defined by

$$\zeta_f^2(p) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(W_f^2(p, j))}{\log(2^{-j})} \right).$$

Second case: Wavelets in the Schwartz class

Definition 14. Leader based Multifractal Formalism 2: Let f be a uniform Hölder function. The wavelet structure function $W_f^3(p, j)$ is defined for $p \in \mathbb{R}$ by

$$W_f^3(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j} (d_\lambda)^p.$$

The scaling function of f is defined by

$$\zeta_f^3(p) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(W_f^3(p, j))}{\log(2^{-j})} \right).$$

Note that we can consider $\zeta_f^3(p)$ even if the wavelets do not belong to the Schwartz class.

In both cases, the same argument as above yields the following multifractal formalism based on the wavelet leaders

$$d_f(H) = \inf_{p \in \mathbb{R}} (d - \zeta_f(p) + Hp), \quad (5.3)$$

where $\zeta_f(p)$ stands either for $\zeta_f^2(p)$ or $\zeta_f^3(p)$ depending on the type of wavelet basis which is used.

Numerically, the determination of the scaling functions $\zeta_f^2(p)$ or $\zeta_f^3(p)$ requires the knowledge of the wavelet coefficients on more scales than the function $\zeta_f^1(p)$; indeed, in order to be trustable, the computation of a wavelet leader at a given scale requires the computation of the wavelet coefficients on several scales below. In the second case, the heuristic argument used in the derivation of the multifractal formalism is backed by mathematical results: It is proved in [29] that the scaling function $\zeta_f^3(p)$ is independent of the wavelet basis (in the Schwartz class) which is chosen and, if f is a uniform Hölder function, then

$$d_f(H) \leq \inf_{p \in \mathbb{R}} (d - \zeta_f^3(p) + Hp). \quad (5.4)$$

One pitfall of using (5.3) in applications is that, as mentioned already, the right-hand side of (5.3) is, by construction, a concave function. Since, in practice, using a Legendre transform of a scaling function is the only way to estimate numerically spectrums of singularities of signals, this may give the (perhaps erroneous) feeling that all spectrums of singularities of signals are concave functions, and therefore that mathematical models that yield concave spectrums are the only relevant ones (this remark also applies to all the variants of the multifractal formalisms that were mentioned above). Let us just mention at this point very simple models of random wavelet series with wavelet coefficients correlated through a Markov chain on the dyadic tree; such models have been proposed to model signals and images; however, they have recently been shown to yield non concave spectrums, see [19] and references therein.

5.3. Robustness for Wavelet-Based Quantities

Let us be more specific concerning the requirement of independence of the wavelet basis, since it is related to our previous discussion on robustness criteria. The scaling functions ζ_f^2 and ζ_f^3 are defined by conditions on the wavelet coefficients. Since the left-hand side of (5.3) is defined independently of any wavelet basis, the multifractal formalism will have no chance to hold if the scaling function depends on the wavelet basis chosen. Naïvely, in order to check this independence, one should first dispose of a description of all possible wavelet bases, which is not realistic. In practice, one checks a stronger (but simpler) requirement which implies that the scaling function considered has some additional stability; indeed, the matrix of the operator which maps an orthonormal wavelet basis onto another orthonormal wavelet basis is invariant under the action of infinite matrices which belong to algebras \mathcal{M}^γ that are defined below; therefore, one can check that the scaling function is also invariant under this action, which is the purpose of Corollary 1 and Proposition 6.

Definition 15. Algebras \mathcal{M}^γ : Let $\gamma > 0$; an infinite matrix $A(\lambda, \lambda')$ indexed by the dyadic cubes belongs to \mathcal{M}^γ if

$$|A(\lambda, \lambda')| \leq \frac{C 2^{-(\frac{d}{2} + \gamma)(j - j')}}{(1 + (j - j')^2)(1 + 2^{\inf(j, j')} \text{dist}(\lambda, \lambda'))^{d + \gamma}}.$$

Matrices of operators which map a smooth wavelet basis onto another one belong to these algebras. It is proved in [49] that the matrix which maps an r -smooth wavelet basis onto another r -smooth wavelet basis belongs to \mathcal{M}^γ for any $\gamma < r$, and that the spaces \mathcal{M}^γ are algebras. More generally, matrices (on wavelet bases) of pseudodifferential operators of order 0, such as the Hilbert transform in dimension 1, or the Riesz transforms in higher dimensions, belong to these algebras (for any $\gamma > 0$ if the wavelets are C^∞). We denote by $\mathcal{Op}(\mathcal{M}^\gamma)$ the space of operators whose matrix on a wavelet basis belongs to \mathcal{M}^γ . The following result is proved in [29].

Proposition 5. Let $p > 0$ and $A \in \mathcal{Op}(\mathcal{M}^\gamma)$ for a $\gamma > 0$. If $\zeta_f^3(p) < p\gamma$, then

$$\zeta_{A(f)}^3(p) \geq \zeta_f^3(p).$$

Applying this proposition to the operator that maps an r -smooth wavelet basis onto another r -smooth wavelet basis, and also to the inverse of this operator, shows that, under the hypotheses of Proposition 5, the scaling function $\zeta_f^3(p)$ is independent of the wavelet basis.

Another important remark is that $\zeta_f^2(p)$ and $\zeta_f^3(p)$ clearly coincide as long as $p > 0$ and $\zeta_f^3(p) < pr$. This follows from the fact, that by definition of d_λ and e_λ , one has,

$$W_f^3(p, j) \leq W_f^2(p, j) \leq 3^d W_f^3(p, j).$$

Thus the following result holds.

Corollary 1. Assume that the wavelet basis used is r -smooth; if $p > 0$ and $\zeta_f^3(p) < pr$, then the scaling function $\zeta_f^3(p)$ is independent of the wavelet basis used and $\zeta_f^2(p) = \zeta_f^3(p)$.

Note that, if the wavelets belong to the Schwartz class, then the previous result holds on the whole range $p > 0$.

Unfortunately, the case $p < 0$ leads to strongly different conclusions (and therefore justifies the introduction of two different scaling functions). In order to state the results in that case, we will need here a different requirement than the one used in Proposition 5.

Definition 16. Quasidiagonal infinite matrix: An infinite matrix $A(\lambda, \lambda')$ is quasidiagonal if A is invertible, and if A and A^{-1} belong to \mathcal{M}^γ for any $\gamma > 0$.

Let $C = \{c_\lambda\}_{\lambda \in \Lambda}$ be a collection of coefficients indexed by the dyadic cubes. A property \mathcal{P} is **robust** if the following condition holds: If $\mathcal{P}(C)$ holds then, for any quasidiagonal operator \mathbf{M} , $\mathcal{P}(\mathbf{M}C)$ holds.

The matrix of an operator which maps a wavelet basis in the Schwartz class onto another one is quasidiagonal, see [49]. Therefore, in order to check that a condition defined on the wavelet coefficients is independent of the wavelet basis (in the Schwartz class) used, one can check the stronger property that it is invariant under the action of quasidiagonal matrices. The following result is proved in [29].

Proposition 6. If $p < 0$, then $\zeta_f^3(p)$ is independent of the wavelet basis in the Schwartz class which is used.

5.4. Illustrations and Examples

5.4.1. Fractional Brownian Motion. The following result shows that both multifractal formalisms based on wavelet leaders yield the correct spectrum of singularities for F.B.M.

Theorem 5.1. Let $\beta \in (0, 1)$ and $B_\beta(t)$ be a generic sample path of F.B.M. of order β . Assume that the wavelet used belongs to the Schwartz class, then, with probability 1,

$$\forall p \in \mathbb{R}, \quad \zeta_{B_\beta}^3(p) = \beta p \quad (5.5)$$

and the limits in the definitions of the scaling functions are true limits; the wavelet leaders based multifractal formalism (5.3) yields the correct spectrum (3.10).

Proof of Theorem 5.1: First, we note that the previous robustness results of Section 5.3 for $\zeta_f^3(p)$ imply that the results do not depend of the wavelet basis in the Schwartz class which is used, and, in particular, we can use the biorthogonal basis generated by the wavelets (4.8) which leads to the decomposition (4.9). Then

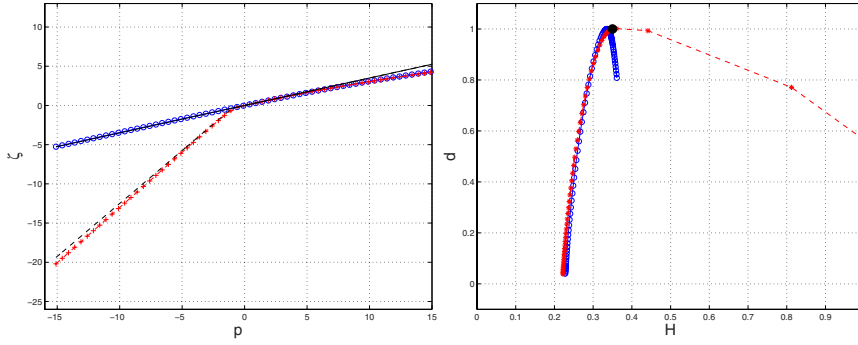


FIGURE 2. **Fractional Brownian Motion.** From a single sample path of fractional Brownian motion synthesized numerically (using the Circulant embedding Matrix technique[8]) with $\beta = 0.35$ (number of sampling points: 2^{18}), one obtains: Left, solid (black) line: theoretical $\zeta_f(p)$, dashed (black) line corresponds to Eq. 3.8, solid (blue) line with 'o': $\zeta_f^3(p)$, dashed (red) line with '+': $\zeta_f^1(p)$. Right, large full (black) dot, theoretical $d(H)$, solid (blue) line with 'o': $d^3(H)$, dashed (red) line with '+': $d^1(H)$. While the wavelet based and leader based formalisms both yield the correct $\zeta_f(p)$ s for positive p s, the leader based one only is able to correctly measure the $\zeta_f(p)$ s for negative p s. The corresponding Legendre transform (solid (blue) line with 'o':) concentrates around the theoretical $d(H)$. Its extension around the correct value gives us an idea of the accuracy of the numerical procedure.

$$\begin{aligned}
 \mathbb{P}(d_\lambda \leq j^{-4\beta} 2^{-\beta j}) &= \prod_{\lambda' \subset 3\lambda} \mathbb{P}(|c_{\lambda'}| \leq j^{-4\beta} 2^{-\beta j}) \\
 &= \prod_{\lambda' \subset 3\lambda} \mathbb{P}(2^{-\beta j'} |\chi_{\lambda'}| \leq j^{-4\beta} 2^{-\beta j}) \\
 &\leq \prod_{\lambda' \subset 3\lambda} j^{-4\beta} 2^{\beta(j'-j)}.
 \end{aligned}$$

We pick the scale $j' = j + \left\lceil \frac{2 \log j}{\log 2} \right\rceil + 1$, and we note that the number of subintervals of scale j' which are subintervals of λ is larger than j^2 , so that

$$\mathbb{P}(d_\lambda \leq j^{-4\beta} 2^{-\beta j}) \leq \left(j^{-4\beta} 2^{\beta(j'-j)} \right)^{j^2},$$

and one easily checks that this quantity is bounded by e^{-j^2} for j large enough. Since $\sum_j \sum_{k=0}^{2^j} e^{-j^2}$ is finite, the Borel-Cantelli lemma implies that for j large enough, all the d_λ are larger than $j^{-4\beta} 2^{-\beta j}$. On the other hand, we already saw that, with probability one, for j large enough, all the $|\chi_\lambda|$ indexed by a dyadic subinterval of $[0, 1]$ are bounded by j , and (5.5) follows from these two estimates.

Fig. 2 compares the wavelet and leader based multifractal formalisms practically applied to a sample path of fractional Brownian motion produced numerically using the so called *Circulant embedding Matrix* synthesis procedure [8]. One clearly sees that the wavelet formalism cannot reach the negative p part of $\zeta(p)$ and hence fails to measure correctly $d(H)$, while the the leader based formalism accurately analyzes both $\zeta(p)$ and $d(H)$. Moreover, it is interesting to note that the wavelet based formalism follows for negative ps the prediction derived from Eq. 4.10 in Proposition 3 (dashed black line).

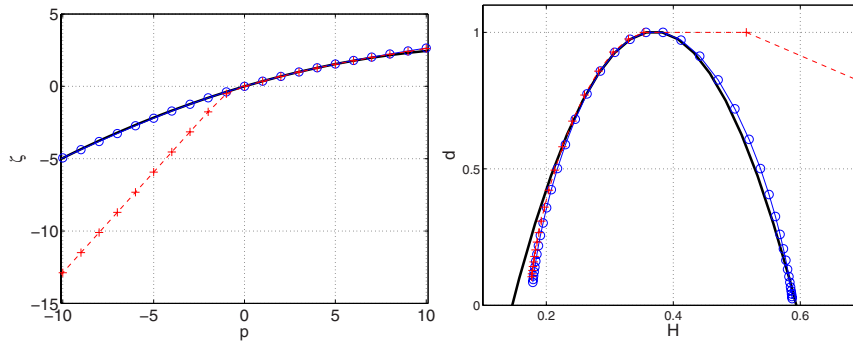


FIGURE 3. **Multiplicative Random Wavelet Cascades.** From (an average of 500 realizations of) a log-normal RWC produced numerically (number of samples: 2^{17}), one obtains: Left, solid (black) line: theoretical $\zeta_f(p)$, solid (blue) line with ‘o’: $\zeta_f^3(p)$, dashed (red) line with ‘+’: $\zeta_f^1(p)$. Right, solid (black) line, theoretical $d(H)$, solid (blue) line with ‘o’: $d^3(H)$, dashed (red) line with ‘+’: $d^1(H)$. While the wavelet based and leader based formalisms both yield the correct $\zeta_f(p)$ s for positive ps , the leader based one only is able to correctly measure the $\zeta_f(p)$ s for negative ps . The corresponding Legendre transforms yield correct measure of $d(H)$ for the lowest hs while only the leader based approach accurately measures the largest hs .

5.4.2. Multiplicative Cascades. The second example is based on random multiplicative cascades (or martingales). Instead of the celebrated cascades construction developed by Mandelbrot [45] and studied theoretically by Kahane and Peyrière

[36] that produce multifractal measures, we chose to illustrate the multifractal formalisms on multiplicative random wavelet cascades (RWC), introduced by Arneodo et al., as they provide us with well defined synthetic multifractal functions or processes (cf. [7]).

RWC are defined through their wavelet coefficient expansion on an orthonormal wavelet basis as:

$$f(x) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} d_f(j, k) \psi(2^j x - k).$$

Following original constructions, the wavelet coefficients $d_f(j, k)$ entering the definition of RWCs are obtained as a product of (positive) multipliers $W_{j,k}$, which consist of mean one independent and identically distributed random variables:

$$d_{rwc}(j, k) = z_{j,k} \prod_{j'=1..j, k'/\lambda(j,k) \subset \lambda(j',k')} W_{j',k'}.$$

The $z_{j,k}$ are random variables taking value $+1$ or -1 with equal probability and ensuring that the wavelet coefficients are randomly chosen positive or negative.

It is known that such constructions yield multifractal processes whose $\zeta(p)$ and hence $d_f(H)$ are entirely determined from the function $-\log_2 \mathbb{E} W^p$ (see [7] for details). For instance, one commonly chose log-normal multipliers, i.e., $-\log_2 \mathbb{E} W^p = mp - \sigma^2 \ln 2 p^2 / 2$, m and σ being two parameters to be chosen.

In Fig. 3, the wavelet based and leader based multifractal formalisms are compared using 500 synthetic realizations of sample paths (number of samples: 2^{17}) of a log-normal RWC. One clearly sees that the wavelet based multifractal formalism fails to measure $\zeta(p)$ for negative p s and $d(H)$ for the largest H s, while the leader based multifractal formalism produces a correct analysis over the entire spectrum.

5.4.3. Two-Dimensional Multiplicative Mandelbrot's Cascades. The third example aims at showing the leader based multifractal formalism at work in higher dimension. We chose to use here 2-dimensional (log normal) multiplicative Mandelbrot's cascades, whose standard definition not recalled here can be found in [45] or e.g., [4, 7, 39]. The corresponding measure is then (fractionally) integrated to produce a 2D function [4]. Fig. 4 compares the wavelet based and leader multifractal formalisms applied to this 2D function. Fig. 4 is obtained from a log normal cascade, with fractional integration of order $1/2$, number of sampling points $= 2^{10} \times 2^{10}$, see [40] for details on the synthesis procedure. Again, the wavelet based multifractal formalism yields an incorrect determination of the scaling exponents for negative p s and of $D(h)$ for its upper (or right) part while the leader based one produces a relevant measure over the entire spectrum. This validates the theoretical and practical straightforward extension of the leader based multifractal formalism to higher dimensions.

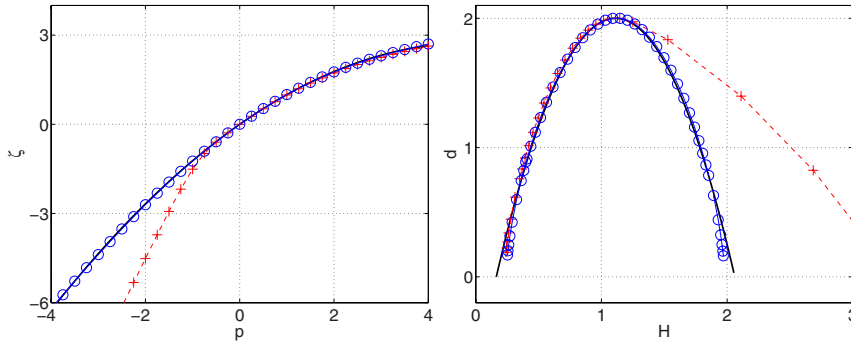


FIGURE 4. **Two-Dimensional Multiplicative Random Mandelbrot Cascades.** From (500 realizations of) a two-dimensional log-normal (fractionally integrated) Mandelbrot's binomial multiplicative cascade produced numerically (number of samples: $2^{10} \times 2^{10}$), one obtains: Left, solid (black) line: theoretical $\zeta_f(p)$, solid (blue) line with 'o': $\zeta_f^3(p)$, dashed (red) line with '+': $\zeta_f^1(p)$. Right, solid (black) line, theoretical $d(H)$, solid (blue) line with 'o': $d^3(H)$, dashed (red) line with '+': $d^1(H)$. While the wavelet based and leader based formalisms both yield the correct $\zeta_f(p)$ s for positive p s, the leader based one only is able to correctly measure the $\zeta_f(p)$ s for negative p s. The corresponding Legendre transforms yield correct measure of $d(H)$ for the lowest h s while only the leader based approach accurately measures the largest h s. This illustrates that the leader based formalism works efficiently and easily in higher dimensions.

5.5. Further Comments, Analysis and Synthesis Routines

At this stage, a number of comments are in order:

In the numerical examples presented here, we have implemented the formalism corresponding to Definition 14, with Daubechies wavelets (i.e., with wavelets that do not belong to the Schwartz class). The numerical results above show that this theoretical requirement can probably be weakened. Moreover in the present numerical implementation, digitalization has two major practical impacts. It implies a finite number of vanishing moments for the mother wavelet so that its belonging to the Schwartz class remains at a theoretical level. The same holds for the theoretical possible choice of a C^∞ mother wavelet. This is under further current investigations.

In numerous papers more focused on practical multifractal analysis (see for instance [3, 39, 40]), the convention $a = 2^j$ is preferred to $a = 2^{-j}$ chosen in the present text. This implies that the limit in the equations defining the ζ_f are taken for $j \rightarrow -\infty$.

All the procedures used in the present work to synthesize processes and signals and to implement multifractal formalism analysis were developed by ourselves¹ in MATLAB or C.

5.6. Practical and Numerical Multifractal Analysis: Comparisons Against Other Multifractal Formalisms and Against the Wavelet Transform Modulus Maxima Approach

5.6.1. Practical and Historical Implementations of Multifractal Formalisms. Because multifractal analysis was first applied to characterize strange attractors in the field of chaos (see e.g., [23]) and dissipation field in hydrodynamic turbulence (see e.g., [58, 48]), the earliest formalism actually used in applications was based on the computation of structure functions based on measures:

$$\forall p \in \mathbb{R}, \Sigma_\mu(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j} \mu(\lambda)^p,$$

a formula that closely resembles that of Definition 6 proposed here.

In hydrodynamic turbulence, one is not only interested in dissipation fields, but also in velocity ones, i.e., in functions. This is why Parisi and Frisch [55] proposed to define a formalism based on the increments $f((k+1)/2^j) - f(k/2^j)$ of the function f under analysis:

$$p > 0, \Sigma_f^1(p, j) = 2^{-j} \sum_k \left| f\left(\frac{k+1}{2^j}\right) - f\left(\frac{k}{2^j}\right) \right|^p$$

Immediately after they appeared, wavelet were read as generalizations both for box-aggregation and increments. For instance, the increments are commonly referred to as the *poor man's wavelet* and the historical Haar wavelet can be seen as a difference of averaged (or aggregated) quantities (see e.g., [34, 50]). Therefore, wavelets act as increments of higher orders and hence generalize the usual increments. Moreover, multiplicative cascades have been used as a standard for the synthesis of multifractal measure [45]. Box aggregation yield a correct multifractal spectrum only for the special class of conservative cascades (see e.g., [37]). This is why both continuous and discrete wavelet transforms have been involved in multifractal analysis since their earlier times, mainly to study turbulence velocity and dissipation fields (see e.g., [47, 13, 4]).

However, it has immediately been observed that most of the early-proposed multifractal formalisms failed to work for negative values of p , a major drawback as the analysis of the full multifractal spectrum theoretically involves the use of both positive and negative ps . To overcome this difficulty, Arneodo and co-authors introduced the use of the *wavelet transform modulus maxima* method (**WTMM**). To date, it remains one of the most widely used tool for empirical multifractal analysis performed in actual applications. The wavelet leader multifractal formalism

¹The authors wish to thank Stéphane Roux, Physics Lab., ENS de Lyon, for having made available to them his codes implementing the wavelet transform modulus maxima technique

proposed here provides us with a new, relevant and efficient multifractal analysis framework.

In the section below, we briefly describe the **WTMM** tool and propose elements of comparisons between the two approaches with no aim to cover a full and detailed analysis of the difference between them.

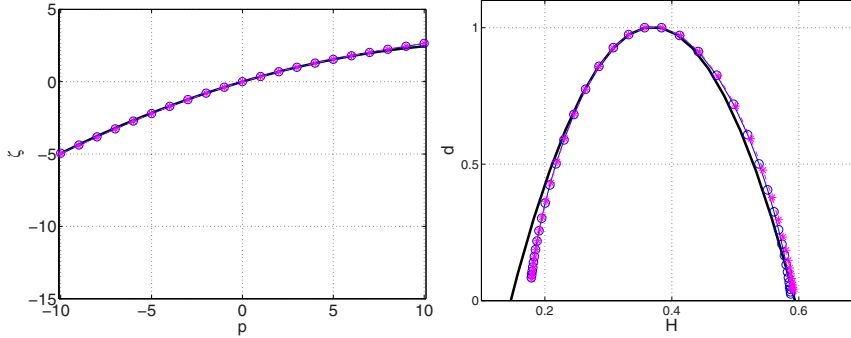


FIGURE 5. **leaders vs WTMM multifractal formalisms.** On the same set of synthetic data as the one used to obtain Fig. 3 (Multiplicative Random Wavelet Cascades), ones obtains: Left, solid (black) line: theoretical $\zeta_f(p)$, solid (blue) line with 'o': $\zeta_f^3(p)$, mixed (magenta) line with '*': $\zeta_f(p)$, produced by the **WTMM** approach. Right, solid (black) line, theoretical $d(H)$, solid (blue) line with 'o': $d^3(H)$, mixed (magenta) line with '*': $d^1(H)$, produced by the **WTMM** approach. Both formalisms are yielding very close and equivalent results at the price, however, of a very different computational cost though.

5.6.2. Wavelet Transform Modulus Maxima. The use of d_λ is reminiscent of the **WTMM** initially introduced by S. Mallat in [44] and developed by A. Arneodo E. Bacry and J.-F. Muzy in the context of multifractal analysis, see [4, 6] and references therein: Assume that ψ is a wavelet, i.e. a well localized function with enough vanishing moments (in practice a derivative, or a second derivative of a Gaussian is often used). One computes the continuous wavelet transform of f

$$C_f(a, b) = a \int f(x) \psi \left(\frac{x - b}{a} \right) dx$$

which is a function defined in the upper half plane $\{(a, b) : a > 0, b \in \mathbb{R}\}$. For each scale a , one spots the local maxima of the functions $b \rightarrow C_f(a, b)$. These local maxima are connected through scales, thus yielding the *wavelet skeleton*. At each local maximum located at position (a, b) in the time scale plane, one associates the supremum of the wavelet transform on the sub-skeleton issued from (a, b) (i.e.

the maximum on the part of the skeleton which is linked to (a, b) and corresponds to values of the scale parameter smaller than a). The $\zeta(p)$ are then obtained using formulas such as those in Definitions 13 and 14, the partition function being computed only with the supremum skeleton values (see, e.g., [4, 5, 6] for details).

Practical results obtained with the **WTMM** approach are illustrated in Fig. 5 (on the same set of synthetic data as the one used to produce Fig. 3) and compared to those produced with the leader based multifractal formalism. One sees that both approaches yield equivalent results. Their merits are further compared below. Historically, the **WTMM** has been the first and remained for a long time the only multifractal formalism yielding correct results for negative ps . Also, it enabled the first attempts to analyze chirp type singularities [5]. However, a number of important differences between the wavelet leaders and **WTMM** approaches can be pointed out.

From a mathematical point of view, the main differences are the following: The wavelet leader based multifractal formalism now benefits of well established theoretical mathematical results as described in previous sections. The situation is much different for the case of the **WTMM**. In the wavelet maxima method, the spacing between the local maxima need not be of the order of magnitude of the scale a or even be regularly spaced; therefore, the scaling function thus obtained can be different from $\zeta_f^3(p)$ (see [27] where counterexamples are constructed). It follows that, up to now, no mathematical results have been proved to hold for the wavelet maxima method. For instance, theoretical results such as the independence of the scaling function with the analyzing wavelet, or the fact that the Legendre transform of the scaling function yields an upper bound for the spectrum of singularities, are not available so far. This is because, as seen before, operators that map a wavelet basis on another one belong to classes of infinite matrices which are easy to describe. On the opposite, a wavelet transform belongs to a specific subspace of $L^2(dadb/a^2)$: The so-called “*Reproducing Kernel Hilbert Spaces*”, which depend on the wavelet, see [22]. Therefore describing specific classes of operators that act on these spaces is much more difficult to handle.

On the computational side, an important drawback of the **WTMM** lies in its computational cost. It is based on the computation of a full continuous wavelet transform followed by the skeletization and maxima tracking procedures. The leaders approach is based on the coefficients on an orthogonal wavelet decomposition and hence benefits from fast decomposition algorithms (cf. [43]). It implies that the wavelet leaders approach can be used for signals of arbitrary length while the **WTMM** is often restricted to much shorter ones.

Along the same line, while the wavelet leaders approach is straightforwardly and without extra difficulties extended to arbitrary higher dimensions (cf. Fig. 4), the definition of the **WTMM** needs to be significantly modified to a more complex procedure before extension to higher dimensions. Those further complications strongly impair the mathematical analysis of the method and substantially increase the corresponding computational cost (see e.g., [37] for discussions).

The statistical performance of the estimators for the $\zeta(p)$ exponents based on these two different approaches are being investigated and compared (see for instance [57]).

Numerical results regarding the leader based analysis of processes containing chirp-type singularities are been proposed in [3, 40] and show that the wavelet leader based formalism correctly measures the corresponding multifractal spectra. This will be further developed in forthcoming works.

6. The Weak-Scaling Exponent

In this section, we investigate whether we can expect (4.5) to yield the spectrum associated with some alternative pointwise exponent. We will show that, though the scaling function $\zeta_f^1(p)$ may depend on the wavelet basis chosen if $p < 0$, nonetheless it is independent of the wavelet basis for $p > 0$ and, when the infimum in (4.5) is reached for $p > 0$, then (4.5) is expected to yield the spectrum of singularities based on the *weak scaling exponent*, which was introduced by Y. Meyer in [50]. This weak scaling exponent coincides with the Hölder exponent in the case of cusp-like singularities, and this will explain why the multifractal formalism based on (4.5) yields the correct increasing part of the spectrum for signals such as Brownian motion. This interpretation will thus allow us to give a new interpretation to the computations done in previous papers which were based on (4.5).

6.1. Characterizations of the Weak Scaling Exponent

Before giving a precise definition, let us first give a feeling on the nature of the information supplied by the weak-scaling exponent. The weak-scaling exponent was introduced as a substitute for the Hölder exponent, which displays a better behaviour under integration: Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function, and denote by $f^{(-1)}$ a primitive of f . It may happen that $h_{f^{(-1)}}(x_0) \neq h_f(x_0) + 1$ as might be expected. A typical example where $h_{f^{(-1)}}(x_0)$ is strictly larger than $h_f(x_0) + 1$ is supplied by the chirp $C_{\alpha,\beta}$ defined in (4.6), when $\beta > 0$; indeed its Hölder exponent at x_0 is α and its Hölder exponent is increased by $1 + \beta$ after one primitivation, as shown immediately by writing

$$C_{\alpha,\beta}^{(-1)}(x) = \int_{x_0}^x \left(\frac{|t - x_0|^{\alpha+\beta+1}}{\beta} \right) \beta |t - x_0|^{-\beta-1} \sin \left(\frac{1}{|t - x_0|^\beta} \right) dt$$

and integrating by parts.

This phenomenon is the source of many difficulties and, in particular, it is one of the causes of failure of the multifractal formalism based on wavelet coefficients, see [5]. Indeed, as already mentioned, counterexamples to the heuristic which led to (4.5) are supplied by functions such as $C_{\alpha,\beta}$. Therefore it is natural to wonder if there is an alternative exponent endowed with the additional property that the exponent of a primitive is just the exponent of the function shifted by 1, and that would be “close” to the Hölder exponent (for instance, they would coincide for cusps such as $|x - x_0|^\alpha$). This is precisely the *weak-scaling exponent* $h_f^{ws}(x_0)$

which has been discovered by Y. Meyer, see [50], and can be characterized by the following properties:

- $h_f^{ws}(x_0) \geq h_f(x_0)$.
- $h_{f^{(-1)}}^{ws}(x_0) = h_f^{ws}(x_0) + 1$
- $h_f^{ws}(x_0)$ is the smallest exponent satisfying the two previous conditions.

These three requirements are not easy to use directly in order to compute the weak-scaling exponent of a function; in practice, one uses a characterization on the wavelet coefficients supplied by Theorem 1.2 of [50].

First, we have to introduce the $\Gamma^s(x_0)$ smoothness criterium.

Definition 17. Let f be a tempered distribution; f belongs to $\Gamma^s(x_0)$ if and only if there exists $s' > 0$ such that f belongs to the two-microlocal space $C^{s,-s'}(x_0)$, which means that the wavelet coefficients of f (taken on a wavelet basis in the Schwartz class) satisfy

$$|c_{j,k}| \leq C 2^{-sj} (1 + |2^j x_0 - k|)^{s'}. \quad (6.1)$$

Note that (4.2) already consisted of a two-microlocal condition.

Definition 18. Let f be a tempered distribution; the weak-scaling exponent of f is

$$h_f^{ws}(x_0) = \sup\{s : f \in \Gamma^s(x_0)\}. \quad (6.2)$$

This definition is independent of the wavelet basis chosen and it coincides with the informal definition given above, see [50].

Definition 19. Weak scaling exponent spectrum: We denote by $E_f^{ws}(H)$ the set of points where the weak-scaling exponent of a distribution f takes the value H . The weak-scaling spectrum of f (denoted by $d_f^{ws}(H)$) is the Hausdorff dimension of $E_f^{ws}(H)$.

In order to derive the multifractal formalism for the weak scaling exponent, the following alternative characterization will be useful (Note that it slightly differs from the wavelet characterization obtained in [32]).

Definition 20. ϵ -leader: Let $\epsilon > 0$. The ϵ -cone of scale j above x_0 is

$$C_j^\epsilon(x_0) = \{\lambda' \subset 3\lambda_j(x_0) \text{ such that } j' \leq (1 + \epsilon)j\}.$$

The ϵ -leader of scale j above x_0 is

$$d_j^\epsilon(x_0) = \sup_{\lambda' \in C_j^\epsilon(x_0)} |c_{\lambda'}|.$$

It is worth noting that the limit $\epsilon \rightarrow +\infty$ corresponds to the definition of the wavelet leaders (cf. Definition 12) while in the limit $\epsilon \rightarrow 0$ one recovers the usual wavelet coefficients.

Proposition 7. Let f be a tempered distribution. The weak scaling exponent of f at x_0 is the supremum of the values of H satisfying

$$\forall \epsilon > 0, \exists J \forall j \geq J \quad d_j^\epsilon(x_0) \leq C 2^{-(H-\epsilon)j}. \quad (6.3)$$

Proof of Proposition 7: Suppose that there exists $s' > 0$ such that (6.1) holds. Let $\epsilon > 0$; then

$$d_j^\epsilon(x_0) \leq \sup_{\lambda' \in C_j^\epsilon(x_0)} 2^{-sj'} (1 + |2^{j'}x_0 - k'|)^{s'} \leq 2^{-sj} 2^{(j'-j)s'} \leq 2^{-sj} 2^{s'\epsilon j}.$$

Since ϵ can be chosen arbitrarily small, (6.3) holds for $H = s$.

Conversely, suppose that (6.3) holds. Since f is a finite order distribution, it follows that

$$\exists u \in \mathbb{R}, \exists C > 0, \forall j, k \quad |c_{j,k}| \leq C.2^{-uj}. \quad (6.4)$$

We can of course assume that u satisfies $H - 1 - u > 0$.

First, if λ' belongs to an ϵ -cone of scale j above x_0 , then

$$j \geq j'/(1 + \epsilon)$$

so that

$$|c_{\lambda'}| \leq C.2^{-j'(H-\epsilon)/(1+\epsilon)}$$

and (6.3) holds for an s arbitrarily close to H and $s' = 0$.

Else, if λ' does not belong to an ϵ -cone of scale j above x_0 then, in particular, it is outside the ϵ -cone of scale $= j'/(1 + \epsilon)$ above x_0 , so that

$$|2^{j'}x_0 - k'| \geq 2^{j'}2^{-j} \geq 2^{j'\epsilon/(1+\epsilon)}.$$

It follows that

$$|c_{\lambda'}| \leq C.2^{-uj'} \leq C.2^{-Hj'} 2^{-(u-H)j'} \leq C.2^{-Hj'} |2^{j'}x_0 - k'|^{(H-u)(1+\epsilon)/\epsilon},$$

and (6.3) holds for $s = H$ and $s' = (H - u)(1 + \epsilon)/\epsilon$.

6.2. Multifractal Formalism for the Weak Scaling Exponent

Proposition 7 states that the weak scaling exponent at x_0 is given by the order of magnitude of the ϵ -leaders above x_0 . Therefore, the following structure and scaling functions are naturally associated with the weak scaling exponent.

Definition 21. Weak scaling exponent Multifractal formalism: Let f be a temperate distribution and assume that the wavelets used belong to the Schwartz class. If $p \in \mathbb{R}$, let

$$d_\lambda^\epsilon = \sup_{\lambda' \subset 3\lambda, j' \leq (1+\epsilon)j} |c_{\lambda'}|,$$

$$W_f^4(p, \epsilon, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j} (d_\lambda^\epsilon)^p,$$

and

$$\zeta_f^4(p, \epsilon) = \liminf_{j \rightarrow +\infty} \left(\frac{\log(W_f^4(p, \epsilon, j))}{\log(2^{-j})} \right). \quad (6.5)$$

The weak scaling function of f is defined by

$$\zeta_f^4(p) = \lim_{\epsilon \rightarrow 0} \zeta_f^4(p, \epsilon). \quad (6.6)$$

Note that, when $\epsilon \rightarrow 0$, the ϵ -leaders are defined by a supremum over a decreasing set, and therefore, they decrease; it follows that, if $p > 0$, then $W_f^4(p, \epsilon, j)$ decreases when $\epsilon \rightarrow 0$, and, if $p < 0$, then $W_f^4(p, \epsilon, j)$ increases when $\epsilon \rightarrow 0$, which implies that the limit exists in (6.6) in all cases.

The same arguments as for the derivation of the previous multifractal formalisms lead to the following multifractal formalism for the weak scaling exponent:

$$d_f^{ws}(H) = \inf_{p \in \mathbb{R}} (d - \zeta_f^4(p) + Hp) \quad (6.7)$$

The following result shows that the Legendre transform of the scaling function yields an upper bound for the weak-scaling spectrum without any uniform regularity assumption, see [32].

Theorem 6.1. *Let f be a tempered distribution. Then its weak scaling spectrum satisfies*

$$d_f^{ws}(H) \leq \inf_{p \in \mathbb{R}} (Hp - \zeta_f^4(p) + d). \quad (6.8)$$

Let us now study more precisely the scaling function $\zeta_f^4(p)$.

Proposition 8. Let $p > 0$. If the wavelets are r -smooth with $r > p\zeta_f^1(p)$, then

$$\zeta_f^4(p) = \zeta_f^1(p).$$

This result implies that the increasing part of the Legendre transforms in (4.5) and (6.8) coincide. This is important in practice since $\zeta_f^1(p)$ is much easier to obtain numerically (because it is not defined as a double limit). Furthermore Proposition 8 shows that the multifractal formalism given by (4.5) is expected to yield the increasing part of the weak scaling spectrum, as announced.

Proof of Proposition 8: Let $p > 0$; since $(d_\lambda^\epsilon)^p \geq |c_{\lambda'}|^p$, it follows that $\zeta_f^4(p) \leq \zeta_f^1(p)$. Conversely,

$$(d_\lambda^\epsilon)^p \leq \sum_{\lambda' \subset 3\lambda, j' \leq (1+\epsilon)j} |c_{\lambda'}|^p$$

so that

$$w_f^4(p, \epsilon, j) \leq w_f^1(p, \epsilon, j) + 2^d w_f^1(p, \epsilon, j+1) + \cdots + 2^{d\epsilon j} w_f^1(p, \epsilon, (1+\epsilon)j);$$

therefore

$$\zeta_f^4(p, \epsilon) \geq (d+1)\epsilon + \zeta_f^1(p, \epsilon).$$

Since this is true $\forall \epsilon > 0$, it follows that $\zeta_f^4(p) \geq \zeta_f^1(p)$.

Let us now give a function space interpretation to $\zeta_f^1(p)$ (hence to $\zeta_f^4(p)$) when $p > 0$. Recall that f belongs to the homogeneous Besov space $\dot{B}_p^{s,\infty}(\mathbb{R}^d)$ if

$$\exists C, \forall j \quad 2^{(sp-d)j} \sum_{\lambda \in \Lambda_j} |c_\lambda|^p \leq C. \quad (6.9)$$

It follows that

$$\zeta_f^1(p) = \sup\{s : f \in \dot{B}_p^{s/p, \infty}(\mathbb{R}^d)\}.$$

Since (6.9) holds as soon as the wavelets are r -smooth with $s > r$, it follows that $\zeta_f^4(p)$ is independent of the r -smooth wavelet basis as soon as $r > p\zeta_f^4(p)$.

If $p < 0$, then $\zeta_f^4(p)$ cannot be given any more a function space interpretation; however, one can prove that it is still independent of the wavelet basis (in the Schwartz class), by using the same arguments as those developed in [29] in order to prove that $\zeta_f^3(p)$ is robust.

One may wonder if $\zeta_f^1(p)$ and $\zeta_f^4(p)$ still coincide for $p < 0$. The example of F.B.M., which we now consider, shows that it is not the case.

6.3. Examples: The weak Scaling Spectrum of Fractional Brownian Motions

First, let us determine the weak scaling exponent of F.B.M. at every point. We use the characterization supplied by Proposition 7, which is independent of the wavelet basis chosen (because it is equivalent to the two-microlocal characterization (6.1), which defines a robust condition, as shown in [26]). Furthermore, using again this robustness property, we can use the decomposition (4.9) on biorthogonal wavelets adapted to F.B.M.. Therefore the wavelet coefficients are $2^{-\beta j} \chi_{j,k}$ where the $\chi_{j,k}$ are I.I.D. standard centered Gaussians. Then, the proof of Theorem 5.1 shows that, for j large enough, the ϵ -leaders are larger than $j^{-4\beta} 2^{-\beta j}$ (because, in the proof, the supremum in the definition of the wavelet leaders is extracted only in the range of scales between j and $j + \left\lceil \frac{2 \log j}{\log 2} \right\rceil + 1$, which is smaller than $j + \epsilon j$ for j large enough). Therefore, the Borel-Cantelli lemma implies that a.s. for j large enough, all the d_λ^ϵ stand between $j^{-4\beta} 2^{-\beta j}$ and $j 2^{-\beta j}$. Thus, the following theorem holds.

Theorem 6.2. *Let $\beta \in (0, 1)$ and $B_\beta(t)$ be a generic sample path of F.B.M. of order β . Assume that the wavelet used is C^2 ; then, with probability 1, the weak scaling exponent of B_β is everywhere β ,*

$$\forall p \in \mathbb{R}, \quad \zeta_{B_\beta}^4(p) = \beta p \quad (6.10)$$

and the liminfs in the definitions of the scaling functions are true limits.

The multifractal formalism (5.3) yields the correct weak scaling spectrum of singularities of F.B.M..

One can note that by inverting the two limits $\epsilon \rightarrow 0$ and $j \rightarrow +\infty$ in Definition 21 and (6.5) and (6.6), one recovers the wavelet coefficient based multifractal formalism, which was shown in Section 4 to yield an incorrect measure of the multifractal spectrum of F.B.M. Hence, F.B.M. provides us with a pedagogical example to emphasize how much the order of the limits matters in multifractal analysis.

7. Conclusion

We conclude this paper by some comments concerning the comparison of the three wavelet-based multifractal formalisms given by (4.5), (5.3) and (6.7), why they may coincide or differ, and how this is related to the presence of “oscillating singularities”, as opposed to “cusp-singularities”.

Let us first discuss what is usually understood by these two types of singularities. As mentioned already, a typical example of a cusp at x_0 is supplied by the function $|x - x_0|^\alpha$ (where α is positive and is not an even integer so that the function is not C^∞ at x_0). This is usually opposed to chirps such as the functions $C_{\alpha,\beta}$ defined in (4.6). These particular examples do not supply us with a general mathematical definition; they can only give some clues in this direction. Unfortunately, there is no general agreement as to what should be the right definition of a chirp. Actually, several definitions have been proposed (see [5, 14, 28, 33, 52]) and simple models of random wavelet series have been shown to display such behaviors, see [9]. Furthermore, C. Melot and A. Fraysse showed that such oscillatory behaviors are not exceptional but “generic” among the functions which have a given scaling function $\zeta_f^1(p)$, see [21, 46]. We won’t compare the merits of these definitions here, but rather discuss the opposite point of view: When can one say that a function displays cusps, since this is the case where we expect the different multifractal formalisms to coincide. The clue for a possible answer lies again in the comparison between the properties of the cusps $|x - x_0|^\alpha$ and the chirps $C_{\alpha,\beta}$: Recall that the Hölder exponent of both functions at x_0 is α , but the weak scaling exponent of the cusp is α whereas it is $+\infty$ for the chirps; following this remark, Y. Meyer proposed in [50] the following general definition for a cusp.

Definition 22. Cusp singularity: Let f be a function which is bounded in a neighbourhood of x_0 , and such that $h_f(x_0) < \infty$; f has a cusp singularity at x_0 if

$$h_f(x_0) = h_f^{ws}(x_0).$$

Note that this definition does not involve the wavelet coefficients of f : however, in order to understand its implications in multifractal analysis, it is necessary to check what it implies for the wavelet leaders. The characterizations supplied by Propositions 4 and 7 show that, if f has a cusp singularity at x_0 , then we can expect that, for any $\epsilon > 0$, the supremum in the quantity $\sup_{\lambda' \subset 3\lambda_j(x_0)} |c_{\lambda'}|$ is actually reached for a λ' whose scale j' satisfies $j \leq j' \leq (1 + \epsilon)j$ (if j is large enough). A typical example of this behavior is supplied by F.B.M.: Indeed, in Section 5.4, we estimated these suprema and actually showed that a.s. every point is a cusp singularity (since everywhere the Hölder exponent and the weak scaling exponent both take the value β). We can wonder why the first multifractal formalism yields a different spectrum, see Proposition 3. A close inspection of the proof of this proposition compared with the proof of Theorem 5.1 shows that, though the suprema of wavelet coefficients on very small subtrees of the form

$$\{\lambda' \subset \lambda : j' \leq (1 + \epsilon)j\}$$

are of the order of magnitude of $2^{-\beta j}$ with a probability extremely close to 1, nonetheless, single wavelet coefficients have a Gaussian distribution and therefore can take very small values with a much larger probability, which becomes non negligible when one considers simultaneously a large number of coefficients (2^j in the present case).

Is it nonetheless possible that (4.5), (5.3) and (6.7) yield the same result? Since the scaling function $\zeta_f^1(p)$ is not robust, this can only be the consequence of a very particular choice of the wavelet basis. In practice, this only happens if an algorithm is used to define the coefficients of the function (or of the stochastic process) on a given wavelet basis, and *the same wavelet basis* is used also in the analysis procedure. Such models have been currently proposed, all of them verifying the following hierarchical property:

$$\lambda' \subset \lambda \implies |c_{\lambda'}| \leq |c_\lambda|. \quad (7.1)$$

Typical examples of wavelet series satisfying this property can be constructed starting with a probability measure μ defined on \mathbb{R}^d and picking, for an $\alpha \geq 0$ and $q > 0$,

$$c_\lambda = 2^{-\alpha j} [\mu(\lambda)]^q,$$

see [10, 7] and references therein. Because of (7.1), all wavelet multifractal formalisms trivially yield the same result but again, under the very artificial assumption that one picks the same synthesizing and analyzing wavelets, since (7.1) will not remain valid if the wavelets are changed. Note that, if the synthesizing and analyzing wavelets differ, then the last two multifractal formalisms will still coincide (because there exists a wavelet basis for which it is the case, and the corresponding scaling functions are robust); by contrast, numerical results show that the first multifractal formalism yields a different spectrum: The decreasing part of the Legendre transform is artificially raised, as in the F.B.M. case (see Proposition 3), because of very small wavelet coefficients whose influence in the structure function is not eliminated by taking suprema of coefficients, as in the wavelet leaders case. (Note that experimentalists are aware of this pitfall: Even if they study “artificial signals” defined through their wavelet coefficients, they take only for granted results which have been validated by using several different wavelet bases.)

8. Appendix: Proof of Theorem 3.2

Let us first prove (3.8). We start by two elementary remarks. If χ is a standard Gaussian, then

$$\mathbb{P}(|\chi| \geq j) \leq e^{-j^2/2};$$

it follows from the Borel-Cantelli Lemma that, with probability 1,

$$\exists J, \quad \forall j \geq J, \quad \forall k = 1, \dots, 2^j, \quad |\chi_{j,k}| \leq j,$$

hence,

$$\text{if } p > 0, \quad \text{then } A(p, j) \leq 2^{-pj/2} j^p. \quad (8.1)$$

On the other hand,

$$\mathbb{P}(|\xi| \leq j^{-2}2^{-j}) \leq \sqrt{\frac{2}{\pi}} j^{-2}2^{-j};$$

it follows that, for a given j , one of the 2^j Gaussians $\xi_{j,k}$ will be smaller than $j^{-2}2^{-j}$ with probability at most j^{-2} . Thus, by the Borel-Cantelli lemma, with probability 1,

$$\exists J, \quad \forall j \geq J, \quad \forall k \quad |\xi_{j,k}| \geq j^{-2}2^{-j}. \quad (8.2)$$

In order to obtain precise estimates on $A(j, p)$, we can use estimates for the distribution of the $|\xi_{j,k}|$ in the neighbourhood of 0; up to a smooth change of variable, it is the same as the distribution of 2^j I.I.D. random variables x_k distributed with the Lebesgue measure on $[0, 1]$ and therefore the estimate of $A(p, j)$ will be the same in both cases. Recall that the *empirical process* is defined as follows: Points $(x_n)_{n \in \mathbb{N}}$ are drawn independently with the Lebesgue measure on $[0, 1]$; therefore the estimates on $A(j, p)$ will be the same in both cases, up to a constant term. The empirical process is the collection of random processes

$$P_t^N = \sum_{n=1}^N 1_{[0, x_n)}(t).$$

Estimates on the joint distribution of the x_n are usually expressed in terms of

$$\alpha_t^N = \sqrt{N} \left(\frac{P_t^N}{N} - t \right)$$

which is the “correct” renormalization of the empirical process since it converges to a non-trivial limit (a Brownian bridge), see [18, 59] and references therein. The increments of the empirical process can be estimated using the following result which is a particular case of Lemma 2.4 of [59].

Lemma 8.1. *There exist two positive constants C'_1 and C'_2 such that, if $0 < l < 1/8$, $Nl \geq 1$ and $8 \leq A \leq C'_1 \sqrt{Nl}$,*

$$\mathbb{P} \left(\sup_{|t-s| \leq l} |\alpha_t^N - \alpha_s^N| > A\sqrt{l} \right) \leq \frac{C'_2}{l} e^{-A^2/64}. \quad (8.3)$$

Note that the condition

$$\sup_{|t-s| \leq l} |\alpha_t^N - \alpha_s^N| \leq A\sqrt{l}$$

implies that the number of points $(x_n)_{n \leq N}$ that fall in the interval $[s, t]$ differs from $N|t-s|$ by at most $A\sqrt{Nl}$.

In the following, we will use Lemma 8.1 repeatedly with intervals of different lengths and positions; however, each time, we will pick $N = 2^j$ and $A = j$; this last choice will allow use to obtain such small probabilities of the opposite events that we can apply the Borel-Cantelli lemma to their complement at the end. First we estimate the number of points that fall in the interval $[9/10, 1]$, so that $l = 1/10$; it follows that, with probability larger than $1 - Ce^{-j^2/64}$, this number differs from

$(1/10)2^j$ by at most $j2^{j/2}\sqrt{10}$. On one hand, if $p > 0$, then $A(p, j) \geq C(p)2^j$ and (3.8) for $p > 0$ follows from this estimate together with (8.1). On the other hand it follows that

$$\text{if } p < 0, \text{ then } A(p, j) \geq C2^j. \quad (8.4)$$

Let us now apply Lemma 8.1 on the interval $[0, j^3 2^{-j}]$, so that $l = j^3 2^{-j}$. With probability larger than $1 - Ce^{-j^2/64} 2^j / j^3$, the number of points in this interval is larger than $j^3 - j\sqrt{j^3 2^{-j} 2^j} \geq j^3/2$, so that

$$\text{if } p < 0, \text{ then } A(p, j) \geq \frac{1}{2} j^{3+3p} 2^{-pj}. \quad (8.5)$$

Let us now obtain upper bounds for $A(p, j)$ when $p < 0$. Let m be a fixed, large integer, and let

$$a_0 = 0, \text{ and } a_k = 2^{-(1-k/m)j} \text{ for } k = 1, \dots, m-1, \text{ and } l_k = a_k - a_{k-1}.$$

We first apply Lemma 8.1 in the first interval $[0, a_1]$. With probability larger than $1 - C2^j e^{-j^2/64}$, the number of points in this interval is bounded by $a_1 2^j + j2^{j/2} \sqrt{a_1}$, taking into account the value of a_1 and (8.2), it follows that the contribution to $A(p, j)$ of the points that fall in this interval is bounded by

$$2 \cdot 2^{j/m} j^{-2p} 2^{-pj}. \quad (8.6)$$

We now apply Lemma 8.1 on the remaining intervals. We obtain that $\forall k = 1, \dots, m-1$, with probability larger than $C'_2 2^j e^{-j^2/64}$ the number of points that fall in the interval $[a_{k-1}, a_k)$ differs from $l_k 2^j$ by at most $j2^{j/2} \sqrt{l_k}$. It follows that, if $p > 0$ then, with probability at most $C'_2 m 2^j e^{-j^2/64}$ the contribution of each interval $[a_{k-1}, a_k)$ to $A(p, j)$ is bounded by

$$2^{-j} l_k 2^j a_k^p + j2^{j/2} \sqrt{l_k} a_{k-1}^p$$

which is bounded by

$$2^{-pj} \left(2^{(p+1)jk/m} + 2^{(p+1/2)jk/m} 2^{-pj/m} \right).$$

If $p < -1$ then each of these terms is bounded by (8.6), and if $-1 \leq p < 0$ then each is bounded by $C2^j$; therefore (3.8) follows in all cases for $p < 0$.

We now prove the second part of Theorem 3.2, i.e. that, in the case of a sample path of Brownian motion, a.s.

$$\forall p \in \mathbb{R}, \quad \eta_B^2 = p/2, \quad (8.7)$$

and that the \liminf in (3.5) is a true limit, which will imply the second assertion of Theorem 3.2.

Let $I_{j,k}$ denote the interval $[k2^{-j}, (k+1)2^{-j})$. The oscillations

$$Osc_B(I_{j,k}) = \sup_{s \in I_{j,k}} B_s - \inf_{x \in I_{j,k}} B_s \quad (k = 0, \dots, 2^j - 1)$$

are 2^j i.i.d. random variables, and we have to estimate

$$\sum_k (\text{Osc}_B(I_{j,k}))^p$$

Proposition 9. Let $O_t = \sup_{s \in [0,t]} B_s - \inf_{x \in [0,t]} B_s$. The law of O_t satisfies

$$\text{if } a \leq \sqrt{t}, \quad \mathbb{P}\{O_t \leq a\} \leq \frac{1}{2\pi} \exp\left(-\frac{\pi^2 t}{a^2}\right)$$

and

$$\text{if } a \geq \sqrt{t}, \quad \mathbb{P}\{O_t \geq a\} \leq \frac{4a}{\sqrt{2\pi t}} \exp\left(-\frac{a^2}{8t}\right).$$

Proof of Proposition 9: Let $B_t^* = \sup_{[0,t]} |B_s|$. We can reduce the problems of estimations of the oscillation to estimations on B_t^* , since

$$B_t^* \leq \sup_{[0,t]} B_s - \inf_{[0,t]} B_s \leq 2B_t^*.$$

We will need two estimations for the law of the random variable B_t^* . We start by recalling (see [16] Proposition 8.4.27) that

$$\mathbb{P}(B_t^* \leq a) = \frac{1}{\sqrt{2\pi t}} \sum_{k \in \mathbb{Z}} (-1)^k \int_{(2k-1)a}^{(2k+1)a} \exp\left(-\frac{u^2}{2t}\right) du \quad (8.8)$$

so that the density of B_t^* is

$$g_t(a) = \frac{2}{\sqrt{2\pi t}} \sum_{k \in \mathbb{Z}} (-1)^k (2k+1) \exp\left(-\frac{(2k+1)^2 a^2}{2t}\right).$$

Let $u = a/\sqrt{t}$ and $f(x) = x \exp(-x^2/2)$, then

$$g_t(a) = \frac{2}{a\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} (-1)^k f((2k+1)u).$$

Since f is in the Schwartz class, the Poisson summation formula yields

$$\sum_{n \in \mathbb{Z}} f(x + an) = \sum_{k \in \mathbb{Z}} \frac{1}{a} \hat{f}\left(\frac{2\pi k}{a}\right) e^{-2i\pi kx/a}.$$

Applying this formula with $2a$ and subtracting, we obtain

$$\sum_{n \in \mathbb{Z}} (-1)^n f(x + an) = \frac{1}{a} \sum_{k \text{ odd}} \hat{f}\left(\frac{\pi k}{a}\right) e^{i\pi kx/a}.$$

We use this formula with $x = u$, $a = 2u$ and $f(x) = x \exp(-x^2/2)$, whose Fourier transform is $-i\xi\sqrt{2\pi} \exp(-\xi^2/2)$; it follows that

$$\begin{aligned} g_t(a) &= \frac{-\pi}{2u^2a} \sum_{k \text{ odd}} i^{k+1} k \exp\left(-\frac{\pi^2 k^2}{8u^2}\right) \\ &= \frac{\pi}{4u^2a} \sum_{l=0}^{\infty} (-1)^l (2l+1) \exp\left(-\frac{\pi^2 (2l+1)^2}{8u^2}\right). \end{aligned}$$

Therefore

$$\begin{aligned} \mathbb{P}(B_t^* \leq a) &= \int_0^a g_t(x) dx = \sum_{l=0}^{\infty} \frac{(-1)^l}{2\pi(2l+1)} \exp\left(-\frac{\pi^2 (2l+1)^2 t}{a^2}\right) \\ &\leq \frac{1}{2\pi} \exp\left(-\frac{\pi^2 t}{a^2}\right), \end{aligned}$$

so that

$$\mathbb{P}\{O_t \leq a\} \leq \mathbb{P}\{B_t^* \leq a\} \leq \frac{1}{2\pi} \exp\left(-\frac{\pi^2 t}{a^2}\right).$$

Separating the term $k = 1$ from the others in (8.8) it follows that

$$\begin{aligned} \mathbb{P}(B_t^* \geq a) &= \frac{2}{\sqrt{2\pi t}} \int_a^{\infty} \exp\left(-\frac{u^2}{2t}\right) du \\ &\quad - \frac{2}{\sqrt{2\pi t}} \sum_{k=1}^{\infty} (-1)^k \int_{(2k-1)a}^{(2k+1)a} \exp\left(-\frac{u^2}{2t}\right) du \end{aligned}$$

which is bounded by $\frac{8a}{\sqrt{2\pi t}} \exp\left(-\frac{a^2}{2t}\right)$. Therefore

$$\mathbb{P}\{O_t \geq a\} \leq \mathbb{P}\{B_t^* \geq a/2\} \leq \frac{4a}{\sqrt{2\pi t}} \exp\left(-\frac{a^2}{8t}\right),$$

hence the second point of Proposition 9 holds.

Let j be fixed and $N = 2^j$. It follows that the event

$$\{\text{One of the oscillations } \text{Osc}_B(I_{j,k}) \text{ is smaller than } 1/(\sqrt{N} \log N)\}$$

has probability less than $Ne^{-\pi(\log N)^2}$; and similarly the event

$$\{\text{One of the oscillations } \text{Osc}_B(I_{j,k}) \text{ is larger than } \log N/\sqrt{N}\}$$

has probability less than $2N(\log N)e^{-\pi(\log N)^2}$. The structure function can be estimated just as in the increments case, but the computations are much easier because the distribution of the suprema decays very strongly near 0, and a direct application of the Borel-Cantelli Lemma shows that (8.7) holds, so that the multifractal formalism yields now the correct spectrum for a.e. sample path of Brownian motion.

References

- [1] P. Abry, P. Goncalves and J. Lévy-Véhel *Lois d'échelle, Fractales et ondelettes*, Vol. 1 and 2, Hermes (2002).
- [2] P. Abry and F. Sellan, *The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation*, Appl. Comput. Harmon. Anal. 3, No.4, pp. 377–383 (1996).
- [3] P. Abry, S. Jaffard, and B. Lashermes. Revisiting scaling, multifractal, and multiplicative cascades with the wavelet leader lens. In *Optic East, Wavelet Applications in Industrial Applications II*, Vol. 5607, pages 103–117, Philadelphia, USA, 2004.
- [4] A. Arneodo, B. Audit, N. Decoster, J.-F. Muzy, C. Vaillant, *Wavelet-based multifractal formalism: applications to DNA sequences, satellite images of the cloud structure and stock market data*, in: The Science of Disasters; A. Bunde, J. Kropp, H. J. Schellnhuber eds., Springer pp. 27–102 (2002).
- [5] A. Arneodo, E. Bacry, S. Jaffard, J.-F. Muzy, *Oscillating singularities on Cantor sets: A grandcanonical multifractal formalism*, J. Statist. Phys., vol. 87 pp. 179–209 (1997)
- [6] A. Arneodo, E. Bacry, J.-F. Muzy, *The thermodynamics of fractals revisited with wavelets*, Physica A Vol.213 p.232-275 (1995).
- [7] A. Arneodo, E. Bacry, and J.-F. Muzy. Random cascades on wavelet dyadic trees. *J. Math Phys.*, 39(8):4142–4164, 1998.
- [8] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, and M. S. Taqqu. Generators of long-range dependent processes: a survey. In P. Doukhan, G. Oppenheim, and M. S. Taqqu, editors, *Long-Range Dependence: Theory and Applications*, pages 579–623, Boston, 2003. Birkhäuser.
- [9] J.-M. Aubry and S. Jaffard *Random wavelet series* Comm. Math. Phys., vol. 227, pp. 483–514 (2002).
- [10] J. Barral and S. Seuret, *From multifractal measures to multifractal wavelet series*, preprint (2005).
- [11] H. Brezis, *Analyse fonctionnelle*, Masson (1983).
- [12] G. Brown, G. Michon et J. Peyrière, *On the multifractal analysis of measures*, J. Statist. Phys., vol. 66 pp. 775–790 (1992).
- [13] A. B. Chhabra and C. Meneveau and R. V. Jensen and K. R. Sreenivasan, Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence, *Physical Review A*, 40(9), 1989, 5284–5294.
- [14] E. Chassande-Mottin and P. Flandrin, *On the time-frequency detection of chirps*, Appl. Comput. Harmon. Anal., Vol. 6 pp. 252–281 (1999).
- [15] A. Cohen and R. Ryan, *Wavelets and multiscale signal processing*, Chapman & Hall (1995).
- [16] D. Dacunha-Castelle and M. Duflo *Probabilités et Statistiques Vol. 2* Masson (1993)
- [17] I. Daubechies. *Orthonormal bases of compactly supported wavelets*, Comm. Pure and App. Math. Vol. 41, pp. 909–996 (1988).
- [18] P. Deheuvels and D. M. Mason *Functional laws of the iterated logarithm for the increments of empirical and quantile processes*, Ann. of Proba., Vol. 20 pp. 1248–1287 (1992).

- [19] A. Durand *Random wavelet series with correlated coefficients*, preprint (2005).
- [20] K. Falconer *Fractal geometry*, John Wiley and sons (1990).
- [21] A. Fraysse, *Résultats de généricité en analyse multifractale*, Thèse de l'Université Paris XII (2005).
- [22] A. Grossmann, J. Morlet and T. Paul *Transforms associated to square integrable group representations. I. General results*, J. Math. Phys Vol 26 (10) pp. 2473-2479 (1985) .
- [23] T. Halsey, M. Jensen, L. Kadanoff, I. Procaccia and B. Shraiman, *Fractal measures and their singularities: The characterization of strange sets*, Phys. Rev. A, vol. 33 pp. 1141–1151 (1986)
- [24] G.H. Hardy, J. E. Littlewood, and G. Pólya *Inequalities*, Cambridge University Press (1952).
- [25] S. Jaffard, *Exposants de Hölder en des points donnés et coefficients d'ondelettes*, C. R. Acad. Sci. Sér. I Math., vol. 308 pp. 79–81 (1989).
- [26] S. Jaffard *Pointwise smoothness, two-microlocalization and wavelet coefficients*, Publ. Mat. **35** (1991), 155–168.
- [27] S. Jaffard, *Multifractal formalism for functions. Part 1: Results valid for all functions and Part 2: Selfsimilar functions*, SIAM J. Math. Anal. **28** (1997), 944–998.
- [28] S. Jaffard, *Oscillation spaces: Properties and applications to fractal and multifractal functions*, J. Math. Phys., vol. 39 pp. 4129–4141 (1998)
- [29] S. Jaffard, *Wavelet techniques in multifractal analysis*, Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus et M. van Frankenhuysen Eds., Proceedings of Symposia in Pure Mathematics, AMS, Vol. 72 Part 2 pp. 91–152 (2004).
- [30] S. Jaffard, *On Davenport expansions*, Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus et M. van Frankenhuysen Eds., Proceedings of Symposia in Pure Mathematics, AMS, Vol. 72 Part 1 pp. 273–303 (2004).
- [31] S. Jaffard, *Pointwise regularity associated with function spaces and multifractal analysis*, Approximation and Probability, T. Figiel and A. Kamont Eds., Banach Center Pub. Vol. 72 pp. 93–110 (2006).
- [32] S. Jaffard and C. Melot, *Wavelet analysis of fractal Boundaries, Part 1: Local regularity and Part 2: Multifractal formalism*, Comm. Math. Phys., Vol. 258 n. 3, pp. 513–565 (2005).
- [33] S. Jaffard and Y. Meyer. *Wavelet methods for pointwise regularity and local oscillation of functions*, Mem. Am. Math. Soc. **123** (1996), 587.
- [34] S. Jaffard, Y. Meyer and R. Ryan *Wavelets: Tools for Science and Technology*, S.I.A.M., (2001)
- [35] J.-P. Kahane, *Some Random Series of Functions*, Cambridge University Press, 1985.
- [36] J.-P. Kahane, and J. Peyrière, *Sur certaines martingales de Benoit Mandelbrot*, Adv. Math., vol. 22, pp. 131–145 (1976).
- [37] P. Kestener and A. Arneodo, *Three-dimensional wavelet-based multifractal method: The need for revisiting the multifractal description of turbulence dissipation data*, Physical Review Letters Vol 91 p. 194501, (2003).

- [38] K.-S. Lau, Ka-Sing and S.-M. Ngai, *Multifractal measures and a weak separation condition*, Adv. Math. 141, No.1, 45-96 (1999).
- [39] B. Lashermes, P. Abry, and P. Chainais. New insights into the estimation of scaling exponents. *Int. J. of Wavelets, Multiresolution and Information Processing*, 2(4):497-523, 2004.
- [40] B. Lashermes, S. Jaffard, and P. Abry. Wavelet Leader based Multifractal Analysis. In *Proc. of the Int. Conf. on Acoust. Speech and Sig. Proc.*, 2005.
- [41] W. Li, *Small ball probabilities for Gaussian Markov processes under the L_p -norm*, Stoch. Proc. Appl. Vol. 92, pp. 87-102 (2001).
- [42] W. Li and Q.-M. Shao, *Gaussian processes: Inequalities, small ball probabilities and applications*, D. N. Shanbhag et al. ed., Stochastic processes: Theory and methods. Amsterdam: North-Holland/ Elsevier. Handb. Stat. Vol. 19, pp. 533-597 (2001).
- [43] S. Mallat *A Wavelet tour of signal processing*, Academic Press (1998)
- [44] S. Mallat and W. L. Hwang *Singularity detection and processing with wavelets*, IEEE Trans. Info. Theo. Vol. 38 pp. 617-643 (1992)
- [45] B. Mandelbrot, *Intermittent turbulence in selfsimilar cascades: divergence of high moments and dimension of the carrier*, J. Fluid Mech., vol. 62 pp. 331-358 (1974).
- [46] C. Melot, *Oscillating singularities in Besov spaces*, J. Math. Pures Appl., Vol. 9, . Ser. 83, pp. 367-416 (2004).
- [47] C. Meneveau, Analysis of turbulence in the orthonormal wavelet representation, J. Fluid Mech., 232, 469-520, 1991.
- [48] C. Meneveau and K. R. Sreenivasan, The multifractal spectrum of the dissipation field in turbulent flows, Nuclear Physics B, 2, 49-76, 1987.
- [49] Y. Meyer, *Ondelettes et opérateurs*. Hermann, 1990.
- [50] Y. Meyer, *Wavelets, Vibrations and Scalings*, CRM Ser. AMS Vol. 9, Presses de l'Université de Montréal (1998).
- [51] Y. Meyer, F. Sellan and M. Taqqu, *Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion*, J. Fourier Anal. Appl. 5, No.5, pp. 465-494 (1999)
- [52] Y. Meyer and H. Xu. *Wavelet analysis and chirps*, Appl. Comput. Harmon. Anal. Vol. 4, pp. 366-379 (1997).
- [53] D. Monrad and H. Rootzén *Small values of Gaussian processes and functional laws of the iterated logarithm*, Proba. Theo. Rel. Fields Vol. 101 p. 173-192 (1995).
- [54] L. Olsen, *A multifractal formalism*, Adv. Math., vol. 116 pp. 92-195 (1995).
- [55] G. Parisi and U. Frisch, *On the singularity structure of fully developed turbulence*; appendix to *Fully developed turbulence and intermittency*, by U. Frisch; Proc. Int. Summer school Phys. Enrico Fermi, 84-88 North Holland (1985).
- [56] R. Riedi, *An improved multifractal formalism and self-similar measures*, J. Math. Anal. Appl. Vol. 189, No.2, pp. 462-490 (1995).
- [57] S. Roux, B. Lashermes, P. Abry and S. Jaffard, Contributions à l'étude des performances statistiques des estimateurs multifractals, XXIème colloque GRETSI, Louvain- la-Neuve, belgium, 2005.

- [58] K.R. Sreenivasan, *Fractals and multifractals in turbulence*, Ann. Rev. Fluid Mech., 23, 1991, 539–600.
- [59] W. Stute, *The oscillation behavior of empirical processes*, Ann. Proba. Vol. 10, pp. 86–107 (1982).

Stéphane Jaffard

Laboratoire d'Analyse et de Mathématiques Appliquées, Université Paris XII,
61 Avenue du Général de Gaulle,
94010 Créteil Cedex, France
and Institut Universitaire de France.
e-mail: jaffard@univ-paris12.fr

Bruno Lashermes

Laboratoire de Physique / École Normale Supérieure de Lyon and CNRS UMR 5672,
46, allée d'Italie,
69364 Lyon cedex, France
e-mail: Bruno.Lashermes@ens-lyon.fr

Patrice Abry

Laboratoire de Physique / École Normale Supérieure de Lyon and CNRS UMR 5672,
46, allée d'Italie,
69364 Lyon cedex, France
e-mail: Patrice.Abry@ens-lyon.fr

Application of Fast Wavelet Transformation in Parametric System Identification

Klaus Markwardt

Abstract. This work deals with discrete embedding of system operators in identification models on basis of Fast Wavelet Transformation (FWT). In particular for FWT-models of linear dynamic systems the missing variables can be calculated with the help of connection coefficients. The application of connection coefficients provides the direct projection of the system operators into the corresponding wavelet space. Here a class of operators is introduced, which satisfies certain permutability relations with respect to dilations and translations. This class contains especially derivation and integration operators and some special convolution operators, like the Hilbert-transform. Such a definition allows the systematic determination of generalized connection coefficients. It gives so the possibility to realize identification procedures for different models and their implementations in a unified pattern. The method can be used for all biorthogonal wavelet systems whose synthesis functions are in the domain of the system operators.

Keywords. FWT, biorthogonal wavelets, connection coefficients, system identification.

1. Biorthogonal Wavelet Systems

Every vector $c \in \mathbb{R}^{\mathbb{Z}}$ with finite many entries is here shortly called a (finite discrete) filter. With appropriate indices $s, u \in \mathbb{Z}$ it can be written in the form

$$c = (\dots, 0, 0, c_s, \dots, c_u, 0, 0, \dots). \quad (1.1)$$

Let's assume that $\{h, \tilde{h}, g, \tilde{g}\}$ is a system of four filters with the following properties

$$\sum_k h_k \tilde{h}_{k+2l} = \delta_{l0}, \quad \sum_k g_k \tilde{g}_{k+2l} = \delta_{l0}, \quad (1.2)$$

$$\sum_k \tilde{h}_k g_{k+2l} = 0, \quad \sum_k h_k \tilde{g}_{k+2l} = 0, \quad (1.3)$$

$$\sum_k h_k = \sum_k \tilde{h}_k = \sqrt{2}, \quad \sum_k g_k = \sum_k \tilde{g}_k = 0. \quad (1.4)$$

The two integers s, u are chosen such that all four filters can be written in the form (1.1) with the same minimal $[s, u] \subset \mathbb{Z}$. It follows

$$\sum_k (-1)^k h_k = 0 \quad \text{and} \quad \sum_k (-1)^k \tilde{h}_k = 0. \quad (1.5)$$

Further conclusions and relations to other biorthogonal filter representations are given in [5], cp. [2], [3], and for the special orthogonal case in [1], [4]. If this filters fulfill some further properties, which are related to the eigenvalues of two Lawton matrices, then by

$$\varphi(t) = \sqrt{2} \sum_k h_k \varphi(2t - k), \quad \psi(t) = \sqrt{2} \sum_k g_k \varphi(2t - k) \quad (1.6)$$

$$\tilde{\varphi}(t) = \sqrt{2} \sum_k \tilde{h}_k \tilde{\varphi}(2t - k) \quad \tilde{\psi}(t) = \sqrt{2} \sum_k \tilde{g}_k \tilde{\varphi}(2t - k) \quad (1.7)$$

$$\int_{-\infty}^{\infty} \varphi(t) dt = \int_{-\infty}^{\infty} \tilde{\varphi}(t) dt = 1. \quad (1.8)$$

scaling functions $\varphi, \tilde{\varphi} \in L^2(\mathbb{R})$ and wavelets $\psi, \tilde{\psi} \in L^2(\mathbb{R})$ are determined uniquely. They have compact support, such that with the above defined s, u

$$\text{supp}(\varphi) \subset [s, u], \quad \text{supp}(\tilde{\varphi}) \subset [s, u] \quad (1.9)$$

is valid. The deduced variants $\varphi_{j,k}(t) = \sqrt{2^{-j}} \varphi(2^{-j}t - k)$, \dots for $i, j, k, l \in \mathbb{Z}$ satisfy the biorthogonal relations

$$\langle \varphi_{j,k}, \tilde{\varphi}_{j,l} \rangle = \delta_{k,l} \quad (1.10)$$

$$\langle \varphi_{i,k}, \tilde{\psi}_{j,l} \rangle = \langle \tilde{\varphi}_{i,k}, \psi_{j,l} \rangle = 0 \quad \text{for } i \geq j \quad (1.11)$$

$$\langle \psi_{i,k}, \tilde{\psi}_{j,l} \rangle = \delta_{i,j} \cdot \delta_{k,l}. \quad (1.12)$$

In particular $\{\psi_{i,k}\}, \{\tilde{\psi}_{i,k}\}$ builds a pair of biorthogonal Riesz bases in $L^2(\mathbb{R})$, cp.[2] and [3]. Using $\tilde{\varphi}, \tilde{\psi}$ in the synthesis and φ, ψ in the analysis we get for the projection on the level J

$$(\tilde{\mathbf{P}}_J f)(t) = \sum_k S_{Jk} \tilde{\varphi}_{J,k}(t) \quad \text{with} \quad S_{Jk} = \langle f, \varphi_{J,k} \rangle. \quad (1.13)$$

Let f_s be a sampled signal. After suitable dilation it can be approximated by (1.13) on the level $J = 0$ using wavelet sampling approximation or an preprocessing method for getting the basic coefficients S_{0k} , cp.[5]. Successive applying of the decomposition formulas (FWT)

$$S_{l+1,k} = \sum_r h_{r-2k} S_{l,r}, \quad W_{l+1,k} = \langle f, \psi_{l+1,k} \rangle = \sum_r g_{r-2k} S_{l,r} \quad (1.14)$$

gives the representation with the wavelet coefficients W_{lk}

$$(\tilde{\mathbf{P}}_0 f)(t) = \sum_k S_{Lk} \tilde{\varphi}_{L,k} + \sum_{l=1}^L \left\{ \sum_k W_{lk} \tilde{\psi}_{l,k} \right\}. \quad (1.15)$$

With increasing l the frequency center of the inner sums decreases step by step by the factor 0.5. Using wavelet packets induced by appropriate $\{h, \tilde{h}, g, \tilde{g}\}$ you get trees with the possibility of stronger locally frequency resolution, cp.[10].

The inverse formula of (1.14) is known as

$$S_{l,k} = \sum_r \tilde{h}_{k-2r} S_{l-1,r} + \sum_r \tilde{g}_{k-2r} W_{l-1,r}. \quad (1.16)$$

2. Connection Coefficients for D-homogeneous Operators

The parameter depending affine groups $T(r)$ and $D(s)$ are defined by

$$T(r) : f(t) \mapsto f(t-r), \quad r \in \mathbb{R} \quad \text{and} \quad D(s) : f(t) \rightarrow \sqrt{s} f(st), \quad s > 0 \quad (2.1)$$

With this notations $\varphi_{l,j} = D(2^{-l}) T(j) \varphi$ and corresponding formulas are valid. This paper is researched into connection coefficients for operators with the following properties.

Definition 2.1. Here will be written $\mathbf{K} \in \mathbb{H}_\lambda(\mathbf{V}_1, \mathbf{V}_2)$ or shorter $\mathbf{K} \in \mathbb{H}_\lambda$ with some $\lambda \in \mathbb{R}$, if $\mathbf{K} : \mathbf{V}_1 \mapsto \mathbf{V}_2$ is a linear operator with \mathbf{V}_1 as a subspace of $L_0^2(\mathbb{R})$ and \mathbf{V}_2 as a subspace of $L_{loc}^2(\mathbb{R})$, whereas

$$T(r) \mathbf{V}_i \subset \mathbf{V}_i \quad \text{and} \quad D(s) \mathbf{V}_i \subset \mathbf{V}_i, \quad (2.2)$$

$$\mathbf{K} T(r) = T(r) \mathbf{K} \quad \text{and} \quad \mathbf{K} D(s) = s^\lambda D(s) \mathbf{K} \quad (2.3)$$

are valid. Operators with the right property in (2.3) will called D-homogeneous of degree λ .

Example. Let \mathbf{V}_1 be the space of functions f with compact support and $f \in H^n(\mathbb{R})$ (Sobolev space of order n). For $\mathcal{D} = \frac{d}{dt}$ the operator \mathcal{D}^n with $n \in \mathbb{N}$ satisfies then $\mathcal{D}^n \in \mathbb{H}_n$. \mathbf{V}_2 can be taken as the space of all functions $g \in L^2(\mathbb{R})$ with compact support.

Example. With $f^{(-1)}(t) = \mathcal{I}(f)(t) := \int_{-\infty}^t f(\tau) d\tau$, $\mathbf{V}_1 = L_0^2(\mathbb{R})$ and

$\mathbf{V}_2 = L_{loc}^2(\mathbb{R})$ results $\mathcal{I}^n \in \mathbb{H}_{-n}$ for $n \in \mathbb{N}$. Obviously the identity operator $\mathbf{I} = \mathcal{I}^0 = \mathcal{D}^0$ satisfies $\mathbf{I} \in \mathbb{H}_0$.

Example. With suitable spaces $\mathbf{V}_1, \mathbf{V}_2$, the convolution operator

$$(\mathbf{K}f)(t) = \int_{-\infty}^{\infty} k(t-\tau) f(\tau) d\tau, \quad (2.4)$$

and the conventions $v_+(t) = \max(v(t), 0)$, $v_-(t) = \min(v(t), 0)$ the following special cases for definition 2.1 result:

1. $k(t)$ is one of the functions $t^{\alpha-1}$, $t_+^{\alpha-1}$, $t_-^{\alpha-1}$ with arbitrary positive α . The spaces are given by $\mathbf{V}_1 = \mathcal{C}_0$, $\mathbf{V}_2 = \mathcal{C}$. Now you get $\mathbf{K} \in \mathbb{H}_{-\alpha}$.
2. The same as above, but with $\alpha > 0.5$ and the changed space $\mathbf{V}_1 = L_0^2$. $\alpha \in \mathbb{N}$ provides the integrals of the previous example.

3. With $k(t) = \pi^{-1} t^{-1}$ and (2.4) in the sense of Cauchy principal value you get the Hilbert-Transformation. Now let \mathbf{V}_1 be a space of Hoelder continuous functions with compact support, $\mathbf{V}_1 \subset \mathcal{C}^\alpha$, $\alpha > 0$ and $\mathbf{V}_2 = \mathcal{C}$.

The discrete embedding of an operator $\mathbf{K} \in \mathbb{H}_\lambda$ over the projection $\tilde{\mathbf{P}}_0$ is given by

$$(\mathbf{K} \tilde{\mathbf{P}}_0 f)(t) = \sum_{\nu} S_{0\nu} \mathbf{K} \tilde{\varphi}_{0,\nu}(t) \xrightarrow{\tilde{\mathbf{P}}_0} \sum_{\nu} S_{0\nu}^K \tilde{\varphi}_{0,\nu}(t) = (\tilde{\mathbf{P}}_0 \mathbf{K} \tilde{\mathbf{P}}_0 f)(t) \quad (2.5)$$

with

$$S_{0l}^K = \sum_{\nu} S_{0\nu} \langle \mathbf{K} \tilde{\varphi}_{0,\nu}, \varphi_{0,l} \rangle \quad \text{or} \quad S_0^K = S_0 K^{(0)}. \quad (2.6)$$

Here $K^{(0)}$ is the matrix induced by the corresponding system of scalar products. It's elements are special connection coefficients, the basic coefficients. Coming from $\varphi_{l,j} = D(2^{-l}) T(j) \varphi, \dots, \dots$ the connection coefficients will defined by

$$\Gamma_{lj}^{mk}(\tilde{f}, f) = \langle \mathbf{K} \tilde{f}_{l,j}, f_{m,k} \rangle. \quad (2.7)$$

Some shorter forms for the basic coefficients are used

$$\Gamma_j^k := \Gamma_{0j}^{0k}(\tilde{\varphi}, \varphi) \quad \text{and} \quad \Gamma_j^k(\tilde{f}, g) := \Gamma_{0j}^{0k}(\tilde{f}, g). \quad (2.8)$$

The following theorems are proofed in [5].

Theorem 2.2. *The connection coefficients of $\mathbf{K} \in \mathbb{H}_\lambda$ satisfy*

$$2^{l\lambda} \Gamma_{lj}^{mk} = \Gamma_{00}^{m-l, k-2^{l-m}j}, \quad 2^{l\lambda} \Gamma_{lj}^{lk} = \Gamma_{00}^{0, k-j} = \Gamma_0^{k-j}, \quad (2.9)$$

$$\Gamma_{l,j+n}^{l,k+n} = \Gamma_{l,j}^{l,k}, \quad \Gamma_{j+n}^{k+n} = \Gamma_j^k, \quad \Gamma_{l,j}^{l,k} = \Gamma_{l,0}^{l,k-j}, \quad \Gamma_j^k = \Gamma_0^{k-j} \quad (2.10)$$

$$\Gamma_{lj}^{mk} = 2^{n\lambda} \Gamma_{l+n,j}^{m+n,k}, \quad (2.11)$$

if $\tilde{f}, f \in \mathbf{V}_1$.

Theorem 2.3. *Both for $\tilde{f} = \tilde{\varphi}$, $\tilde{f} = \tilde{\psi}$ and $f = \varphi$, $f = \psi$ follows with $\mathbf{K} \in \mathbb{H}_\lambda$*

$$\sum_{\nu=s}^u h_\nu \Gamma_{lj}^{m-1, \nu+2k}(\tilde{f}, \varphi) = \Gamma_{lj}^{mk}(\tilde{f}, \varphi) = 2^\lambda \sum_{\nu=s}^u h_\nu \Gamma_{l+1,j}^{m, \nu+2k}(\tilde{f}, \varphi) \quad (2.12)$$

$$\sum_{\nu=s}^u g_\nu \Gamma_{lj}^{m-1, \nu+2k}(\tilde{f}, \varphi) = \Gamma_{lj}^{mk}(\tilde{f}, \psi) = 2^\lambda \sum_{\nu=s}^u g_\nu \Gamma_{l+1,j}^{m, \nu+2k}(\tilde{f}, \varphi) \quad (2.13)$$

$$\sum_{\nu=s}^u \tilde{h}_\nu \Gamma_{l-1, \nu+2j}^{mk}(\tilde{\varphi}, f) = \Gamma_{lj}^{mk}(\tilde{\varphi}, f) = 2^\lambda \sum_{\nu=s}^u \tilde{h}_\nu \Gamma_{l, \nu+2j}^{m+1, k}(\tilde{\varphi}, f) \quad (2.14)$$

$$\sum_{\nu=s}^u \tilde{f}_\nu \Gamma_{l-1, \nu+2j}^{m,k}(\tilde{\varphi}, f) = \Gamma_{lj}^{mk}(\tilde{\psi}, f) = 2^\lambda \sum_{\nu=s}^u \tilde{f}_\nu \Gamma_{l, \nu+2j}^{m+1, k}(\tilde{\varphi}, f) \quad (2.15)$$

(For the equations on the left sides only linearity is necessary.) If the basic coefficients are known, than step by step all other connection coefficients can be verified.

3. Calculation of Basic Coefficients

The proofs of the following theorems you find in [5] for the general case and in [7] for the special orthogonal case. They generalize some results in [1] and [12].

Theorem 3.1. For $\mathbf{K} \in \mathbb{H}_\lambda$ and $\mathcal{N} = u - s$, cp.(1.1), the basic coefficients satisfy the system of equations

$$\Gamma_0^k = 2^\lambda \sum_{\nu} \alpha_{2k-\nu} \Gamma_0^\nu, \quad 2k - \mathcal{N} \leq \nu \leq 2k + \mathcal{N}, \quad \forall k \in \mathbb{Z} \quad (3.1)$$

with $\alpha_\nu = \sum_{\mu} h_\mu \tilde{h}_{\nu+\mu}$. The relations $\alpha_0 = 1$, $\sum_{\nu} \alpha_{2\nu-1} = 1$, $\alpha_{2\nu} = 0$ for $\nu \neq 0$ and $\alpha_\nu = 0$ for $|\nu| > \mathcal{N}$ are valid.

The following definitions are basic for researching the system (3.1).

A submatrix $L = L(h, \tilde{h})$ of (3.1) is given by

$$L_{ij} = \alpha_{2i-j} \quad \text{with} \quad i, j = 1 - \mathcal{N}, \dots, \mathcal{N} - 1 \quad (3.2)$$

and called here **Lawton matrix** of 2. kind.

An operator \mathbf{K} will called **causal**, if

$$f(t) = 0 \quad \text{for} \quad t < T \quad \text{results ever} \quad (Kf)(t) = 0 \quad \text{for} \quad t < T. \quad (3.3)$$

For $\mathbf{K} \in \mathbb{H}_\lambda$ the case $T = 0$ in (3.3) is sufficient.

Theorem 3.2. If $\mathbf{K} \in \mathbb{H}_\lambda$ is causal, then

$$\Gamma_0^k(\tilde{\varphi}, \varphi) = 0 \quad \text{for} \quad k < 1 - \mathcal{N} \quad (3.4)$$

is valid. If furthermore for $\nu > \mathcal{N} - 1$ the basic coefficients $G_\nu := \Gamma_0^\nu$ are known then with the Lawton matrix L in (3.2) follows

$$\sum_{\nu=1-\mathcal{N}}^{\mathcal{N}-1} (L_{k\nu} - 2^{-\lambda} \delta_{k\nu}) \Gamma_0^\nu(\tilde{\varphi}, \varphi) = - \sum_{\nu=\mathcal{N}}^{3\mathcal{N}-2} \tilde{\alpha}_{2k-\nu} G_\nu, \quad k = 1 - \mathcal{N}, \dots, \mathcal{N} - 1 \quad (3.5)$$

This system has then an unique solution if $\mu = 2^{-\lambda}$ is no eigenvalue of the Lawton matrix L . If the right side of (3.5) vanishes and the eigenvalue $\mu = 2^{-\lambda}$ has order 1, then the solution becomes unique if an additional inhomogeneous linear equation can be found for the unknowns Γ_0^ν .

Remark 3.3. Making stronger the condition (1.5) regarding $\{\tilde{h}_k\}$ by

$$\sum_k (-1)^k k^\nu \tilde{h}_k = 0 \quad \text{for} \quad \nu = 0, 1, \dots, n, \quad (3.6)$$

(1.13) changes for better approximation properties with increasing n . In particular the moments of the dual wavelets vanish up to order n .

$$M_\nu(\psi) = \int_{-\infty}^{\infty} t^\nu \psi(t) dt = 0 \quad \text{for all} \quad \nu = 0, 1, \dots, n, \quad (3.7)$$

Theorem 3.4. *Let be the scaling coefficients \tilde{h}_k satisfied additionally (3.6) and $\tilde{\varphi} \in H^n(\mathbb{R})$. Then the basic coefficients of the differentiation operator \mathcal{D}^n $\Gamma_0^\nu(n) = \langle \tilde{\varphi}^{(n)}, \varphi_{0,\nu} \rangle$ fulfill the homogeneous equations*

$$\sum_{\nu=1-\mathcal{N}}^{\mathcal{N}-1} (L_{k\nu} - 2^{-n}\delta_{k\nu}) \Gamma_0^\nu(n) = 0, \quad k = 1 - \mathcal{N}, \dots, 0, \dots, \mathcal{N} - 1 \quad (3.8)$$

and the inhomogeneous equation

$$\sum_{\nu=1-\mathcal{N}}^{\mathcal{N}-1} \nu^n \Gamma_0^\nu(n) = (-1)^n n! \quad (3.9)$$

Remark 3.5. Corresponding to the integration operator \mathcal{I}^n , which is causal, you have

$$\Gamma_0^\nu(-n) = \langle \mathcal{I}^n \tilde{\varphi}, \varphi_{0\nu} \rangle = 0 \quad \text{for } \nu \leq -\mathcal{N}, \quad \forall n \in \mathbb{N}.$$

First you must calculate

$$G_\nu(-n) = \Gamma_0^\nu(-n) \quad \text{with } n \in \mathbb{N}, \quad \text{for } \nu \geq \mathcal{N}.$$

on the right side of (3.5). In [5] recurrence formulas are developed for calculating the $G_\nu(-n)$, which at last only use discrete moments m_n of the scaling filters h, \tilde{h}

$$m_n(h) = \sum_{\nu=s}^u \nu^n h_\nu.$$

This formulas provide the $G_\nu(-n)$ as polynomials in ν of order $(n-1)$. Note that in [5] has been worked with different normed filters $a = \sqrt{2}h, \dots$. The first three $G_\nu(-n)$ are given by

$$G_\nu(-1) = \Gamma_0^\nu(-1) = 1 \quad \text{for } \nu \geq \mathcal{N},$$

$$G_\nu(-2) = \Gamma_0^\nu(-2) = \nu + \frac{\sqrt{2}}{2} [m_1(h) - m_1(\tilde{h})] \quad \text{for } \nu \geq \mathcal{N},$$

$$G_\nu(-3) = \Gamma_0^\nu(-3) = \frac{1}{2}\nu^2 + \frac{\sqrt{2}}{2} [m_1(h) - m_1(\tilde{h})] \nu + C(3, \mathcal{I}) \quad \text{for } \nu \geq \mathcal{N}$$

the last with

$$C(3, \mathcal{I}) = \frac{\sqrt{2}}{12} [m_2(h) + m_2(\tilde{h})] + \frac{1}{6} [(m_1(h))^2 - 3m_1(h)m_1(\tilde{h}) + (m_1(\tilde{h}))^2].$$

The remaining Γ_0^ν are found now by (3.5) after controlling that $\mu = 2^n$ for $n = 1, 2, 3$ is no eigenvalue of the Lawton matrix L . In the orthogonal case you obtain $\tilde{h} = h$, $\varphi = \tilde{\varphi}$, $\psi = \tilde{\psi}$ and you get the original Lawton matrix in [4]. The greatest eigenvalue becomes 1 if the wavelets build an orthogonal system, such that theorem 3.2 can be applied. Applying formulas given in [7] this special case would used for an identification algorithm in the engineering work [11].

In [5] is shown that the wavelet based quadrature filter $\{\Gamma_0^\nu(-1)\}$ can be decomposed, such that it's application is no more extensive then that of a good

quadrature formula. We think that this wavelet adapted filter gives better results in identification procedures.

Theorem 3.2 is also applicable in the case of different biorthogonal B-spline produced wavelets (cp.[2] and [3]), such that identification algorithms can also be realized with the help of this systems, cp. [10].

Remark 3.6. Corresponding to the differentiation operator \mathcal{D}^n theorems 3.2 and 3.4 are applicable for customary orthogonal or biorthogonal wavelets, like Daubechies wavelets, Symlets, Coiflets and B-Spline produced wavelets. If the order of these wavelets becomes sufficiently high then 2^{-n} becomes an eigenvalue of order 1. Practically this can be used in the cases $n = 1$ and $n = 2$. This extends the possibilities given in [1] and [12] on the biorthogonal case.

4. Applications and Outlook

Here an extract of [5] (generalization of [7]) is given. This works were written to give engineers some special background in parametric system identification. But sure they are useful also in other fields of applied mathematics, for instance in solving operator equations. In [11] such basics were used to work out a wavelet-based method that allows the direct identification of system parameters from data generated in dynamic tests of a structure. Of specific interest was an algorithm that requires the response data solely in the form of accelerations. The method was successful applied to multi-degree-of-freedom systems and to existing structural elements and structures in civil engineering. The parameters of linear time-invariant systems could be identified with a relatively high accuracy, even if the simulated data were contaminated by noise. The identified stiffness terms agreed better with the original values than the viscous damping parameters. Intended is the improvement of this applications by the use of biorthogonal wavelets and wavelet packets, cp. [10] and [9].

References

- [1] G. Beylkin & B. Torresani, *Implementation of Operators via Filter Banks*. Applied and Computational Harmonic Analysis 3, 1996, 164–185.
- [2] A. Cohen, I. Daubechies & J.C. Feauveau, *Biorthogonal bases of compactly supported wavelets*. Communications on Pure and Applied Mathematics, Vol.XLV, 485–560, 1992.
- [3] A. Cohen, *Biorthogonal Wavelets*. Wavelets in Analysis and Applications II, 123–152, Academic Press 1992.
- [4] W.M. Lawton, *Necessary and sufficient conditions for constructing orthonormal wavelet bases*. J. Math. Phys.32(1), 1440–1443, 1991.
- [5] K. Markwardt, *Die schnelle Wavelet-Transformation als Grundlage für Verfahren zur System- und Parameteridentifikation*. Manuscript, pages 1–122 (2005), not published yet.

- [6] K. Markwardt, *Biorthogonale Wavelet-Systeme in der Parameteridentifikation*. Proceedings of the 16th IKM 2003, Gürlebeck, Hempel & Könke (eds.), Weimar, Germany, June 10-12, 2003.
- [7] K. Markwardt, *Systemanalyse und Parameteridentifikation mit Hilfe der schnellen Wavelet-Transformation*. Forschungsbericht, 1–76, 2002, ISM, BUW, Weimar.
- [8] K. Markwardt, *Betrachtungen zur Anwendung der Wavelet-Transformation in der Systemidentifikation*. 6. Institutskolloquium, Bericht 1/00, 2000, 107–126, ISM, Bucher, Burkhardt & Vormwald (eds.).
- [9] M. Brehm, K. Markwardt and V. Zabel, *Applications of Wavelet Packets in System Identification*. Jahrestagung GAMM, 2005, Universit du Luxembourg.
- [10] M. Brehm, K. Markwardt and V. Zabel, *Applications of Biorthogonal Wavelets in System Identification*. ECCOMAS 2004. Jyväskylä, Neittaanmäki, Rossi, K. Majava & O. Pironneau (eds.).
- [11] V. Zabel, *Applications of Wavelet Analysis in System Identification*. Dissertation, 2003, Institute Structural Mechanics, Faculty of Civil Engineering, Bauhaus-University Weimar
- [12] Resnikoff, *Wavelet Analysis*. Springer, 1998

Klaus Markwardt
Coudraystr. 13 B
Bauhaus-University Weimar
Institute for Mathematics/Physics
99421 Weimar
Federal Republic of Germany
e-mail: klaus.markwardt@bauing.uni-weimar.de

Image Denoising by a Novel Digital Curvelet Reconstruction Algorithm

Jian Bai and Xiang-Chu Feng

Abstract. For an anisotropic image, wavelets lose their effects on singularity detection because discontinuities across edges are spatially distributed. Based on the idea of curvelet, a new digital curvelet reconstruction algorithm is proposed. Our algorithm provides sparser representations and keeps low computational complexity. When applying it to the image denoising, much better results than the original algorithm are obtained.

Keywords. Curvelet, ridgelet, digital curvelet, Radon transform, fast Slant Stack.

1. Introduction

Many image processing tasks take advantage of sparse representations of image data where most information is packed into a small number of samples. The wavelet transform is the most popular choices for this purpose. The success of wavelets is mainly due to the good performance for piecewise smooth functions in one dimension. Unfortunately, such is not the case in two dimensions. To overcome the weakness of wavelets in higher dimensions, E.J.Candes and D.L.Donoho introduced the concepts of ridgelet transform and curvelet transform[1,2]. Ridgelet system is very good at representing the smooth functions with line singularities. Essentially, Radon transform was used to the original image, and then the wavelet analysis in the Radon domain was applied. Curvelets are derived from theory of ridgelet. They are based on multiscale ridgelets combined with a spatial bandpass filtering operation to isolate different scales. Curvelet transform filters image into subbands, so there are many subband images with different frequencies. For the lowpass image, apply wavelet analysis to it. For the bandpass images, apply multiscale ridgelets to them. The length and width of curvelets at fine scales are related by a scaling law: $width \approx length^2$. Compare with the traditional isotropic relation in wavelet transform, they have anisotropic relation. This makes they are very good at representing the smooth functions with curve singularities. Thus curvelets have

great potentialities in the image denoising, enhancement and other application cases.

The digital realization of curvelet was proposed in [3,4]. J.L. Starck applied it to the image denoising and the very good results were obtained. He used the generalized Radon inverse transform to implement ridgelet reconstruction. Our algorithm computes the orthogonal projection of the original block image into the space spanned by the thresholded ridgelet bases. When the thresholded ridgelet bases span $n \times n$ (where n is the size of block image) space, the block image can be completely reconstructed. The orthogonal projection is the best approximation of the origin block image, so the new algorithm provides sparser representations. The new algorithm still keeps low computational complexity. It can be computed in order $O(N \log N)$, where $N = n^2$. In this paper we use it to digital curvelet reconstruction. The very high quality and signal-to-noise ratios are illustrated by denoising the noisy Lenna and barbara images.

2. Digital Curvelet Transform

Curvelet bases $r_\mu = r_\mu(x_1, x_2)$ are defined as tight frame of $L^2(R^2)$ in [1,3]. It is the sequence of the following steps: a. subband decomposition b. smooth partitioning c. renormalization d. ridgelet analysis. Curvelet transform uses ridgelet transform as a component step, so we firstly describe a digital ridgelet transform.

2.1. Digital Ridgelet

There are a number of methods that can implement discrete ridgelet transform. We use the method in [5]. Essentially, use fast Slant Stack algorithm to implement Radon transform, and then apply Meyer wavelet transform in t . The digital ridgelet analysis operator R applied to an $n \times n$ image $(I(u, v) : -\frac{n}{2} \leq u, v \leq \frac{n}{2})$ is the $4n^2$ ridgelet coefficients:

$$RI = (\langle I, \rho_\lambda \rangle; \lambda \in \mu, \mu = (j, k, s, l)) \quad (2.1)$$

The following theorem shows the relation between the digital ridgelet transform R and discrete Slant Stack S followed by the Meyer wavelet transform W .

Theorem 1 [5] The digital ridgelet transform is the 1-dimensional Meyer-wavelet transform in t , of the Slant Stack Radon transform:

$$(RI)(j, k, s, l) = \langle\langle SI(\cdot, s, l), \psi_{j,k}(\cdot) \rangle\rangle \quad (2.2)$$

Let R^* be the adjoint of R , then $R^* = S^* \cdot W^{-1}$, where W^{-1} is the Meyer wavelet reconstruction operator, S^* is the adjoint of S . R and R^* can be computed in the order $O(N \log N)$.

2.2. Digital Curvelet Transform

Now we describe a strategy for realizing a digital implementation of the curvelet transform.

Subband Decomposition: Image size is 256×256 . We use Daubechies 4 wavelet to divide it into 8 subbands, at levels $j = 0, 1 \dots 7$. The curvelet subband $s = 1$ corresponds to wavelet subbands $j = 0, 1, 2, 3$. Curvelet subband $s = 2$ corresponds to wavelet subbands $j = 4, 5, 6$. Curvelet subband $s = 3$ corresponds to wavelet subband $j = 7$. Hence, for image f , we get lowpass image f_1 and bandpass images f_2, f_3 .

Partitioning: For f_2 , partition it by 32×32 blocks $f_{2,i,j}$. There are 64 blocks $f_{2,i,j}$ ($1 \leq i \leq 8, 1 \leq j \leq 8$).

For f_3 , partition it by 16×16 blocks $f_{3,m,n}$. There are 256 blocks $f_{3,m,n}$ ($1 \leq m \leq 16, 1 \leq n \leq 16$).

Ridgelet Analysis: For f_2 , use 64×64 overlay blocks $g_{2,p,q}$ to cover it, where ($1 \leq p \leq 7, 1 \leq q \leq 7$).

$$g_{2,p,q} = f_2 \left([(p-1) \times 32 + 1, (p+1) \times 32] \times [(q-1) \times 32 + 1, (q+1) \times 32] \right).$$

For f_3 , use 32×32 overlay blocks $g_{3,x,y}$ to cover it, where ($1 \leq x \leq 15, 1 \leq y \leq 15$),

$$g_{3,x,y} = f_3 \left([(x-1) \times 16 + 1, (x+1) \times 16] \times [(y-1) \times 16 + 1, (y+1) \times 16] \right).$$

Apply ridgelet transform to $g_{2,p,q}$ and $g_{3,x,y}$, getting the ridgelet coefficients $\alpha_{2,p,q,\lambda}$ and $\alpha_{3,x,y,\lambda}$, where λ is ridgelet index.

2.3. Reconstruction Algorithm

The above steps enjoy exact reconstruction, so the reconstruction algorithm is divided into the following steps:

Ridgelet Reconstruction: For ridgelet coefficients of each block image α_η , we apply the inverse transform in [5], where $R^{-1} = S^{-1}W^{-1}$ and S^{-1} is the generalized Radon inverse transform operator in [6].

Smooth Partition: We calculate a pixel value, $f_{2,i,j}$ from its four corresponding block values, namely, 64×64 blocks B_1, B_2, B_3 and B_4 , in the following way:

$$h_1 = w(i_2/l) B_1(i_1, j_1) + w(1 - i_2/l) B_2(i_2, j_1)$$

$$h_2 = w(i_2/l) B_3(i_1, j_2) + w(1 - i_2/l) B_4(i_2, j_2)$$

$$\Delta f_{2,i,j} = w(j_2/l) h_1 + w(1 - j_2/l) h_2$$

where $w(x) = \cos^2(\pi x/2)$ is the smooth function. $l=32$ is the size of block, $i_2 = i_1 - l$, $j_2 = j_1 - l$, $i_1, j_1 > l$. Similarly, for $f_{3,m,n}$, we have $\Delta f_{3,m,n}$.

Subband Reconstruction: For each subband we get the reconstructed subband images $\Delta f_1, \Delta f_2, \Delta f_3$ and undo the bank of subband filters, using the reproducing formula: $\bar{f} = \Delta_1(\Delta f_1) + \Delta_2(\Delta f_2) + \Delta_3(\Delta f_3)$.

3. New Reconstruction Algorithm

We use the following algorithm to implement ridgelet reconstruction.

Apply the ridgelet transform to the original block image I , threshold the ridgelet coefficients, getting the space V spanned by the thresholded ridgelets, index set T and the ridgelet coefficient matrix $M = \begin{cases} \langle I, \rho_\lambda \rangle & \lambda \in T \\ 0 & \lambda \notin T \end{cases}$.

The new algorithm computes the orthogonal projection of the original block image I into the space V . It can be divided into three steps:

Step1: Apply the adjoint ridgelet transform to M , getting $b = R^*M$. In fact, the Mallat algorithm is used to obtain $W^{-1}M$; the adjoint Slant Stack in [6] is used to compute $b = S^*(W^{-1}M)$.

Step 2: Solve f from equation:

$$L_v f = b \quad (3.1)$$

where $L_v f = R^*(\overline{Rf})$, $\overline{Rf} = \begin{cases} Rf & \lambda \in T \\ 0 & \lambda \notin T \end{cases}$.

We use conjugate gradient to solve Eq3.1. Details are in [7]. The whole procedure takes $O(N \log N)$ which is the same as the cost of the original algorithm. For the block reconstructed image f , we have the following results:

Theorem 2 The block reconstructed image by the new algorithm is $f = \sum_{\lambda \in T} \langle I, \rho_\lambda \rangle \tilde{\rho}_\lambda$, where $\{\tilde{\rho}_\lambda\}_{\lambda \in T}$ is the dual frame of $\{\rho_\lambda\}_{\lambda \in T}$ which is the frame of V . f is the orthogonal projection of I into the subspace V .

Proof. According to $R^*(1_\lambda) = \rho_\lambda$, where 1_λ is characteristic matrix,

$$1_\lambda(\mu) = \begin{cases} 1 & \mu = \lambda \\ 0 & \mu \neq \lambda \end{cases}, \text{ so } M = \sum_{\lambda \in T} \langle I, \rho_\lambda \rangle 1_\lambda,$$

we have $R^*M = \sum_{\lambda \in T} \langle I, \rho_\lambda \rangle \rho_\lambda$, $L_V f = \sum_{\lambda \in T} \langle f, \rho_\lambda \rangle \rho_\lambda$, where L_V is the frame operator of space V . It is invertible. $\{\tilde{\rho}_\lambda\}_{\lambda \in T}$ which is the dual frame of $\{\rho_\lambda\}_{\lambda \in T}$ can be defined by the inverse operator $L_V^{-1} : \tilde{\rho}_\lambda = L_V^{-1} \rho_\lambda$, so $f = L_V^{-1} R^*M = \sum_{\lambda \in T} \langle I, \rho_\lambda \rangle \tilde{\rho}_\lambda$.

To prove that f is the orthogonal projection of I into V , it is enough to test for the arbitrary $p \in T$ that $\langle I - f, \rho_p \rangle = 0$. In fact, $\langle I - f, \rho_p \rangle = \langle I, \rho_p \rangle - \sum_{\lambda \in T} \langle I, \rho_\lambda \rangle \langle \tilde{\rho}_\lambda, \rho_p \rangle$.

We know the dual frame in V has property:

$$\sum_{\lambda \in T} \langle \tilde{\rho}_\lambda, \rho_p \rangle \rho_\lambda = \rho_p$$

Therefore $\langle I - f, \rho_p \rangle = 0$. That ends the proof. \square

For each block image with size $n \times n$, we use the new algorithm to implement ridgelet reconstruction, so a new digital curvelet reconstruction algorithm is obtained. Our algorithm introduces an extra redundancy factor of $16J + 1$. The article [4] used the generalized Radon inverse transform. In order to completely reconstruct the block image, $4n^2$ coefficients must be used. The new algorithm can completely reconstruct the block image when n^2 linearly independent bases are used.

4. Numerical Experiments

In this paper we add Gaussian white noise to Lenna and Barbara images with size 256×256 . Similar to reference [4], we apply the hard-thresholding to it and use Monte-Carlo simulations to get the threshold. Let y_λ be the noisy coefficients, σ is noise standard deviation, $\overline{\sigma_\lambda}$ are estimated by evaluating curvelet transform of a few standard white noise images.

$$\overline{y_\lambda} = \begin{cases} y_\lambda & \text{if } |y_\lambda|/\sigma \geq k\overline{\sigma_\lambda} \\ 0 & \text{if } |y_\lambda|/\sigma < k\overline{\sigma_\lambda} \end{cases}.$$

In our experiments, we choose a subband-dependent value for k ; we have $k = 3$ for the second subband ($s = 2$) while $k = 4$ for the third subband ($s = 3$).

The following figures show the results, where the signal-to-noise ratios unit is dB. The new algorithm outperforms original algorithm by 0.72db for Lenna and 0.78db for Barbara images. We can see that the edges and textures are much clearer by the new algorithm and the visual effect is greatly improved.

5. Conclusion

A new digital curvelet reconstruction algorithm is proposed. Instead of applying the generalized Radon inverse transform in [3,4], we compute the orthogonal projection of the original block image into the space spanned by the thresholded ridgelet bases. Because the orthogonal projection is the best approximation of the origin image, so the new algorithm provides sparser representations when applying it to the nonlinear approximation. The experiments of the image denoising show that the new algorithm exhibits much higher quality and SNR than the original algorithm and the visual effect is surprisingly good.



Figure 1: (Top left) Original image and (top right) noisy image 13.95db, (bottom left) denoising by the new algorithm 20.08db and (bottom right) denoising by the original algorithm 19.36db



Figure 2: (Top left) Original image and (top right) noisy image 13.95db, (bottom left) denoising by the new algorithm 18.81db and (bottom right) denoising by the original algorithm 18.03db

References

- [1] E. J. Candes and D. L. Donoho, Curvelets—A Surprising Effective Nonadaptive Representation For Objects With Edges. *Curve and Surface Fitting: Saint-Malo 1999*, A. Cohen, C. Rabut, and L. Schumaker, eds., Vanderbilt University Press, (Nashville, TN), 1999.
- [2] D. L. Donoho, Orthonormal Ridgelet and Linear Singularities. *SIAM J. Math. Anal.*, 31(5), 1062-1099, 2000.
- [3] D. L. Donoho and M.R. Duncan, Digital curvelet Transform: Strategy, Implementation and Experiments. *Proc. SPIE*, 4056, 12-29, 2000.
- [4] J. L. Starck, E. J. Candes, D. L. Donoho, The Curvelet Transform for Image Denoising. *IEEE Trans. Image Processing*, 11(6), 670-684, 2002.
- [5] D. L. Donoho and Ana Georgina Flesia, Digital Ridgelet Transform based on True Ridge Functions. <http://www.stat.Stanford.edu/donoho/Report>.
- [6] A.Averbuch, R. R. Coifman, D.L. Donoho, M. Israeli, J. Walden, Fast Slant Stack: A notion of Radon Transform for Data in a Cartesian Grid which is Rapidly Computible, Algebraically Exact, Geometrically Faithful and Invertible. <http://www-stat.stanford.edu/~beamlab>
- [7] K. Grochenig, Acceleration of Frame Algorithm. *IEEE Trans. Signal Processing*, 41(12), 3331-3340, 1993.

Jian Bai
Department of Applied Mathematics
Xidian University
Xi'an
710071
China
e-mail: keywhite26@126.com

Xiang-Chu Feng
Department of Applied Mathematics
Xidian University
Xi'an
710071
China

Condition Number for Under-Determined Toeplitz Systems

Huaian Diao and Yimin Wei

Abstract. In this note, we prove that the structured condition number for Toeplitz under-determined systems with full row rank is not better than unstructured condition number in probability sense.

Mathematics Subject Classification (2000). Primary 15A09; Secondary 65F15, 65F10.

Keywords. Moore-Penrose inverse, Structured condition number, Under-determined Toeplitz systems.

1. Introduction

In signal and image processing, by using wavelet [1, 2, 3], the resulting system sometimes is singular. There are few papers concerning the singular linear system. In this note, we show that the structured condition number for Toeplitz under-determined systems with full row rank is not better than unstructured condition number in probability sense.

Structured matrices, for example, Toeplitz, Hankel, Vandermonde and Cauchy matrices, are very important in applied mathematics, computer science and engineering [4], such as in system theory, signal processing, network theory and linear prediction etc.

The problem for structured matrices, such as structured low rank approximation [5], is difficult because the matrix structure should be considered. Recent developments in structured matrices can be founded in [6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

This work was completed with the support of the National Natural Science Foundation of China under grant No. 10471027 and Shanghai Education Committee.
Correspondence to: Yimin Wei, E-mail: ymwei@fudan.edu.cn.

Jin and Yuan [10] studied the first kind integral equation

$$\int_0^{+\infty} k(x-y)\sigma(y)dy = g(x)$$

by the wavelet method. The integral equation is discretized with respect to two different wavelet bases, one of them is a Toeplitz system, and the other one is a system with condition number $\kappa = \mathbf{O}(1)$ after a diagonal scaling.

Condition numbers measure the sensitivity of the output of a problem with respect to small perturbations of the input data. Now let us consider the following nonsingular linear system

$$Ax = b,$$

where $A \in \mathbf{R}^{n \times n}$ and $b \in \mathbf{R}^n$.

As we know the linear system has unique solution $x = A^{-1}b$. When there are small perturbations caused by algorithm on both A and b , how does x change? This has been studied intensively. Suppose that the perturbations on A and b are ΔA and Δb respectively, satisfying $\|\Delta A\|_2 \leq \epsilon \|A\|_2$, and $\|\Delta b\|_2 \leq \epsilon \|b\|_2$, and $A + \Delta A$ remains nonsingular. Let $x + \Delta x$ be the solution of the perturbed system

$$(A + \Delta A)(x + \Delta x) = b + \Delta b,$$

then we have [16, 17],

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \kappa_{A,b}(A, b) \frac{\epsilon}{1 - \epsilon \kappa(A)},$$

where

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2, \quad \kappa_{A,b}(A, b) = \|A\|_2 \|A^{-1}\|_2 + \frac{\|A^{-1}\|_2 \|b\|_2}{\|x\|_2}$$

which is proved to be the condition number in [18]. Here $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2$,

where $\|x\|_2$ is the spectral norm of x .

A condition number plays an important role in numerical analysis [19]. The condition number describes the sensitivity of the linear system solution x with respect to the perturbations on A and b . It is also important for the rectangular linear systems.

First, let us introduce the definition of Moore-Penrose inverse for rectangular matrices. For any $A \in \mathbf{R}^{m \times n}$, there exists a unique $n \times m$ complex matrix A^\dagger , called *Moore-Penrose inverse* of A , satisfying the following four matrix equations [20, 21, 22]

$$AXA = A, \quad XAX = X, \quad (AX)^H = AX, \quad (XA)^H = XA.$$

Here, for a complex matrix M , M^H denotes its conjugate transpose matrix.

It follows from the singular value decomposition [23] that we can write A as

$$A = U \begin{bmatrix} D & 0 \end{bmatrix} V^H, \quad (1.1)$$

where $U \in \mathbf{R}^{m \times m}$ and $V \in \mathbf{R}^{n \times n}$ are unitary matrices, and $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m)$ where $m = \text{rank}(A)$ and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m > 0$ are the singular values of A . It is easy to check that A^\dagger can be expressed by:

$$A^\dagger = V \begin{bmatrix} D^{-1} \\ 0 \end{bmatrix} U^H. \quad (1.2)$$

If the coefficient matrix of linear system is rectangular with full row rank, then the problem becomes *under-determined system* [24]. This consists of, given $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$, finding $x \in \mathbf{R}^n$ to solve the following under-determined system [24]

$$Ax = b, \quad (1.3)$$

with the minimum spectral norm solution. It is well known that

$$x = A^\dagger b, \quad (1.4)$$

where A^\dagger is the Moore-Penrose inverse of A .

Denote $x + \Delta x$ the minimum spectral norm solution to

$$(A + \Delta A)(x + \Delta x) = b + \Delta b. \quad (1.5)$$

with $\|\Delta A\|_2 \leq \epsilon \|A\|_2$, $\|\Delta b\|_2 \leq \epsilon \|b\|_2$ and ϵ small enough. And

$$x + \Delta x = (A + \Delta A)^\dagger (b + \Delta b).$$

The unstructured condition number of (1.3) with respect to a weight matrix $E \in \mathbf{R}^{m \times n}$ and a weight vector $f \in \mathbf{R}^n$ is defined by

$$\kappa_{E,f}(A, b) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|\Delta x\|_2}{\epsilon \|x\|_2} : \begin{aligned} &(A + \Delta A)(x + \Delta x) = b + \Delta b, \\ &\Delta A \in \mathbf{R}^{m \times n}, \Delta b \in \mathbf{R}^n, \\ &\|\Delta A\|_2 \leq \epsilon \|E\|_2, \|\Delta b\|_2 \leq \epsilon \|f\|_2 \end{aligned} \right\}. \quad (1.6)$$

When A is a structured matrix, such as Toeplitz matrix, which is constant along its diagonals, for examples, 3×4 Toeplitz matrix is given by

$$A = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ a_{-1} & a_0 & a_1 & a_2 \\ a_{-2} & a_{-1} & a_0 & a_1 \end{bmatrix},$$

where $a_i \in \mathbf{C}$, $i = -2, -1, \dots, 3$. We expect that the small perturbation matrix ΔA on A keeping the same structures as A . The condition number for structured nonsingular matrices has been studied in [25, 26]. Rump obtained the explicit formula of *structured condition number* and got the ratio between structured condition number and unstructured condition number. Structured condition number can also be defined similar to (1.6):

$$\kappa_{E,f}^{struct}(A, b) = \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|\Delta x\|_2}{\varepsilon \|x\|_2} : \begin{aligned} &(A + \Delta A)(x + \Delta x) = b + \Delta b, \\ &\Delta A \in \mathbf{R}_{struct}^{m \times n}, \Delta b \in \mathbf{R}^n, \\ &\|\Delta A\|_2 \leq \varepsilon \|E\|_2, \|\Delta b\|_2 \leq \varepsilon \|f\|_2 \end{aligned} \right\}, \quad (1.7)$$

where $\mathbf{R}_{struct}^{m \times n}$ denotes $m \times n$ structured matrices.

In practice, we hope that the structured condition number is smaller than the unstructured one. But from many numerical experiments, this does not seem to be true. In this note, we will show that the structured condition number is rarely better than unstructured condition number for full row rank under-determined Toeplitz system from the probability viewpoint.

2. Main Result

In the following, we first seek the upper bound for unstructured condition number. It is well known that Golub and Van Loan [16, Theorem 5.7.1] presented the following results.

Lemma 2.1. *Let $A \in \mathbf{R}^{m \times n}$ and $0 \neq b \in \mathbf{R}^m$. Suppose that $\text{rank}(A) = m < n$ and $\Delta A \in \mathbf{R}^{m \times n}$ and $\Delta b \in \mathbf{R}^m$ satisfy*

$$\epsilon = \max\{\|\Delta A\|_2/\|A\|_2, \|\Delta b\|_2/\|b\|_2\} < \sigma_m(A).$$

Then

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq 3\kappa_2(A)\epsilon + \mathbf{O}(\epsilon^2), \quad (2.1)$$

where $\kappa_2(A) = \|A\|_2\|A^\dagger\|_2$.

From Lemma 2.1 the upper bound for the unstructured condition number of the under-determined system (1.3) can be obtained easily.

Theorem 2.2. *For the under-determined system (1.3), when the weighted matrix $E = A$ and weighted vector $f = b$, the unstructured condition number for (1.3) has an upper bound*

$$\kappa_{A,b}(A, b) \leq 3\kappa_2(A).$$

Proof. From (2.1) and the definition of unstructured condition number (1.6), let $\epsilon \rightarrow 0$, then we draw the conclusion.

In this paper we only focus on the case that the structure is Toeplitz matrix. For nonsingular Toeplitz matrices, Böttcher and Grudsky [25] showed that the structured condition number is rarely better than usual condition number. Now how about the structured condition number for structured under-determined system? Is it always smaller than unstructured case? We will answer this question in the following context.

The following theorem gives a new lower bound of $\kappa_{A,b}^{\text{Toep}}(A, b)/\kappa_{A,b}(A, b)$. Using the similar method in [25], we can prove the following lemma for rectangular matrices.

Lemma 2.3. *Let $0 \neq x = (x_0, x_1, \dots, x_{n-1})^T \in \mathbf{R}^n$ and suppose that the polynomial $x(t) = x_0 + x_1 t + \dots + x_{n-1} t^{n-1}$ has exactly ℓ zeros (counted with multiplicities) on \mathbf{T} , here \mathbf{T} is the unite circle. Thus,*

$$x(t) = \Pi_{j=1}^{\ell} (t - u_j) z(t), \quad u_j \in \mathbf{T}, \quad \mathbf{z}(t) \neq \mathbf{0} \quad \text{for } t \in \mathbf{T},$$

where $z(t)$ is a polynomial of degree $n - \ell - 1$. Put $\min |z| = \min_{t \in \mathbf{T}} |z(t)|$. Then there exists a Toeplitz matrix T_0 such that

$$T_0 x = u_m, \quad \|T_0\|_2 \leq 2^{\ell/2} m^{\ell+1/2} \frac{1}{\min_{t \in \mathbf{T}} |z(t)|} := M,$$

where u_m is the m -th column of unitary matrix U in (1.1).

Now let us investigate the lower bound for the structured condition number for under-determined system (1.3).

Theorem 2.4. *As the notation above, we have*

$$\frac{\kappa_{A,b}^{\text{Toep}}(A, b)}{\kappa_{A,b}(A, b)} \geq \frac{\min |z|}{2^{\ell/2} m^{\ell+1/2}} \frac{1}{3\|x\|_2}.$$

Proof. Since $x = A^\dagger b = A^T (AA^T)^{-1} b$, from (1.5) we have

$$x + \Delta x = (A + \Delta A)^T [(A + \Delta A)(A + \Delta A)^T]^{-1} (b + \Delta b).$$

Then we obtain that

$$\Delta x = (I - A^\dagger A)(\Delta A)^T (AA^T)^{-1} b + A^\dagger (\Delta b - \Delta A x) + \mathbf{O}(\epsilon^2). \quad (2.2)$$

Taking spectral norm and noting that $\mathbf{Im}(A^\dagger) \perp \mathbf{Im}(I - A^\dagger A)$, we have

$$\begin{aligned} \|\Delta x\|_2 &= \|(I - A^\dagger A)(\Delta A)^T (AA^T)^{-1} b\|_2 + \|A^\dagger (\Delta b - \Delta A x)\|_2 + \mathbf{O}(\epsilon^2) \\ &\geq \|A^\dagger (\Delta b - \Delta A x)\|_2 + \mathbf{O}(\epsilon^2). \end{aligned} \quad (2.3)$$

From Lemma 2.3, choosing $\Delta A_0 = \epsilon \|A\|_2 \frac{1}{M} T_0 \in \mathbf{R}_{struct}^{m \times n}$ and $\Delta b_0 = 0$, where $M = 2^{\ell/2} m^{\ell+1/2} \frac{1}{\min_{t \in \mathbf{T}} |z(t)|}$, we know $\|\Delta A_0\|_2 \leq \|A\|_2$, and from (2.3), we have

$$\begin{aligned}
 \kappa_{A,b}^{\text{Toep}}(A, b) &= \lim_{\epsilon \rightarrow 0} \sup_{\substack{\|\Delta A\|_2 \leq \epsilon \|A\|_2 \\ \|\Delta b\|_2 \leq \epsilon \|b\|_2}} \frac{\|\Delta x\|_2}{\epsilon \|x\|_2} \\
 &\geq \lim_{\epsilon \rightarrow 0} \left(\frac{\|A^\dagger \Delta A_0 x\|_2}{\epsilon \|x\|_2} + \mathbf{O}(\epsilon) \right) \\
 &= \left\| A^\dagger \frac{\|A\|_2 T_0}{M} x \right\|_2 \frac{1}{\|x\|_2} \\
 &= \frac{1}{M} \|A^\dagger T_0 x\|_2 \|A\|_2 \frac{1}{\|x\|_2} \\
 &= \frac{1}{M} \|A^\dagger u_n\|_2 \|A\|_2 \frac{1}{\|x\|_2} \\
 &= \frac{1}{M} \|A^\dagger\|_2 \|A\|_2 \frac{1}{\|x\|_2} \\
 &= \frac{\kappa_2(A)}{M \|x\|_2}.
 \end{aligned}$$

Thus we have estimated the lower bound for structured condition number

$$\kappa_{A,b}^{\text{Toep}}(A, b) \geq \frac{\kappa_2(A)}{2^{\ell/2} m^{\ell+1/2} \frac{1}{\min_{t \in \mathbf{T}} |z(t)|} \|x\|_2}.$$

From Theorem 2.2, we get the ratio lower bound between structured condition number and unstructured condition number

$$\frac{\kappa_{A,b}^{\text{Toep}}(A, b)}{\kappa_{A,b}(A, b)} \geq \frac{\min |z|}{2^{\ell/2} m^{\ell+1/2}} \frac{1}{3 \|x\|_2}.$$

Now we can get the probability analysis of the structured condition number in the next theorem.

Theorem 2.5. *Let $x_0, x_1, \dots, x_{n-1} \in \mathbf{C}$ be independent standard normal variables and put $x = (x_0, x_1, \dots, x_{n-1})^H$. There exists universal positive constant ϵ and n_0 such that*

$$\text{Prob} \left(\frac{\kappa_{A,b}^{\text{Toep}}(A, b)}{\kappa_{A,b}(A, b)} \geq \frac{\epsilon}{n^{3/2}} \right) > \frac{99}{100},$$

for $n \geq n_0$.

Proof. Let $u_0, v_0, \dots, u_{n-1}, v_{n-1}$ be independent standard normal. Then $x_0 = u_0 + \mathbf{i}v_0, \dots, x_{n-1} = u_{n-1} + \mathbf{i}v_{n-1}$ and let

$$u(t) = u_0 + u_1 t + \dots + u_{n-1} t^{n-1}, \quad v(t) = v_0 + v_1 + \dots + v_{n-1} t^{n-1}, \quad x(t) = u(t) + \mathbf{i}v(t),$$

where $\mathbf{i} = \sqrt{-1}$.

In [27], Konyagin and Schlag proved that there is a universal constant $D \in (0, \infty)$ such that

$$\limsup_{m \rightarrow \infty} \text{Prob} \left(\min_{t \in \mathbf{T}} |u(t)| \leq \frac{\delta}{\sqrt{n}} \right) \leq D\delta,$$

where δ is an arbitrary positive real number. Then we can choose small δ such that $D\delta < 0.004$. Then there exists an n_0 such that

$$\text{Prob} \left(\min_{t \in \mathbf{T}} |x(t)| < \frac{\delta}{\sqrt{n}} \right) \leq \text{Prob} \left(\min_{t \in \mathbf{T}} |u(t)| < \frac{\delta}{\sqrt{n}} \right) < 0.005, \quad \text{for any } n \geq n_0. \quad (2.4)$$

It follows from Theorem 2.4 that we have

$$\frac{\kappa_{A,b}^{\text{Toep}}(A, b)}{\kappa_{A,b}(A, b)} \geq \frac{1}{\|x\|_2} \min_{t \in \mathbf{T}} |x(t)| \frac{1}{3\sqrt{m}} \geq \frac{1}{\|x\|_2} \min_{t \in \mathbf{T}} |x(t)| \frac{1}{3\sqrt{n}}.$$

Since u_i, v_j are independent standard normal variables, we know $\|x\|_2^2 = \|u\|_2^2 + \|v\|_2^2$ is χ_{2n}^2 -distributed. From this the expected value $\mathbf{E}(\|x\|_2^2)$ is $2n$, and from Chebychev's inequality we arrive at

$$\text{Prob} \left(\frac{1}{\|x\|_2} \leq \frac{1}{20\sqrt{n}} \right) = \text{Prob} \left(\|x\|_2^2 \geq 400n \right) \leq \frac{2n}{400n} = 0.005. \quad (2.5)$$

Thus we have

$$\text{Prob} \left(\frac{1}{\|x\|_2} \leq \frac{1}{20\sqrt{n}} \right) \leq 0.005.$$

Then from (2.4) and (2.5), we obtain

$$\text{Prob} \left(\min_{t \in \mathbf{T}} |x(t)| < \frac{\delta}{\sqrt{n}} \quad \text{or} \quad \frac{1}{\|x\|_2} \leq \frac{1}{20\sqrt{n}} \right) < 0.005 + 0.005 = 0.01.$$

And so

$$\text{Prob} \left(\min_{t \in \mathbf{T}} |x(t)| \geq \frac{\delta}{\sqrt{n}} \quad \text{and} \quad \frac{1}{\|x\|_2} > \frac{1}{20\sqrt{n}} \right) > 0.99.$$

Therefore we know that

$$\begin{aligned} \text{Prob} \left(\frac{\kappa_{A,b}^{\text{Toep}}(A, b)}{\kappa_{A,b}(A, b)} \geq \frac{\delta}{60n^{3/2}} \right) &\geq \text{Prob} \left(\frac{1}{\|x\|_2} \min_{t \in \mathbf{T}} |x(t)| \frac{1}{3\sqrt{n}} \geq \frac{\delta}{60n^{3/2}} \right) \\ &> 0.99. \end{aligned}$$

If we choose $\epsilon = \frac{\delta}{60}$, then we complete the proof.

3. Concluding Remarks

In this note, we show that the structured condition numbers for Toeplitz under-determined systems with full row rank are rarely better than unstructured one in probability sense. In [28, 29], Diao and Wei investigated the structured condition number for the Drazin and group inverse solution of singular square linear systems [30] and got the expressions for the structured condition number. Xu, Wei

and Qiao [31] considered the condition numbers for rank-deficient structured least squares problem. We will compare the structured condition numbers with the unstructured one for the singular case in a forthcoming paper.

Acknowledgements. The authors would like to thank Prof. Tao QIAN and the referee for their valuable comments on our manuscript.

References

- [1] I. Daubechies, *Ten Lectures on Wavelets*, SIAM Press 1992.
- [2] Y. Meyer, *Wavelets - Algorithms and Applications*, SIAM Press, 1993.
- [3] Ke Chen, *Matrix Preconditioning Techniques and Applications*, Cambridge University Press, 2005.
- [4] R. Chan and M. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Review, **38**(1996), 427–482.
- [5] M. Chu, R. Funderlic and R. Plemmons, *Structured low rank approximation*, Linear Algebra Appl., **366**(2003), 157–172.
- [6] R.H. Chan, J.G. Nagy and R.J. Plemmons, *Circulant preconditioned for Toeplitz least squares iterations*, SIAM J. Matrix Anal., **15**(1994), 80–97.
- [7] G. Heinig, *Fast algorithms for Toeplitz least squares problems*, Current Trends in Operator Theory and its Applications, pp.167-197, Oper. Theory Adv. Appl., 149, Birkhauser, Basel, 2004.
- [8] X. Jin, *Developments and Applications of Block Toeplitz Iterative Solvers*, Kluwer Academic Publishers and Science Press, Beijing, 2002.
- [9] X. Jin and Y. Wei, *Numerical Linear Algebra and its Applications*, Science Press, Beijing, 2004.
- [10] X. Jin and J. Yuan, *A wavelet method for the first kind integral equations with kernel $k(x - y)$* , Taiwanese Journal of Mathematics, **2**(1998), 427-434.
- [11] M. Ng, *Iterative Methods for Toeplitz Systems*, Oxford University Press, New York, 2004.
- [12] V. Olshevsky (Editor), *Structured Matrices in Mathematics, Computer Science, and Engineering I and II*, Contemporary Mathematics, Vol. 280 and 281, AMS Publications, 2001.
- [13] V. Olshevsky (Editor), *Fast Algorithms for Structured Matrices: Theory and Applications*, Contemporary Mathematics, Vol. 323, AMS Publications, 2003.
- [14] H. Park and L. Elden, *Schur-type methods for solving least squares problems with Toeplitz structure*, SIAM J. Sci. Comput., **22**(2000), 406–430.
- [15] Y. Wei, J. Cai and M. Ng, *Computing Moore-Penrose inverses of Toeplitz matrix by Newton's iteration*, Math. Comput. Modelling, **40**(2004), 181–191.
- [16] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd Edition, Johns Hopkins University Press, 1996.
- [17] S.M. Rump, *Structured perturbations Part I: Normwise distances*, SIAM J. Matrix Anal. Appl., **25**(2003), 1–30.

- [18] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd Edition, SIAM, Philadelphia, 2002.
- [19] F. Cucker, H. Diao and Y. Wei, *Smoothed analysis of some condition numbers*, Numer. Linear Algebra Appl., **13**(2006), 71–84.
- [20] A. Ben-Israel, and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, 2nd Edition, Springer-Verlag, New York, 2003.
- [21] S.L. Campbell and C.D. Meyer, *Generalized Inverse of Linear Transformations*, Pitman, London, 1979; Dover, New York, 1991.
- [22] G. Wang, Y. Wei, and S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing, 2004.
- [23] G.W. Stewart and J. Sun, *Matrix Perturbation Analysis*, Academic Press, New York, 1990.
- [24] J. Demmel and N. Higham, *Improved error bounds for underdetermined system solvers*, SIAM J. Matrix Anal. Appl., **14** (1993), 1–14.
- [25] A. Böttcher, S.M. Grudsky, *Structured condition numbers of large Toeplitz matrices are rarely better than usual condition numbers* Numer. Linear Algebra Appl., **12**(2005), 95–102.
- [26] S.M. Rump, *Structured perturbation Part II: Componentwise distances*, SIAM J. Matrix Anal. Appl., **25**(2003), 31–56.
- [27] S.V. Konyagin and W. Schlag, *Lower bounds for the absolute value of random polynomials on a neighborhood of the unit circle*, Trans. Amer. Math. Soc., **351**(1999), 4963–4980.
- [28] H. Diao, *Structured perturbations of Drazin inverse*, Appl. Math. Comput., **158** (2004), 419–432.
- [29] H. Diao and Y. Wei, *Structured perturbations of group inverse and singular linear system with index one*, J. Comput. Appl. Math., **173**(2005), 93–113.
- [30] Y. Wei and H. Wu, *Convergence properties of Krylov subspace methods for singular linear systems with arbitrary index*, J. Comput. Appl. Math., **114** (2004), 305–318.
- [31] W. Xu, Y. Wei and S. Qiao, *Condition numbers for structured least squares*, BIT, **46** (2006), 203–225.

Huaian Diao
City University of Hong Kong
83 Tat Chee Avenue, Kowloon Tong
Hong Kong, P.R. of China
e-mail: 50007445@student.cityu.edu.hk

Yimin Wei
School of Mathematical Sciences
Fudan University
Shanghai, 200433, P.R. of China
& Key Laboratory of Mathematics for Nonlinear Sciences
(Fudan University), Ministry of Education.
e-mail: ymwei@fudan.edu.cn

Powell–Sabin Spline Prewavelets on the Hexagonal Lattice

Jan Maes and Adhemar Bultheel

Abstract. In this paper we give an explicit construction of compactly supported prewavelets on differentiable, twodimensional, piecewise polynomial quadratic finite element spaces of $L_2(\mathbb{R}^2)$, sampled on the hexagonal grid. The obtained prewavelet basis is stable in the Sobolev spaces \mathcal{H}^s for $|s| < \frac{5}{2}$. In particular, the prewavelet basis is generated by one single function vector ψ consisting of three generating functions ψ_1, ψ_2, ψ_3 that are globally invariant by a rotation of $2\pi/3$.

Mathematics Subject Classification (2000). Primary 65T60; Secondary 65D07.

Keywords. Powell–Sabin splines, prewavelets, local support.

1. Introduction

Spline multiresolution analysis has become a standard mathematical tool in a large number of areas including signal processing and image compression, computer graphics, and numerical solutions of differential and integral equations. The construction of multivariate wavelets on arbitrary triangulations is very challenging. In fact, even the case of continuous piecewise linear wavelets construction is unexpectedly complicated, see [6, 7, 8, 9, 10, 11, 17]. In this paper, we present an explicit construction of C^1 continuous piecewise quadratic prewavelets with compact support on the hexagonal lattice in \mathbb{R}^2 . Our main motivation comes from the numerical treatment of elliptic variational problems.

Let \mathcal{H}^s , $s \in \mathbb{R}$ denote the usual Sobolev space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}^s}$, and let $a(\cdot, \cdot)$ be a symmetric bilinear form such that $a(v, v) \sim \|v\|_{\mathcal{H}^s}^2$. We always

This work is partially supported by the Flemish Fund for Scientific Research (FWO Vlaanderen) project MISS (G.0211.02), and by the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with the authors.

mean by $A \sim B$ that A can be bounded above and below by constant multiples of B . Let \mathcal{H}^{-s} denote the dual of \mathcal{H}^s . Consider the elliptic variational problem:

$$\text{Given } f \in \mathcal{H}^{-s}, \text{ find } u \in \mathcal{H}^s \text{ such that } a(u, v) = f(v), \quad \forall v \in \mathcal{H}^s. \quad (1.1)$$

Suppose V is a finite dimensional subspace of \mathcal{H}^s with basis $\{B_j : j = 1, \dots, N\}$. The Galerkin approximate solution $u_V = \sum_{j=1}^N c_j B_j$ is the unique solution of (1.1) with \mathcal{H}^s replaced by V and it can be found by solving the linear system

$$(a(B_i, B_j))_{i,j=1}^N (c_j)_{j=1}^N = (f(B_i))_{i=1}^N. \quad (1.2)$$

The stiffness matrix $A = (a(B_i, B_j))_{i,j=1}^N$ for a typical nodal basis is ill-conditioned. Suppose that there exist constants $0 < \underline{\alpha}, \bar{\alpha} < \infty$ (Riesz bounds) such that

$$\underline{\alpha} \sum_{j=1}^N |c_j|^2 \leq \left\| \sum_{j=1}^N c_j B_j \right\|_{\mathcal{H}^s}^2 \leq \bar{\alpha} \sum_{j=1}^N |c_j|^2, \quad (1.3)$$

then it is well-known that $\kappa(A) = \mathcal{O}(\bar{\alpha}/\underline{\alpha})$, [2]. So let $\{B'_j : j = 1, \dots, N\}$ be another basis for V . Then we can solve the linear system (1.2) with the basis functions B_j replaced by B'_j . The new system can be considered to arise from (1.2) by preconditioning. Hence, we are interested in basis functions B'_j for which $\kappa((a(B'_i, B'_j))_{i,j=1}^N) = \mathcal{O}(1)$.

Several (pre-)wavelet constructions on polygonal domains generate Riesz bases on \mathcal{H}^s for s in a range around zero, see e.g. [4, 18]. However, for most constructions at least a subset of the wavelets is only continuous, so that the range of stability is restricted to $s < \frac{3}{2}$. The only available wavelet Riesz basis on general polygons for $s \geq \frac{3}{2}$ is the basis constructed in [3], but, from a practical point of view, this basis is very hard to construct. The wavelet basis that we present in this paper is easy to construct, and it forms a Riesz basis for \mathcal{H}^s , $|s| < \frac{5}{2}$, which makes it a useful basis for preconditioning fourth order elliptic problems. It is however restricted to uniform triangulations of the hexagonal grid. Another construction that provides a Riesz basis for \mathcal{H}^s , $|s| < \frac{5}{2}$, is the cubic spline prewavelet basis on the uniform four-directional mesh [1]. The dimension of the cubic spline space in [1] is larger than the Powell–Sabin spline space that we use here, which gives more degrees of freedom in constructing prewavelets with certain properties, but at the cost of the construction complexity. The extra degrees of freedom are used to obtain symmetry properties. In comparison, the Powell–Sabin spline prewavelets have hexagonal symmetry for free.

Section 2 gives a short overview of Powell–Sabin splines on the hexagonal lattice and introduces a stable basis. In Section 3 we construct Powell–Sabin spline prewavelets using the two-step construction method from [17]. Finally we visualize the prewavelets in Section 4.

2. Preliminaries

Consider the hexagonal lattice Δ in \mathbb{R}^2 which is defined as the image of \mathbb{Z}^2 by a linear transformation corresponding to the matrix

$$\Gamma = \begin{bmatrix} 1 & -1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix},$$

and let Δ^* be its refinement by drawing in the additional grid lines $y = l$, $y = \frac{\sqrt{3}}{3}(x+m)$, and $y = -\frac{\sqrt{3}}{3}(x+n)$, $l, m, n \in \mathbb{Z}$. In fact, Δ^* is the Powell–Sabin 6-split of Δ , see Figure 1. Define $S_2^1(\Delta^*)$ as the space of real-valued functions in $C^1(\mathbb{R}^2)$ whose restrictions on each triangle of the triangulation Δ^* are bivariate quadratic polynomials. Then each function $\phi \in S_2^1(\Delta^*)$ is called a uniform Powell–Sabin (PS) spline.

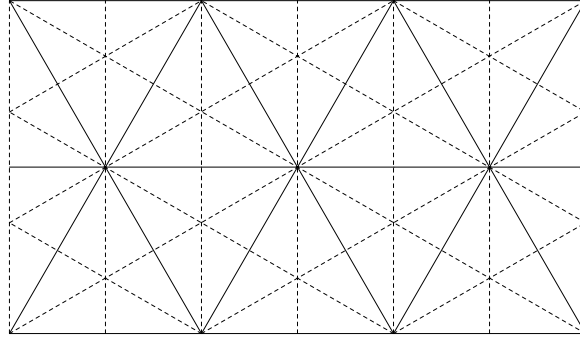


FIGURE 1. Hexagonal lattice Δ (black lines) with Powell–Sabin 6-split Δ^* (black and dotted lines)

Let $k \in \Delta$, then the interpolation problem

$$\left[\phi(l), \frac{\partial}{\partial x} \phi(l), \frac{\partial}{\partial y} \phi(l) \right] = \delta_{k,l} \begin{bmatrix} \alpha, \beta, \gamma \end{bmatrix}, \quad l \in \Delta, \quad (2.1)$$

has a unique solution $\phi \in S_2^1(\Delta^*)$, see [16]. This allows to define a function vector $\phi = [\phi_1, \phi_2, \phi_3]^T$ that generates a multiresolution analysis (MRA). Let each ϕ_i , $i = 1, 2, 3$, be the unique solution of (2.1) with

$$\alpha = \frac{1}{3}, \quad \beta = \frac{2}{3} - \delta_{i1} \frac{6}{3}, \quad \gamma = ((-1)^{i-1} - \delta_{i1}) \frac{2\sqrt{3}}{3},$$

then the integer translates under Γ of the basis functions ϕ_i form a basis for $S_2^1(\Delta^*)$. Furthermore they form a partition of unity, see [5]. It can be checked that the three basis functions ϕ_1, ϕ_2, ϕ_3 are invariant by a rotation of $2\pi/3$.

Define the 2×2 dilation matrix D given by

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

then we consider the refinement $\Delta_j := D^{-j}\Delta$, which can be obtained by midedge subdivision, and the corresponding PS 6-split $\Delta_j^* := D^{-j}\Delta^*$. This yields nested subspaces $V_j = S_2^1(\Delta_j^*) \subset L_2(\mathbb{R}^2)$, $j \in \mathbb{Z}$, such that

$$V_j \subset V_{j+1}, \quad j \in \mathbb{Z}, \quad (2.2)$$

the closure of their union is $L_2(\mathbb{R}^2)$, and their intersection contains only the zero function. In general, the basis functions on all standard refinements Δ_j of Δ can be written as

$$\phi_{j,k}(u) = \phi(D^j(u - k)), \quad k \in \Delta_j, \quad u \in \mathbb{R}^2,$$

and the set $\{2^j \phi_{j,k} \mid k \in \Delta_j\}$ forms a Riesz basis of V_j , i.e., for $\mathbf{c} = \{\mathbf{c}_k\}_{k \in \Delta_j} \in l_2^{3 \times 1}(\Delta_j)$, $\mathbf{c}_k := (c_{1,k}, c_{2,k}, c_{3,k})^T$,

$$\left\| \sum_{k \in \Delta_j} \mathbf{c}_k^T 2^j \phi_{j,k} \right\|_{L_2(\mathbb{R}^2)} \sim \|\mathbf{c}\|_{l_2^{3 \times 1}(\Delta_j)}, \quad (2.3)$$

see [13, 14]. With $l_2^{3 \times 1}(\Delta_j)$ we denote the Banach space of all sequences of 3×1 vectors \mathbf{c}_k for which $\sqrt{\sum_{k \in \Delta_j} \|\mathbf{c}_k\|_2^2} < \infty$.

Because of the properties listed above we say that the sequence of closed subspaces $\{V_j\}_{j \in \mathbb{Z}}$ of $L^2(\mathbb{R}^2)$ forms a MRA of multiplicity 3, and the function vector ϕ is called scaling vector. The nestedness (2.2) of the MRA implies that ϕ needs to satisfy a matrix refinement equation of the form

$$\phi(u) = \sum_{k \in \mathbb{Z}^2} \mathbf{A}_k \phi(Du - \Gamma k), \quad u \in \mathbb{R}^2, \quad (2.4)$$

where \mathbf{A}_k are 3×3 mask coefficient matrices. Moreover, $\mathbf{A}_{(-1,-1)}$ and $\mathbf{A}_{(0,-1)}$ are given by

$$\frac{1}{4} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \frac{1}{4} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix},$$

$\mathbf{A}_{(-1,0)}$, $\mathbf{A}_{(0,0)}$ and $\mathbf{A}_{(0,1)}$ are given by

$$\frac{1}{4} \begin{bmatrix} 0 & 2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \frac{1}{6} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}, \quad \frac{1}{4} \begin{bmatrix} 0 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix},$$

and $\mathbf{A}_{(0,1)}$ and $\mathbf{A}_{(1,1)}$ are given by

$$\frac{1}{4} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}, \quad \frac{1}{4} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix},$$

see e.g. [21].

3. Prewavelets

We are interested in complement spaces W_j of V_j such that

$$V_{j+1} = V_j \oplus^{\perp_{L_2}} W_j,$$

hence

$$L_2(\mathbb{R}^2) = \bigoplus_{j \in \mathbb{Z}}^{\perp_{L_2}} W_j.$$

We will refer to the basis functions of W_j as prewavelets. Note that the prewavelets have three vanishing moments since constant, linear and quadratic polynomials are contained in the space V_j .

The construction of PS prewavelets consists of two steps, cfr. [17]. First we construct a dual basis

$$\{\tilde{\phi}_{j,k} \mid k \in \Delta_j\} \subset V_{j+1} \quad (3.1)$$

for the basis $\{\phi_{j,k} (\|\phi_{j,k}\|_{L_2}^2)^{-1} \mid k \in \Delta_j\}$ of V_j , i.e.

$$\langle \tilde{\phi}_{j,k}, \phi_{j,l}^T \rangle_{L_2} = \begin{cases} \|\phi_{j,k}\|_{L_2}^2 & k = l, \\ 0 & k \neq l, \end{cases} \quad (3.2)$$

with $\|\phi_{j,k}\|_{L_2}^2 := \text{diag}(\langle \phi_{j,k}, \phi_{j,k}^T \rangle_{L_2})$. Since we search for a dual basis in the space V_{j+1} which is much larger than V_j , this basis is not unique. We will use this freedom to select a dual basis that is locally supported. Then, as the second step, for $l \in \Delta_{j+1} \setminus \Delta_j$ we define

$$\psi_{j,l} := \phi_{j+1,l} - \sum_{k \in \Delta_j} \langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2} (\|\phi_{j,k}\|_{L_2}^2)^{-1} \tilde{\phi}_{j,k}. \quad (3.3)$$

Note that, because of (3.1) and (3.2), $\psi_{j,l} \in V_{j+1} \cap V_j^{\perp_{L_2}}$.

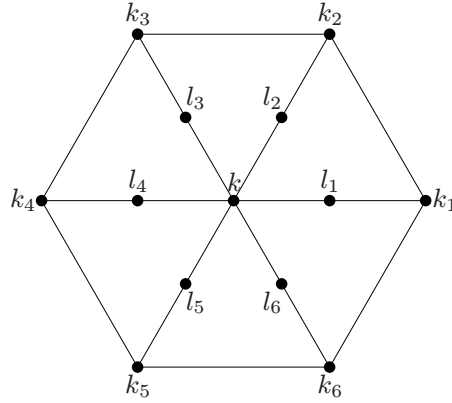


FIGURE 2. Grid around k

Let $k \in \Delta_j$ and define $k_1, \dots, k_6 \in \Delta_j$ and $l_1, \dots, l_6 \in \Delta_{j+1}$ as in Figure 2. We look for a dual function vector $\tilde{\phi}_{j,k}$ of the form

$$\tilde{\phi}_{j,k} := \tilde{\mathbf{A}}_0 \phi_{j+1,k} + \sum_{i=1}^6 \tilde{\mathbf{A}}_i \phi_{j+1,l_i}$$

that satisfies (3.2), where the $\tilde{\mathbf{A}}_i$ are 3×3 mask coefficient matrices. By construction we know that the support of each scaling vector $\phi_{j,k}$ is the set of all triangles in Δ_j that have $k \in \Delta_j$ as a common vertex. Note that, by construction, the support of $\tilde{\phi}_{j,k}$ equals the support of $\phi_{j,k}$, see Figure 2. Therefore it is sufficient to solve the system

$$\begin{cases} \langle \tilde{\phi}_{j,k}, \phi_{j,k}^T \rangle_{L_2} = \|\phi_{j,k}\|_{L_2}^2, \\ \langle \tilde{\phi}_{j,k}, \phi_{j,k_i}^T \rangle_{L_2} = 0 \end{cases}, \quad i = 1, \dots, 6, \quad (3.4)$$

to obtain a dual basis. We have 63 equations and 63 unknown matrix entries. Since the system (3.4) is non-singular we find a unique solution,

$$\begin{aligned} \tilde{\mathbf{A}}_0 &= \begin{bmatrix} \frac{382603}{41310} & -\frac{318649}{82620} & -\frac{318649}{82620} \\ -\frac{318649}{82620} & \frac{382603}{41310} & -\frac{318649}{82620} \\ -\frac{318649}{82620} & -\frac{318649}{82620} & \frac{382603}{41310} \end{bmatrix}, \\ \tilde{\mathbf{A}}_1 &= \begin{bmatrix} \frac{1691}{2754} & -\frac{551}{18360} & -\frac{551}{18360} \\ -\frac{551}{18360} & \frac{1691}{2754} & -\frac{551}{18360} \\ -\frac{551}{18360} & -\frac{551}{18360} & \frac{1691}{2754} \end{bmatrix}, \\ \tilde{\mathbf{A}}_2 &= \begin{bmatrix} \frac{10241}{82620} & \frac{4123}{18360} & -\frac{2204}{20655} \\ \frac{4123}{18360} & \frac{10241}{82620} & -\frac{2204}{20655} \\ -\frac{2204}{20655} & -\frac{2204}{20655} & \frac{10241}{82620} \end{bmatrix}, \\ \tilde{\mathbf{A}}_3 &= \begin{bmatrix} -\frac{608}{20655} & -\frac{779}{1836} & \frac{11761}{165240} \\ -\frac{779}{1836} & -\frac{608}{20655} & \frac{11761}{165240} \\ \frac{11761}{165240} & \frac{11761}{165240} & -\frac{608}{20655} \end{bmatrix}, \\ \tilde{\mathbf{A}}_4 &= \begin{bmatrix} \frac{6346}{20655} & -\frac{82783}{165240} & -\frac{82783}{165240} \\ -\frac{82783}{165240} & \frac{6346}{20655} & -\frac{82783}{165240} \\ -\frac{82783}{165240} & -\frac{82783}{165240} & \frac{6346}{20655} \end{bmatrix}, \\ \tilde{\mathbf{A}}_5 &= \begin{bmatrix} -\frac{608}{20655} & \frac{11761}{165240} & -\frac{779}{1836} \\ \frac{11761}{165240} & -\frac{608}{20655} & -\frac{779}{1836} \\ -\frac{779}{1836} & -\frac{779}{1836} & \frac{11761}{165240} \end{bmatrix}, \\ \tilde{\mathbf{A}}_6 &= \begin{bmatrix} \frac{10241}{82620} & -\frac{2204}{20655} & \frac{4123}{18360} \\ \frac{4123}{18360} & -\frac{2204}{20655} & \frac{10241}{82620} \\ -\frac{2204}{20655} & \frac{4123}{18360} & \frac{10241}{82620} \end{bmatrix}. \end{aligned}$$

Theorem 3.1. *The set*

$$\{2^{j+1}\tilde{\phi}_{j,k} \mid k \in \Delta_j\} \cup \{2^{j+1}\phi_{j+1,l} \mid l \in \Delta_{j+1} \setminus \Delta_j\} \quad (3.5)$$

is a Riesz basis for the space V_{j+1} .

Proof. We know that $\{2^{j+1}\phi_{j+1,k} \mid k \in \Delta_{j+1}\}$ is a Riesz basis for V_{j+1} , see (2.3). Furthermore we have that the basis $\{2^{j+1}\phi_{j+1,k} \mid k \in \Delta_{j+1}\}$ can be transformed into the set (3.5) by a matrix operation of the form

$$\mathbf{M} = \begin{bmatrix} \mathbf{K} & \mathbf{L} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix}, \quad (3.6)$$

where \mathbf{I}_n denotes the $n \times n$ identity matrix with $n = \#\{l \in \Delta_{j+1} \setminus \Delta_j\}$ and $\mathbf{0}$ is a zero matrix. The matrix \mathbf{K} has a block-structure of the form

$$\mathbf{K} = \begin{bmatrix} \tilde{\mathbf{A}}_0 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \tilde{\mathbf{A}}_0 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \tilde{\mathbf{A}}_0 & \\ \vdots & \vdots & & \ddots \end{bmatrix},$$

and the matrix \mathbf{L} is a sparse block-structured matrix with blocks $\tilde{\mathbf{A}}_i$, $i = 1, \dots, 6$. Note that

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{K}^{-1} & -\mathbf{K}^{-1}\mathbf{L} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix}.$$

Clearly \mathbf{M} is uniformly bounded and has full rank which proves our claim. \square

For all j , for all $l \in \Delta_{j+1} \setminus \Delta_j$, we define the prewavelets $\psi_{j,l}$ as in (3.3). Note that these prewavelets are compactly supported. It is easily checked that $W_j = \text{span}\{\psi_{j,l} \mid l \in \Delta_{j+1} \setminus \Delta_j\}$. Indeed, suppose $w \in W_j$. Then by Theorem 3.1 we have

$$w = \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \phi_{j+1,l} + \sum_{k \in \Delta_j} \mathbf{c}_k^T \tilde{\phi}_{j,k},$$

or

$$\begin{bmatrix} 0 & 0 & 0 \end{bmatrix} = \langle w, \phi_{j,k}^T \rangle_{L_2} = \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2} + \mathbf{c}_k^T \|\phi_{j,k}\|_{L_2}^2,$$

which yields

$$\mathbf{c}_k^T = - \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2} (\|\phi_{j,k}\|_{L_2}^2)^{-1}$$

and we get

$$w = \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \psi_{j,l}.$$

Theorem 3.2. *The set*

$$\{2^{j+1}\psi_{j,l} \mid l \in \Delta_{j+1} \setminus \Delta_j\} \quad (3.7)$$

is a Riesz basis for the space W_j .

Proof. It is sufficient to show that

$$\left\| \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T 2^{j+1} \psi_{j,l} \right\|_{L_2} \sim \|\mathbf{c}\|_{l_2(\Delta_{j+1} \setminus \Delta_j)}^{3 \times 1}.$$

By Theorem 3.1 and (3.3) it holds that

$$\begin{aligned} & \left\| \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T 2^{j+1} \boldsymbol{\psi}_{j,l} \right\|_{L_2}^2 \sim \sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \mathbf{c}_l \\ & + \sum_{k \in \Delta_l} \left(\sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2} (\|\phi_{j,k}\|_{L_2}^2)^{-1} \right)^T \\ & \cdot \left(\sum_{l \in \Delta_{j+1} \setminus \Delta_j} \mathbf{c}_l^T \langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2} (\|\phi_{j,k}\|_{L_2}^2)^{-1} \right). \end{aligned}$$

Both for fixed $k \in \Delta_j$ or for fixed $l \in \Delta_{j+1} \setminus \Delta_j$ the number of nonzero inner products $\langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2}$ is bounded above by an absolute constant. Furthermore it is straightforward to check that

$$0 \leq \langle \phi_{j+1,l}, \phi_{j,k}^T \rangle_{L_2} (\|\phi_{j,k}\|_{L_2}^2)^{-1} \leq C$$

with C an absolute constant. This concludes the proof. \square

Remark 3.3. Note that the prewavelets $\boldsymbol{\psi}_{j,l}$ and the dual functions $\tilde{\boldsymbol{\phi}}_{j,k}$ can be obtained as scaled translates of a finite subset of them. Indeed, we have that

$$\begin{aligned} \boldsymbol{\psi}_{j,l}(u) &= \boldsymbol{\psi}(D^j(u-l)), \quad l \in \Delta_{j+1} \setminus \Delta_j, \quad u \in \mathbb{R}^2, \\ \tilde{\boldsymbol{\phi}}_{j,k}(u) &= \tilde{\boldsymbol{\phi}}(D^j(u-k)), \quad k \in \Delta_j, \quad u \in \mathbb{R}^2, \end{aligned}$$

with $\boldsymbol{\psi}$ and $\tilde{\boldsymbol{\phi}}$ a generating wavelet vector resp. a generating dual scaling vector. Therefore the inner products appearing in the construction above have to be computed only once. As a consequence of the hexagonal symmetry of $\boldsymbol{\phi}$, both $\tilde{\boldsymbol{\phi}}$ and $\boldsymbol{\psi}$ have also hexagonal symmetry, i.e. the corresponding generating functions are invariant by a rotation of $2\pi/3$.

Remark 3.4. We are interested in determining the exact range of Sobolev exponents s for which the prewavelet basis forms a Riesz basis for \mathcal{H}^s . It turns out that the range of such s is determined by the Sobolev regularity $s(\boldsymbol{\phi})$ of the scaling vector $\boldsymbol{\phi}$, i.e. the wavelet system is a Riesz basis for \mathcal{H}^s for all s with $-s(\boldsymbol{\phi}) < s < s(\boldsymbol{\phi})$ and that this interval is sharp, cfr. [15]. It is well-known that PS splines are in the Sobolev spaces \mathcal{H}^s for any $s < 5/2$, see [13] for a detailed proof.

Remark 3.5. The above construction can be used to construct prewavelets on bounded domains in \mathbb{R}^2 due to the explicit construction that we use.

Remark 3.6. It is also possible to extend this construction to arbitrary polygonal domains in \mathbb{R}^2 . In that case, one should use a triadic refinement scheme instead of the dyadic scheme that we use here, see [20]. Unfortunately it remains an open problem whether the system (3.4) is non singular for arbitrary settings. However, using perturbation arguments, one can argue that a non-trivial solution to (3.4) should always exist in some sense.

Remark 3.7. A PS spline wavelet basis on arbitrary polygonal domains in \mathbb{R}^2 does exist, see [19]. In [12] it is shown that this wavelet basis is not a Riesz basis for L_2 . Fortunately it does form a Riesz basis for \mathcal{H}^s with s in a range around 2 which makes it suitable for preconditioning fourth order elliptic equations.

4. Visualization of the Functions

Finally we visualize a scaling function, a dual scaling function and a prewavelet, see Figures 3, 4, and 5.

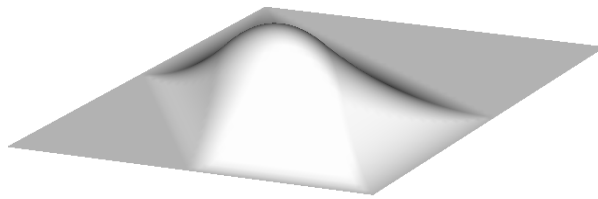


FIGURE 3. Scaling function

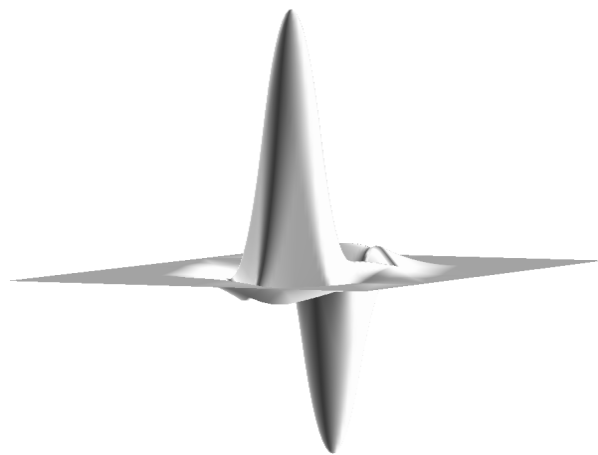


FIGURE 4. Dual scaling function

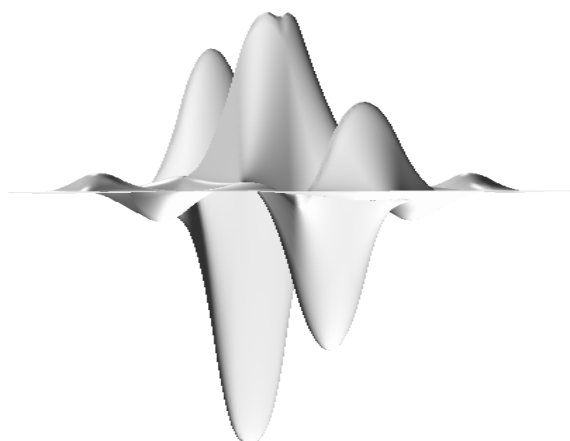


FIGURE 5. Prewavelet

References

- [1] M. D. Buhmann, O. Davydov, and T. N. T. Goodman. Cubic spline prewavelets on the four-directional mesh. *Foundations of Computational Mathematics*, 3:113–133, 2003.
- [2] W. Dahmen and A. Kunoth. Multilevel preconditioning. *Numer. Math.*, 63:315–344, 1992.
- [3] W. Dahmen and R. Schneider. Wavelets on manifolds I: Construction and domain decomposition. *SIAM J. Math. Anal.*, 31:184–230, 1999.
- [4] W. Dahmen and R. Stevenson. Element-by-element construction of wavelets satisfying stability and moment conditions. *SIAM J. Numer. Anal.*, 37(1):319–352, 2000.
- [5] P. Dierckx. On calculating normalized Powell–Sabin B-splines. *Comput. Aided Geom. Design*, 15(3):61–78, 1997.
- [6] G.C. Donovan, J. S. Geronimo, and D. P. Hardin. Intertwining multiresolution analysis and the construction of piecewise polynomial wavelets. *SIAM J. Math. Anal.*, 27:1791–1815, 1996.
- [7] G.C. Donovan, J. S. Geronimo, and D. P. Hardin. Compactly supported, piecewise affine scaling functions over triangulations. *Constr. Approx.*, 16:201–219, 2000.
- [8] M. S. Floater and E. G. Quak. Piecewise linear prewavelets on arbitrary triangulations. *Numer. Math.*, 82:221–252, 1999.
- [9] M. S. Floater and E. G. Quak. Piecewise linear prewavelets over Type-2 triangulations. *Comput. Supplement*, 14:89–103, 2001.
- [10] D. P. Hardin and D. Hong. Construction of wavelets and prewavelets over triangulations. *J. Comput. Appl. Math.*, 155:91–109, 2003.
- [11] U. Kotyczka and P. Oswald. Piecewise linear prewavelets of small support. In C. K. Chui and L. L. Schumaker, editors, *Approximation Theory VIII, Vol 2*, pages 235–242. World Scientific, Singapore, 1995.

- [12] J. Maes and A. Bultheel. Stable lifting construction of non-uniform biorthogonal spline wavelets with compact support. TW Report 437, Department of Computer Science, Katholieke Universiteit Leuven, Belgium, October 2005.
- [13] J. Maes and A. Bultheel. Stable multiresolution analysis on triangles for surface compression. 2005. Submitted.
- [14] J. Maes, E. Vanraes, P. Dierckx, and A. Bultheel. On the stability of normalized Powell–Sabin B-splines. *J. Comput. Appl. Math.*, 170(1):181–196, 2004.
- [15] P. Oswald. *Multilevel finite element approximation: Theory and applications*. B.G. Teubner, Stuttgart, 1994.
- [16] M. J. D. Powell and M. A. Sabin. Piecewise quadratic approximations on triangles. *ACM Trans. on Math. Software*, 3:316–325, 1977.
- [17] R. Stevenson. Piecewise linear (pre-)wavelets on non-uniform meshes. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, pages 306–319. Springer-Verlag, Heidelberg, 1996.
- [18] R. Stevenson. Locally supported, piecewise polynomial biorthogonal wavelets on non-uniform meshes. *Constr. Approx.*, 19(4):477–508, 2003.
- [19] E. Vanraes, J. Maes, and A. Bultheel. Powell–Sabin spline wavelets. *Intl. Journal of Wavelets, Multiresolution and Information Processing*, 2(1):23–42, 2004.
- [20] E. Vanraes, J. Windmolders, A. Bultheel, and P. Dierckx. Automatic construction of control triangles for subdivided Powell–Sabin splines. *Comput. Aided Geom. Design*, 21(7):671–682, 2004.
- [21] J. Windmolders, E. Vanraes, P. Dierckx, and A. Bultheel. Uniform Powell–Sabin spline wavelets. *J. Comput. Appl. Math.*, 154(1):125–142, 2003.

Jan Maes
Department of Computer Science
Katholieke Universiteit Leuven
Celestijnenlaan 200A
B-3001 Heverlee
Belgium
e-mail: jan.maes@cs.kuleuven.be

Adhemar Bultheel
Department of Computer Science
Katholieke Universiteit Leuven
Celestijnenlaan 200A
B-3001 Heverlee
Belgium
e-mail: adhemar.bultheel@cs.kuleuven.be