The contribution of wavelets in multifractal analysis

Stéphane Jaffard, Patrice Abry, Stéphane G. Roux, Béatrice Vedel, Herwig
Wendt

The contribution of wavelets in multifractal analysis

S. Jaffard[*], P. Abry[†], S. Roux[*], B. Vedel[†*], H. Wendt[*]

Abstract: We show how wavelet techniques allow to derive irregularity properties of functions on two particular examples: Lacunary Fourier series and some Gaussian random processes. Then, we work out a general derivation of the multifractal formalism in the sequence setting, and derive some of its properties.

Contents

[*]Address: Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050 du CNRS, Université Paris Est, 61 Avenue du Général de Gaulle, 94010 Créteil Cedex, France.

[†]Address: CNRS UMR 5672 Laboratoire de Physique, ENS de Lyon, 46, allée d'Italie, F-69364 Lyon cedex, France

The purpose of multifractal analysis is twofold: on the mathematical side, it allows to determine the size of the sets of points where a function has a given Hölder regularity; on the signal processing side, it yields new collections of parameters associated to the signal considered and which can be used for classification, model selection, or for parameter selection inside a parametric setting. The main advances in the subject came from a better understanding of the interactions between these two motivations. The seminal ideas of this subject were introduced by N. Kolmogorov in the years 1940, in the study of turbulence. Though they could have been used in other contexts at that time, they remained confined to this specific subject up to the mid 80's. One reason is that the "scaling function" identified by Kolmogorov as a key-tool in the study of turbulence, was not clearly interpreted as a function-space index. Therefore, the subject could no benefit from the important advances performed from the 50s to the 80s in real analysis and function space theory. The situation changed completely in the mid 80s for several reasons:

- The interpretation of the scaling function as a description of the statistical repartition of the pointwise Hölder singularities of the signal by G. Parisi and U. Frisch supplied new motivations that were no more specific to turbulence analysis. As a consequence, these general methods were applied in many other settings.

- The wavelet formulation of the scaling function supplied ways to rewrite it that were fitted to modern signal processing; indeed, they were numerically more stable, they led to alternative, more robust definitions of the scaling function, and they allowed a mathematical analysis of these methods.

Our purpose in these lecture notes is to describe these developments, and to give an introduction to the recent research topics in this area. We will also mention several open questions. It is partly a review based on [1, 19, 20, 21, 23, 26], but it also contains original results. Recent applications of these techniques in signal and image processing can be found in [2, 3, 27, 41, 42].

1   Kolmogorov's scaling law and function spaces

Let us start by a short and partial description of the seminal work of Kolmogorov in fully developed turbulence. The streamwise component of turbulent flow velocity spatial field exhibits very irregular fluctuations over a large range of scales, whose statistical moments furthermore behave, within the so-called inertial scale range, like power laws with respect to the scale h; this velocity measured at a given point is therefore a function of time only, which we denote by $v(t)$. This power-law behavior is written

$$\int |v(t+h) - v(t)|^p dt \quad \sim \quad h^{\eta(p)}. \tag{1}$$

This statement means that the function $\eta(p)$ can be determined as a limit when $h \to 0$ on a log-log plot; it is called the scaling function of the velocity v. Characterization and understanding of the observed scaling properties play a central role in the theoretical description of turbulence, and Kolmogorov in 1941 expected a linear scaling function for turbulent flows: $\eta(p) = p/3$. This prediction has been refined by Obukhov and Kolmogorov in 1962 who predicted a (quadratic) non-linear behavior of the scaling exponents. The non-linear behavior of $\eta(p)$ was confirmed by various experimental results and other models have been proposed leading to different scaling functions $\eta(p)$. Let us now give the function space interpretation of this initial scaling function. This is done with the help of the spaces $\mathrm{Lip}(s, L^p)$ defined as follows. (We give definitions in the d-dimensional setting, since we will deal with several variable functions later on.)

Definition 1 *Let $s \in (0,1)$, and $p \in [1,\infty]$; $f \in Lip(s, L^p(\mathbb{R}^d))$ if $f \in L^p$ and $\exists C > 0$ such that $\forall h > 0$,*

$$\| f(x+h) - f(x) \|_{L^p} \leq C h^s. \tag{2}$$

Note that, if s is larger than 1, one uses differences of higher order in (2). It follows from this definition that, if $\eta(p) < p$,

$$\eta(p) = \sup\{s : f \in \mathrm{Lip}(s/p, L^p(\mathbb{R}^d))\}. \tag{3}$$

Remarks: The condition $\eta(p) < p$ has to hold because this interpretation is valid only if the smoothness exponent s in Definition 1 is less than 1; otherwise, the scaling function should be defined using higher order differences. The spaces $\mathrm{Lip}(s, L^p)$ are defined only for $p \geq 1$, however, in applications, one also considers values of the scaling function for $p < 1$; therefore, one would like to extend the function space interpretation of the scaling function to smaller values of p in a proper mathematical way. Finally, the scaling law in (1) is much more precise than what is given by the function space interpretation (3): Indeed, the order of magnitude of the integral $\int |v(t+h) - v(t)|^p dt$ might oscillate indefinitely between two power laws $h^{\eta_1(p)}$ and $h^{\eta_2(p)}$ when $h \to 0$. One way to avoid this problem is to define the scaling function directly by the function space interpretation, i.e. by (3). Thus, if

$$S(f, p, h) = \int_{\mathbb{R}^d} |f(x+h) - f(x)|^p dx, \quad \text{then} \quad \eta(p) = \liminf_{h \to 0} \frac{\log(S(f, p, h))}{\log(h)}. \tag{4}$$

Using (4) as a definition has the further advantage of allowing to drop the assumption $\eta(p) < p$. In practice, the scaling function is determined by plotting, for each p, $\log(S(f, p, h))$ as a function of $\log(h)$; it can be sharply estimated only if one obtains a plot which is close to a straight line on a sufficient number of scales. This means in particular that, if in (4) the liminf is not a real limit, then one cannot expect to determine the scaling function. Note also that (3) and (4) are not changed if one modified the right hand-side of (2) by logarithmic corrections, i.e. if it were of the form $|h|^s |\log(1/h)|^\beta$ for instance. This remark allows to give alternative interpretations of the scaling function based on other function spaces, and to extend it to values of p that lie between 0 and 1. Let us recall the definition of other families of alternative function spaces currently used.

Definition 2 *Let $s \geq 0$ and $p \geq 1$. A function f belongs to the Sobolev space $L^{p,s}(\mathbb{R}^d)$ if $f \in L^p$ and if $(Id - \Delta)^{s/2} f \in L^p$, where the operator $(Id - \Delta)^{s/2}$ is defined as follows: $g = (Id - \Delta)^{s/2} f$ means that $\hat{g}(\xi) = (1 + |\xi|^2)^{s/2} \hat{f}(\xi)$ (the function $(1 + |\xi|^2)^{s/2}$ being C^∞ with polynomial increase, $(1 + |\xi|^2)^{s/2} \hat{f}(\xi)$ is well defined if f is a tempered distribution).*

This definition amounts to say that f and its fractional derivatives of order at most s belong to L^p. Let us now recall the definition of Besov spaces. If s is large enough, Besov spaces can be defined by conditions on the finite differences $\Delta_h^M f$ which are defined as follows.

Let $f : \mathbb{R}^d \to \mathbb{R}$ and $h \in \mathbb{R}^d$. The first order difference of f is

$$(\Delta_h^1 f)(x) = f(x + h) - f(x).$$

If $n > 1$, the differences of order n are defined recursively by

$$(\Delta_h^n f)(x) = (\Delta_h^{n-1} f)(x + h) - (\Delta_h^{n-1} f)(x).$$

Definition 3 *Let p, q and s be such that $0 < p \leq +\infty$, $0 < q \leq +\infty$ and $s > d(\frac{1}{p} - 1)_+$; then $f \in B_p^{s,q}(\mathbb{R}^d)$ if $f \in L^p$ and if, for $M > s$,*

$$\int_{|h| \leq 1} |h|^{-sq} \|(\Delta_h^M f\|)_p^q \frac{dh}{|h|^d} \leq C. \tag{5}$$

If $s \geq 0$ and $p \geq 1$, then the following embeddings hold

$$B_p^{s,1} \hookrightarrow L^{p,s} \hookrightarrow B_p^{s,\infty}, \quad \text{and} \quad B_p^{s,1} \hookrightarrow \mathrm{Lip}(s, L^p(\mathbb{R}^d)) \hookrightarrow B_p^{s,\infty}$$

and, if $s \geq 0$, $p > 0$ and $0 < q_1 < q_2$, then $\forall \epsilon > 0$,

$$B_p^{s+\epsilon,\infty} \hookrightarrow B_p^{s,q_1} \hookrightarrow B_p^{s,q_2} \hookrightarrow B_p^{s,\infty}.$$

Thus $B_p^{s,q}$ is "very close" to $L^{p,s}$ and $\mathrm{Lip}(\alpha, L^p)$. (Recall also that $B_\infty^{\alpha;\infty} = C^\alpha(\mathbb{R}^d)$.) In fact, all the previous families of function spaces coincide "up to logarithmic corrections" in the sense mentioned previously. More precisely, they are equivalent in the following sense:

Definition 4 *Let A_p^s and B_p^s denote two families of function spaces. They are equivalent families of function spaces in the range (p_1, p_2) if*

$$\forall p \in (p_1, p_2), \quad \forall \epsilon > 0 \qquad A_p^{s+\epsilon} \subset B_p^s \subset A_p^{s-\epsilon}.$$

The scaling function of a function f in the scale A_p^s is

$$\eta^A(p) = \sup\{s : f \in A_p^{s/p}\}. \tag{6}$$

4

Thus, the scaling function associated with two equivalent families coincide. This abstract setting will also be useful in Section 5 where we will consider families of function spaces defined for sequences instead of functions. The embeddings between $\text{Lip}(s, L^p)$, Sobolev and Besov spaces imply that they are equivalent families of function spaces in the range $(1, \infty)$; and, when q varies, Besov spaces are equivalent families in the range $(0, \infty)$. Therefore, in (3), one can replace $\text{Lip}(\alpha, L^p)$ spaces by Sobolev or Besov spaces without altering the definition of the scaling function. The notion of equivalent families of function spaces has a practical motivation: Since the scaling function of a signal is numerically determined by a slope in a log-log plot, one cannot numerically draw a difference between two scales that are equivalent.

Since Besov spaces are defined for $p > 0$, the definition of the scaling function with Besov spaces allows to extend it in a natural way to all values of $p > 0$:

$$\forall p > 0, \ \ \forall q > 0, \qquad \eta(p) = \sup\{s : f \in B_q^{s/p, p}\}. \tag{7}$$

It follows that, for $p > 0$, any of the equivalent definitions of Besov or Sobolev spaces which have been found can be used in the determination of the scaling function of a signal. Wavelet characterizations are the ones that are now preferred in practice. We will see how the extend $\eta(p)$ to negative values of p (this will require a detour via wavelet-based formulas for scaling functions).

Let us mention another problem posed by function-space modeling when applied to real-life signals: Data are always available on a finite length; therefore modeling should use function spaces defined on an interval (or a domain, in several dimensions). This leads to several complications, especially when dealing with wavelets, and is not really relevant, when boundary phenomena are not of interest. Therefore, one uses wavelet bases and function spaces defined on $\mathbb{R}$ or $\mathbb{R}^d$.

2　Pointwise regularity

2.1　Hölder exponents

Pointwise regularity is a way to quantify, using a positive parameter α, the fact that the graph of a function may be more or less "ruguous" at a point x_0.

Definition 5 *Let α be a nonnegative real number, and $x_0 \in \mathbb{R}^d$; a function $f : \mathbb{R}^d \to \mathbb{R}$ belongs to $C^\alpha(x_0)$ if there exists $C > 0$, $\delta > 0$ and a polynomial P satisfying $\deg(P) < \alpha$ such that*

$$if \ \ |x - x_0| \leq \delta, \qquad |f(x) - P(x - x_0)| \leq C|x - x_0|^\alpha. \tag{8}$$

The Hölder exponent of f at x_0 is

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

The polynomial P is unique; the constant term of P is necessarily $f(x_0)$; P is called the Taylor expansion of f at x_0 of order α. The notion of Hölder exponent is adapted

to functions whose regularity changes abruptly from point to point. When it is not the case, the more stable notion of *local Hölder exponent* can be used:

$$H_f(x_0) = \inf\{\alpha : \ \exists \delta > 0, \ \ f \in C^{\alpha}([x_0 - \delta, x_0 + \delta])\},$$

see [28]. However, up to now, this notion has had no impact on multifractal analysis: Indeed, most functions with a nonconstant pointwise Hölder exponent which have been considered have a constant local Hölder exponent (in which case, the local Hölder exponent does not allow to draw distinctions between different local behaviors).

Note that (8) implies that f is bounded in a neighbourhood of x_0; therefore, the Hölder exponent can be defined only for locally bounded functions. The Hölder exponent is defined point by point and describes the local regularity variations of f. Some functions have a constant Hölder exponent; they are called *monohölder functions*. It is the case of the Weierstrass functions that we will consider in Section 3. We will see that it is also the case for Brownian motion, and Fractional Brownian Motions. Such functions display a "very regular irregularity". On the opposite, "multifractal functions" have a very irregular Hölder exponent which cannot be estimated numerically point by point.

2.2 Other notions of pointwise regularity

The notion of pointwise Hölder regularity is pertinent only if applied to locally bounded functions. We will derive pointwise irregularity results for solutions of PDEs where the natural function space setting is L^p or Sobolev spaces. In such cases, one has to use the following extension of pointwise smoothness, which was introduced by Calderón and Zygmund in 1961, see [7].

Definition 6 *Let $B(x_0, r)$ denote the open ball centered at x_0 and of radius r; let $p \in [1, +\infty)$ and $\alpha > -d/p$. Let f be a tempered distribution on $\mathbb{R}^d$; f belongs to $T_{\alpha}^p(x_0)$ if it coincides with an L^p function in a ball $B(x_0, R)$ for an $R > 0$, and if there exist $C > 0$ and a polynomial P of degree less than α such that*

$$\forall r \leq R, \ \ \left(\frac{1}{r^d} \int_{B(x_0, r)} |f(x) - P(x - x_0)|^p dx \right)^{1/p} \leq Cr^{\alpha}. \tag{9}$$

The p-exponent of f at x_0 is

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

Note that the Hölder exponent corresponds to the case $p = +\infty$, and the condition on the degree of P implies its uniqueness. This definition is particularly useful when dealing with functions which are not locally bounded: it is a natural substitute for pointwise Hölder regularity when functions in L_{loc}^p are considered. In particular, the p-exponent can take values down to $-d/p$, and therefore it allows to model behaviors which are locally of the form $1/|x - x_0|^{\alpha}$ for $\alpha < d/p$. For example, this is relevant in fully developed turbulence where singularities of negative Hölder exponent corresponding

to thin vorticity filaments can be observed, see [4]. We will use this notion in Section 3, in order to derive everywhere irregularity results for solutions of PDEs.

Pointwise Hölder regularity can also be considered in the setting of measures.

Definition 7 *Let $x_0 \in \mathbb{R}^d$ and let $\alpha \geq 0$. A probability 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,*

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

Let x_0 belong to the support of μ; then the lower Hölder exponent of μ at x_0 is

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

The upper Hölder exponent of μ at x_0 is

$$\tilde{h}_\mu(x_0) = \inf\{\alpha \quad \text{such that, for } r \text{ small enough}, \quad \mu(B(x_0, r)) \geq r^\alpha\}.$$

Note that the Hölder exponent of a measure is sometimes called the *local dimension*. We will need to deduce the Hölder exponent at every point from discrete quantities, which, in practice, will be indexed by the dyadic cubes.

Definition 8 *Let $j \in \mathbb{Z}$; a dyadic cube of scale j is of the form*

$$\lambda = \left[\frac{k_1}{2^j}, \frac{k_1 + 1}{2^j}\right) \times \ldots \times \left[\frac{k_d}{2^j}, \frac{k_d + 1}{2^j}\right), \tag{10}$$

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

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., if $\lambda_j(x_0)$ is given by (10), then it is the cube

$$3\lambda_j(x_0) = \left[\frac{k_1 - 1}{2^j}, \frac{k_1 + 2}{2^j}\right) \times \ldots \times \left[\frac{k_d - 1}{2^j}, \frac{k_d + 2}{2^j}\right).$$

Using dyadic cubes in analysis has two advantages: For a fixed j, they form a partition of $\mathbb{R}^d$, and they are naturally endowed with a tree structure which is inherited from the notion of inclusion: A dyadic cube of scale j is exactly composed of 2^j dyadic "children" of scale $j + 1$. This tree sructure will play a key-role in the notion of *wavelet leader* in Section 4.3, where pointwise Hölder regularity of functions will be characterized in terms of quantities defined on the dyadic cubes. It is also the case for the pointwise exponents of measures we introduced: One immediately checks that they can be derived from the knowledge of the quantities $\omega_\lambda = \mu[3\lambda]$:

- Let μ be a nonnegative measure defined on $\mathbb{R}^d$. Then

$$\forall x_0, \qquad h_\mu(x_0) = \liminf_{j \to +\infty} \left(\frac{\log(\mu[3\lambda_j(x_0)])}{\log(2^{-j})}\right). \tag{11}$$

- Similarly,

$$\forall x_0, \qquad \tilde{h}_\mu(x_0) = \limsup_{j \to +\infty} \left(\frac{\log\left(\mu[3\lambda_j(x_0)]\right)}{\log(2^{-j})} \right). \tag{12}$$

We will now study some Gaussian processes and lacunary Fourier series, which show why wavelet-type techniques yield irregularity results on technically easy examples.

2.3   Brownian motion and related noncentered stochastic processes

Brownian motion (defined on $\mathbb{R}^+$) is (up to a multiplicative constant) the only random process with independent and stationary increments which has continuous sample paths; i.e., if $y > x$, $B(y) - B(x)$ is independent of the $B(t)$ for $t \leq x$, and has the same law as $B(y - x)$. P. Lévy and Z. Ciesielski obtained a remarkable decomposition of the Brownian motion restricted on $[0, 1]$. Let $\Lambda(x)$ be the "hat function" defined by

$$\begin{aligned} \Lambda(x) = & \quad x & \text{if} \quad & x \in [0, 1/2] \\ = & \quad 1 - x & \text{if} \quad & x \in [1/2, 1] \\ = & \quad 0 & & \text{else.} \end{aligned}$$

The Brownian motion restricted on $[0, 1]$ can be written

$$B(x) = \chi_0 x + \sum_{j \geq 0} \sum_{k=0}^{2^j - 1} 2^{-j/2} \chi_{j,k} \Lambda(2^j x - k) \tag{13}$$

where χ_0 and the $\chi_{j,k}$ are Gaussian independant identically distributed (IID) random variables of variance 1. The set of functions 1, x, and the

$$\Lambda(2^j x - k), \quad j \geq 0, \ k = 0, \ldots, 2^j - 1, \tag{14}$$

which appear in (13), is called the *Schauder basis* on the interval $[0, 1]$. Note that (14) has the same algorithmic form as wavelet bases (see Section 4.1); indeed, the Schauder basis is obtained by taking the primitives of the Haar basis (and renormalizing them correctly). Therefore it is not surprising that the technique we introduce now in order to estimate the pointwise smoothness of the Brownian motion anticipates the wavelet techniques of Section 4.3.

Definition 9 *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a given continuous function. The decentered Brownian motion of expectation f is the stochastic process*

$$X(x) = f(x) + B(x).$$

This denomination is justified by the fact that $\forall x, \ \mathbb{E}(X(x)) = f(x)$.

Theorem 1 *Let f be an arbitrary continuous function on $[0, 1]$. With probability 1, the sample paths of X satisfy*

$$\forall x_0 \in [0, 1] \quad \limsup_{x \to x_0} \frac{|X(x) - X(x_0)|}{\sqrt{|x - x_0|}} > 0.$$

Therefore the Hölder exponent of X satisfies

$$a.s. \qquad \forall x \qquad h_X(x) = \inf\left(h_f(x), \frac{1}{2}\right). \tag{15}$$

Remark: This theorem can be interpreted as an example of a generic result in the sense of *prevalence.* Such results, where one proves that a property holds "almost surely" in a given function space E require the proof of results of this type, which hold for the sum of an arbitrary function $f \in E$ and of a stochastic process whose sample path almost surely belong to E, see [12, 16, 17]. Here, the corresponding prevalent result is that the Hölder exponent of almost every continuous function is everywhere less than $1/2$. The prevalent notion of genericity offers an alternative which is often preferred to the previous notion of "quasi-sure" in the sense of Baire categories.

In order to prove Theorem 1, we will need two ancillary lemmas. The first one yields the explicit formula of the coefficients of a continuous function on the Schauder basis.

Lemma 1 *Let f be a continuous function on $[0,1]$; then*

$$f(x) = f(0) + x(f(1) - f(0)) + \sum_{j \geq 0} \sum_{k=0}^{2^j - 1} C_{j,k} \Lambda(2^j x - k)$$

where

$$C_{j,k} = 2f\left(\frac{k + 1/2}{2^j}\right) - f\left(\frac{k}{2^j}\right) - f\left(\frac{k+1}{2^j}\right).$$

We give the idea of the proof of this well known result: One checks by recursion on J that

$$P_J(f)(x) = f(0) + x(f(1) - f(0)) + \sum_{j=0}^{J} \sum_{k=0}^{2^j - 1} C_{j,k} \Lambda(2^j x - k)$$

is the continuous piecewise linear function which coincides with f at the points $l/2^{J+1}$; the uniform convergence of $P_J(f)$ to f follows from the uniform continuity of f on $[0,1]$.

Lemma 2 *Let f be a continuous function on $[0,1]$. Let $k_j(x)$ be the integer k such that $x \in [\frac{k}{2^j}, \frac{k+1}{2^j}[$. If*

$$|f(x) - f(x_0)| \leq C\sqrt{|x - x_0|}, \tag{16}$$

then its Schauder coefficients on $[0,1]$ satisfy

$$|C_{j,k_j(x_0)}| \leq 4C2^{-j/2}.$$

Proof of Lemma 2: Using lemma 1,

$$|C_{j,k_j(x_0)}| = \left|2f\left(\frac{k_j(x_0) + 1/2}{2^j}\right) - f\left(\frac{k_j(x_0)}{2^j}\right) - f\left(\frac{k_j(x_0) + 1}{2^j}\right)\right|$$

$$\leq 2\left|f\left(\frac{k_j(x_0) + 1/2}{2^j}\right) - f(x_0)\right| + \left|f\left(\frac{k_j(x_0)}{2^j}\right) - f(x_0)\right| + \left|f\left(\frac{k_j(x_0) + 1}{2^j}\right) - f(x_0)\right|,$$

9

which, using (16) is bounded by $4C2^{-j/2}$.

Proof of Theorem 1: Denote by $C_{j,k}$ the coefficients of X on the Schauder basis. We call a (C, j_0)-*slow point* a point x_0 where the sample path of $X(x)$ satisfies (16) for any x such that $|x - x_0| \leq 2^{-j_0}$. If x_0 is such a point, using Lemma 2, for all couples $(j, k_j(x_0))$ such that $j \geq j_0$, we have

$$|C_{j,k_j(x_0)}| \leq 4C2^{-j/2}, \tag{17}$$

Let $f_{j,k}$ denote the Schauder coefficients of f. Using (13), (17) can be rewritten

$$\forall j \geq j_0, \qquad |2^{-j/2}\chi_{j,k_j(x_0)} + f_{j,k_j(x_0)}| \leq 4C2^{-j/2}.$$

For any j, k, let

$$p_{j,k} = \mathbb{P}\left(|2^{-j/2}\chi_{j,k} + f_{j,k}| \leq 4C2^{-j/2}\right).$$

Since the $\chi_{j,k}$ are standard Gaussians,

$$p_{j,k} = \sqrt{\frac{2}{\pi}} \int_{-4C-2^{j/2}f_{j,k}}^{4C-2^{j/2}f_{j,k}} e^{-x^2/2}dx \leq 8C\sqrt{\frac{2}{\pi}} := p_C.$$

Let j_0 be given, $j \geq j_0$ and λ be a dyadic interval of length 2^{-2j}. If $l \leq 2j$, let $k_l(\lambda)$ denote the integer k such that $\lambda \in [\frac{k}{2^l}, \frac{k+1}{2^l}[$. Let E_λ be the event defined by:

$$\forall l \in \{j+1, \cdots, 2j\}, \quad |C_{l,k_l(\lambda)}| \leq 4C2^{-l/2}.$$

Since the Schauder coefficients of B are independent, the probability of E_λ satisfies $\mathbb{P}(E_\lambda) \leq (p_C)^j$. Since there are 2^{2j} such intervals λ, the probability that at least one of the events E_λ occurs is bounded by $(p_C)^j 2^{2j}$. Since $p_C \leq 8C\sqrt{\frac{2}{\pi}}$, we see that, if $C < \frac{1}{32}\sqrt{\frac{2}{\pi}}$, this probability tends to 0, and therefore a.s. there is no (C, j_0)-slow point in the interval $[0, 1]$; the result is therefore true for any j_0, hence the first part of the theorem holds.

The second part of the theorem is a consequence of the following classical results:

- $h_{f+g}(x) \geq \inf(h_f(x), h_g(x))$;

- if $h_f(x) \neq h_g(x)$, then $h_{f+g}(x) = \inf(h_f(x), h_g(x))$;

- the Hölder exponent of the Brownian motion is everywhere $1/2$;

(15) follows for the points where the Hölder exponent of f differs from $1/2$. Else, $h_{f+B} \geq 1/2$, and the first part of the theorem implies that it is at most $1/2$.

An extension of Theorem 1 to the Fractional Brownian Motion will be proved in Section 4.4.

3 Lacunary Fourier series

Our second example of estimation of Hölder exponents is supplied by lacunary Fourier series. Informally, a lacunary Fourier series satisfies the following property: "Most" of its coefficients vanish. We will obtain a general result of irregularity based on the *Gabor-wavelet transform*, and we will apply it to the case of multidimensional nonharmonic Fourier series. Examples of applications to everywhere irregularity results for solutions of PDEs will then be derived.

3.1 A pointwise irregularity criterium

We will use the following notation: If λ, $x \in \mathbb{R}^d$, $\lambda \cdot x$ denotes the usual scalar product of λ and x.

Definition 10 *Let $\phi : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a function in the Schwartz class such that $\hat{\phi}(\xi)$ is supported in the unit ball centered at 0 and such that $\hat{\phi}(0) = 1$. The Gabor-wavelet transform of a function or a tempered distribution f defined on $\mathbb{R}^d$ is defined by*

$$d(a, b, \lambda) = \frac{1}{a^d} \int_{\mathbb{R}^d} f(x) e^{-i\lambda \cdot (x-b)} \phi \left(\frac{x - b}{a} \right) dx. \tag{18}$$

Note that, if we pick λ of the form $\lambda = \lambda_0/a$ where λ_0 is fixed and does not belong to $supp(\phi)$, then this definition boils down to the continuous wavelet transform in several dimensions; the purpose of adding the extra factor $e^{-i\lambda \cdot (x-b)}$ is to bring an additional frequency shift which will prove useful.

Proposition 1 *Let f be a tempered distribution; let $p \in (1, +\infty]$ and assume that f belongs to L^p. Let $\alpha > -d/p$; if $f \in T_\alpha^p(x_0)$, then there exists $C' > 0$, which depends only on ϕ and α such that*

$$\forall a > 0, \quad \forall \lambda : \quad |\lambda| \geq \frac{1}{a}, \quad |d(a, b, \lambda)| \leq CC' a^\alpha \left(1 + \frac{|x_0 - b|}{a} \right)^{\alpha + d/p}, \tag{19}$$

where C is the $T_\alpha^p(x_0)$ constant that appears in (9).

Proof of Proposition 1: If $\omega_{a,\lambda}(x) = a^{-d} e^{-i\lambda \cdot x} \phi(x/a)$, then

$$\widehat{\omega_{a,\lambda}}(\xi) = \hat{\phi}(a(\xi + \lambda)),$$

so that, as soon as $|\lambda| > 1/a$, $\widehat{\omega_{a,\lambda}}$ and all its derivatives vanishes at 0. It follows that

$$d(a, b, \lambda) = \frac{1}{a^d} \int_{\mathbb{R}^d} (f(x) - P(x - x_0)) e^{i\lambda \cdot x} \phi \left(\frac{x - b}{a} \right) dx.$$

For $n \geq 0$, let $B_n = B(b, 2^n a)$, $\Delta_n = B_{n+1} - B_n$ and $\Delta_0 = B_0$. We split $d(a, b, \lambda)$ as a sum of integrals I_n over Δ_n. Let q denote the conjugate exponent of p; by Hölder's inequality,

$$\forall n \geq 0, \quad |I_n| \leq \frac{1}{a^d} \| f(x) - P(x - x_0) \|_{L^p(B_{n+1})} \left\| \phi \left(\frac{x - b}{a} \right) \right\|_{L^q(\Delta_n)}.$$

Since $B_{n+1} \subset B(x_0, |x_0 - b| + 2^{n+1}a)$, and since ϕ has fast decay, $\forall D$ large enough,

$$|I_n| \leq \frac{CC'(D)}{a^d} (|x_0 - b| + 2^{n+1}a)^{\alpha + d/p} a^{d/q} (2^{-Dn})^{1/q}.$$

Therefore

$$
\begin{aligned}
|d(a, b, \lambda)| \quad &\leq CC'(D) a^{-d/p} \sum_{n=0}^{\infty} (|x_0 - b| + 2^{n+1}a)^{\alpha + d/p} 2^{-Dn/q} \\
&\leq CC' \left(|x_0 - b|^{\alpha + d/p} a^{-d/p} + a^{\alpha} \right),
\end{aligned}
$$

hence Proposition 1 holds.

3.2 Application to nonharmonic Fourier series

Let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence of points in $\mathbb{R}^d$. We will consider series of the form

$$f(x) = \sum_{n \in \mathbb{N}} a_n e^{i\lambda_n \cdot x}, \tag{20}$$

where $(a_n)_{n \in \mathbb{N}}$ is a sequence of complex numbers with, at most, polynomial increase. Of course, we can (and will) assume that the λ_n are distinct. Note that we do not assume that the λ_n are integers; one usually refers to (20) as *nonharmonic Fourier series.*

Definition 11 *Let (λ_n) be a sequence in $\mathbb{R}^d$. The gap sequence associated with (λ_n) is the sequence (θ_n) defined by*

$$\theta_n = \inf_{m \neq n} |\lambda_n - \lambda_m|.$$

The sequence (λ_n) is separated if $\inf_n \theta_n > 0$.

Note that θ_n is the distance between λ_n and its closest neighbour. We will always assume in this section that (λ_n) is separated and (a_n) increases at most polynomially, which implies the convergence of (20) in the space of tempered distributions.

Proposition 2 *Let f be given by (20) and let x_0 be a given point of $\mathbb{R}^d$, $p > 1$, $\alpha > -d/p$ and assume that f belongs to L^p in a neighbourhood of x_0. If $f \in T_{\alpha}^p(x_0)$, then there exists C' which depends only on α such that*

$$\forall n \in \mathbb{N} \quad \text{if } |\lambda_n| \geq \theta_n, \quad \text{then} \quad |a_n| \leq \frac{CC'}{\theta_n^{\alpha}}, \tag{21}$$

where C is the constant that appears in (9).

Remark: This is indeed an everywhere irregularity result: Let

$$H = \sup\{\alpha : \quad (21) \text{ holds}\};$$

Proposition 2 implies that the p-exponent of f is everywhere smaller than H. (Since the p-exponent is larger than the Hölder exponent, if f is locally bounded it also implies that the Hölder exponent of f is everywhere smaller than H.)

Proof of Proposition 2: Let us estimate the Gabor-wavelet transform of f at particular points, and for a function ϕ such that $\hat{\phi}(\xi)$ is radial, supported in the unit ball centered at 0, and such that $\hat{\phi}(0) = 1$. Let

$$D_m = d\left(\frac{1}{\theta_m}, x_0, \lambda_m, \right). \tag{22}$$

On one hand,

$$D_m = (\theta_m)^d \int \left(\sum_n a_n e^{i(\lambda_n - \lambda_m) \cdot x} \phi(\theta_m(x - x_0)) \right) dx$$

$$= \sum_n a_n \hat{\phi}\left(\frac{\lambda_m - \lambda_n}{\theta_n} \right) e^{i(\lambda_n - \lambda_m) \cdot x_0}; \tag{23}$$

since $\hat{\phi}$ vanishes outside of $B(0,1)$, the definition of θ_n implies that $\hat{\phi}\left(\dfrac{\lambda_m - \lambda_n}{\theta_n} \right) = \delta_{n,m}$, so that $D_m = a_m$. on the other hand, if $f \in T_\alpha^p(x_0)$, then Proposition 1 implies that, for any m such that $|\lambda_m| \geq \theta_m$, $|D_m| \leq C\theta_m^{-\alpha}$; Proposition 2 follows.

3.3   Everywhere irregularity of solutions of Schrödinger's equation

As a consequence of the previous results, let us show that the solutions of a simple linear PDE display a remarkable property of everywhere irregularity if the initial condition is not smooth. We consider the one-dimensional Schrödinger equation

$$i\frac{\partial \psi}{\partial t} = -\frac{\partial^2 \psi}{\partial x^2}, \qquad \text{for} \quad (x,t) \in \mathbb{R} \times \mathbb{R} \tag{24}$$

$$\text{with initial condition:} \quad \psi(x,0) = \psi_0(x) = \sum_{n \in \mathbb{Z}} a_n e^{inx}. \tag{25}$$

The general solution of (24) can be written

$$\psi(x,t) = \sum_{n \in \mathbb{Z}} a_n e^{inx} e^{-in^2 t}. \tag{26}$$

Note that (26) is of the form $\displaystyle\sum_{n \in \mathbb{Z}} a_n e^{i\lambda_n \cdot X}$ where $\lambda_n = (n, -n^2)$, and $X = (x,t)$, so that the gap sequence θ_n associated with λ_n satisfies

$$\forall n \in \mathbb{Z}, \qquad \theta_n \geq |n| + 1. \tag{27}$$

13

Let $p > 1$ and $\alpha \geq -2/p$. It follows from Proposition 2 that

$$\text{if} \quad \psi \in T_\alpha^p(x_0, t_0), \quad \text{then} \quad \forall n, \quad |a_n| \leq \frac{C}{(|n|+1)^\alpha}. \tag{28}$$

But, if $|a_n| \leq C(|n|+1)^{-\alpha}$, then $\psi_0(x)$ belongs to the periodic Sobolev space H^s, as soon as $s < \alpha - 1/2$. One can also consider the trace $\tilde{\psi}_{x_0}$ of ψ at a given point x_0, as a function of t, i.e. formally, $\tilde{\psi}_{x_0}(t) = \psi(x_0, t)$. The solution is still a one-dimensional lacunary Fourier series with $\lambda_n = n^2$. We obtain that, if $\tilde{\psi}_{x_0}(t) \in T_\alpha^1(t_0)$, then $|a_n| \leq C/n^\alpha$. Hence the following corollary holds.

Corollary 1 *Let $s > -5/2$, let $\psi(x,t)$ be a solution of (24), and assume that $\psi_0 \notin H^s$. Then*

$$\forall \alpha > s + 1/2, \quad \forall (x_0, t_0), \quad \forall p > 1, \quad \psi \notin T_\alpha^p(x_0, t_0).$$

Furhermore, as regards irregularity in the time direction,

$$\forall \alpha > s + 1/2, \quad \forall (x_0, t_0), \quad \forall p > 1, \quad \psi_{x_0} \notin T_\alpha^p(x_0, t_0).$$

In particular, if (24) has a bounded solution, $\forall \alpha > s + 1/2, \forall (x_0, t_0), \psi_{x_0} \notin C^\alpha(t_0)$.

The fact that solutions of the Schrödinger equation are everywhere irregular somehow means that their graph is a fractal; in a similar spirit, the fractal properties of the graph of the fundamental solution of (24) have been investigated by K. Oskolkov, see [36] and references therein.

4 Wavelets, function spaces and Hölder regularity

Orthonormal wavelet bases are a privileged tool to study multifractal functions for several reasons. A first one, exposed in this section, is that classical function spaces (such as Besov or Sobolev spaces) can be characterized by conditions bearing on the wavelet coefficients, see Section 4.2; furthermore, pointwise regularity can also be characterized by simple local decay conditions on the wavelet coefficients, see Section 4.3. Another reason concerns the wavelet formulation of the multifractal formalism, which leads to the construction of new function spaces, directly defined by conditions on the wavelet coefficients. We will just recall some properties of orthonormal and biorthogonal wavelet bases that will be useful in the following. We refer the reader for instance to [10, 11, 29, 34] for detailed expositions of this subject.

4.1 Orthonormal and biorthogonal wavelet bases

Orthonormal wavelet bases are of the following form: There exists a function $\varphi(x)$ and $2^d - 1$ functions $\psi^{(i)}$ with the following properties: The functions $\varphi(x - k)$ ($k \in \mathbb{Z}^d$) and the $2^{dj/2}\psi^{(i)}(2^j x - k)$ ($k \in \mathbb{Z}^d$, $j \in \mathbb{Z}$) form an orthonormal basis of $L^2(\mathbb{R}^d)$. This basis is r-smooth if φ and the $\psi^{(i)}$ are C^r and if the $\partial^\alpha \varphi$, and the $\partial^\alpha \varphi \psi^{(i)}$, for $|\alpha| \leq r$, have fast decay.

Therefore, $\forall f \in L^2$,

$$f(x) = \sum_{k \in \mathbb{Z}^d} C_k \varphi(x - k) + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}^d} \sum_i c_{j,k}^i \psi^{(i)}(2^j x - k); \qquad (29)$$

the $c_{j,k}^i$ are the wavelet coefficients of f

$$c_{j,k}^i = 2^{dj} \int_{\mathbb{R}^d} f(x)\psi^{(i)}(2^j x - k)dx, \qquad (30)$$

and

$$C_k = \int_{\mathbb{R}^d} f(x)\varphi(x - k)dx. \qquad (31)$$

Remarks: In (29), we do not choose the L^2 normalisation for the wavelets, but rather an L^∞ normalisation which is better fitted to the study of Hölder regularity. The L^1 normalisation of (30) follows accordingly.

Note that (30) and (31) make sense even if f does not belong to L^2; indeed, if one uses smooth enough wavelets, these formulas can be interpreted as a duality product betweeen smooth functions (the wavelets) and distributions. We will see the examples of Sobolev and Besov spaces.

We will also need decompositions on *biorthogonal wavelet bases*, which are a useful extension of orthonormal wavelet bases. 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

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

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

$$f = \sum_{n=1}^{\infty} \langle f | f_n \rangle e_n. \qquad (32)$$

Biorthogonal wavelet bases are couples of Riesz bases of L^2 which are, of the form: on one side,

$$\varphi(x - k), \ (k \in \mathbb{Z}^d) \quad \text{and} \quad 2^{dj/2}\psi^{(i)}(2^j x - k), \ (k \in \mathbb{Z}^d, \ j \in \mathbb{Z})$$

and, on the other side,

$$\tilde{\varphi}(x - k) \ (k \in \mathbb{Z}^d) \quad \text{and} \quad 2^{dj/2}\tilde{\psi}^{(i)}(2^j x - k), \ (k \in \mathbb{Z}^d, \ j \in \mathbb{Z}).$$

Therefore, $\forall f \in L^2$,

$$f(x) = \sum_{k \in \mathbb{Z}^d} C_k \varphi(x - k) + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}^d} \sum_i c_{j,k}^i \psi^{(i)}(2^j x - k); \qquad (33)$$

15

where
$$c^i_{j,k} = 2^{dj} \int_{\mathbb{R}^d} f(x)\tilde{\psi}^{(i)}(2^j x - k)dx, \quad \text{and } C_k = \int_{\mathbb{R}^d} f(x)\tilde{\varphi}(x - k)dx. \qquad (34)$$

We will see that biorthogonal wavelet bases are particularly well adapted to the decomposition of the Fractional Brownian Motion; indeed, well chosen biorthogonal wavelet bases allow to decorrelate the wavelet coefficients of these processes (the wavelet coefficients become independent random variables), and therefore greatly simplifies their analysis.

We will use more compact notations for indexing wavelets. Instead of using the three indices (i, j, k), we will use dyadic cubes. Since i takes $2^d - 1$ values, we can assume that it takes values in $\{0, 1\}^d - (0, \ldots, 0)$; we introduce:

- $\lambda \ (= \lambda(i, j, k)) \ = \dfrac{k}{2^j} + \dfrac{i}{2^{j+1}} + \left[0, \dfrac{1}{2^{j+1}}\right)^d.$

- $c_\lambda = c^i_{j,k}$

- $\psi_\lambda(x) = \psi^{(i)}(2^j x - k).$

The wavelet ψ_λ is essentially localized near the cube λ; more precisely, when the wavelets are compactly supported

$$\exists C > 0 \quad \text{such that} \quad \forall i, j, k, \quad supp\,(\psi_\lambda) \subset C\,\lambda$$

(where $C\,\lambda$ denotes the cube of same center as λ and C times wider). Finally, Λ_j will denote the set of dyadic cubes λ which index a wavelet of scale j, i.e. wavelets of the form $\psi_\lambda(x) = \psi^{(i)}(2^j x - k)$ (note that Λ_j is a subset of the dyadic cubes of side 2^{j+1}). We take for norm on $\mathbb{R}^d$

$$\text{if } x = (x_1, \ldots, x_d), \quad |x| = \sup_{i=1,\ldots,d} |x_i|;$$

so that the diameter of a dyadic cube of side 2^{-j} is exactly 2^{-j}.

Among the families of wavelet bases that exist, two will be particularly useful for us:

- Lemarié-Meyer wavelets, such that φ and $\psi^{(i)}$ both belong to the Schwartz class;

- Daubechies wavelets, such that the functions φ and $\psi^{(i)}$ can be chosen arbitrarily smooth and with compact support.

If the wavelets are r-smooth, they have a corresponding number of vanishing moments, see [34]:
$$\text{If } |\alpha| < r, \quad \text{then} \quad \int_{\mathbb{R}^d} \psi^{(i)}(x)x^\alpha dx = 0.$$

Therefore, if the wavelets are in the Schwartz class, all their moments vanish.

16

4.2 Wavelets and function spaces

A remarkable property of wavelet bases is that they supply bases not only in the L^2 setting, but also for most function spaces that are used in analysis. When considering wavelet characterization of function spaces, a first natural question is to understand in which sense the wavelet series of a function, or of a distribution converges, i.e. in which sense wavelets are *bases* of the corresponding space; before giving the two standard definitions of bases, (depending whether E is separable or not, see [38]), we need to recall the notions of quasi-norm and quasi-Banach space.

Definition 12 *Let E be a vector space. A quasi-norm on E is a nonnegative function satisfying*

$$\exists C, \ \forall x, y \in E, \quad \| x + y \| \leq C(\| x \| + \| y \|),$$
$$\forall \lambda \in \mathbb{R}, \ \forall x \in E, \quad \| \lambda x \| = \| x \|,$$
$$\forall x \in E, \quad \| x \| = 0 \implies x = 0.$$

A quasi-Banach space is a vector space endowed with a quasi-norm, and which is complete for the corresponding topology.

Besov spaces for $p < 1$ or $q < 1$ are typical example of quasi-Banach spaces, indeed (5), or equivalently (37), only define quasi-norms if p or q are less than 1.

Definition 13 *Let E be a Banach, or a quasi-Banach space. A sequence e_n is a basis of E if the following condition holds: For any element f in E, there exists a unique sequence c_n such that the partial sums $\sum_{n \leq N} c_n e_n$ converge to f in E. It is an unconditional basis if furthermore*

$$\exists C > 0, \ \forall \epsilon_n \ \text{ such that } |\epsilon_n| \leq 1, \ \forall c_n, \ \| \sum c_n \epsilon_n e_n \|_E \leq C \| \sum c_n e_n \|_E . \tag{35}$$

If the space E is not separable (it is the case for instance of Besov spaces when p or q is infinite, or C^α spaces, since $C^\alpha = B_{\alpha,\infty}^\infty$) then, of course, it cannot have a basis in the previous sense. In this case, the following weaker notion often applies.

Definition 14 *Assume that E is the dual of a separable space F; a sequence $e_n \in E$ is a weak* basis of E if, $\forall f \in E$, there exists a unique sequence c_n such that the partial sums $\sum_{n \leq N} c_n e_n$ converge to f in the weak* topology. It is unconditional if furthermore (35) holds.*

Let F be either the dual of E (in the basis setting of Definition 13) or a predual of E (in the weak* basis setting); we will furthermore always assume in the following that,

$$\text{if } f = \sum c_n e_n, \quad \text{then there exist } g_n \in F \quad \text{such that } c_n = \langle f | g_n \rangle. \tag{36}$$

The g_n are called the biorthogonal system of the e_n; indeed, this notion extends the previous definition of biorthogonality in the non-Hilbert setting (in the Hilbert case where $E = F$, (36) boils down to (32)). Note that if E is a Banach space, if $F = E^*$ and if the e_n form a basis according to Definition 13, then (36) is automatically verified,

see [38]; it is also verified if the e_n are a wavelet basis, in which case $g_n = e_n$ for L^2 orthonomal wavelet bases, (or g_n is another wavelet basis in the wavelet biorthogonal case). Note that, for wavelets, the L^2 biorthogonal system is also the biorthogonal system for the (E, F) duality; indeed, by uniqueness if $\mathcal{S}_0$ is dense in either E or F, then the $(\mathcal{S}_0, \mathcal{S}_0')$ duality, the (L^2, L^2) duality and the (E, F) duality coincide for finite linear combinations of wavelets; therefore (36) holds for all functions of E by density, and the duality product $\langle f | g_n \rangle$ in (36) can be understood in any of the three settings. These considerations explain why the usual L^2 wavelet decomposition (29) also makes sense in other function space settings.

Examples of non-separable spaces for which wavelets are weak* bases include the Hölder spaces $C^s(\mathbb{R}^d)$, and, more generally, the Besov spaces $B_p^{s,q}$ with $p = +\infty$ or $q = +\infty$. We now give the wavelet characterizations of the spaces that will be useful for us. These characterizations supply equivalent norms or quasi-norms for the corresponding spaces, see [34].

Proposition 3 *Let ψ_λ be an r-smooth wavelet basis with $r > \sup(s, s + d(\frac{1}{p} - 1))$. Let $s > 0$ and $p, q \in (0, \infty]$. A function f belongs to the Besov space $B_p^{s,q}(\mathbb{R}^d)$ if and only if $(c_k) \in l^p$ and*

$$\sum_{j \in \mathbb{Z}} \left(\sum_{\lambda \in \Lambda_j} \left[2^{(s-d/p)j} |c_\lambda| \right]^p \right)^{q/p} \leq C \tag{37}$$

(using the usual convention for l^∞ when p or q in infinite).

A function f belongs to $L^{p,s}(\mathbb{R}^d)$ (for $1 < p < +\infty$) if and only if $(c_k) \in l^p$ and

$$\left(\sum_{\lambda \in \Lambda} |2^{sj} c_\lambda|^2 1_\lambda(x) \right)^{1/2} \in L^p(\mathbb{R}^d). \tag{38}$$

Remarks: When p or q is infinite, it may come as a surprise that nonseparable spaces are characterized by expansion properties on a countable set of functions. However, Proposition 3 does **not** state that the partial sums of the wavelet series (29) converge in the corresponding space (which would indeed be in contradiction with nonseparability); it only yields a quantity which is equivalent to the Besov or Sobolev norm. Note however that, in the separable case, partial sums of (29) do converge in the corresponding space. Wavelets are a basis of Sobolev or Besov spaces when $p < \infty$ and $q < \infty$; else they are a weak* basis.

4.3 Wavelet characterizations of pointwise regularity

Pointwise Hölder regularity is characterized in terms of the following quantities.

Definition 15 *The wavelet leaders are*

$$d_\lambda = \sup_{\lambda' \subset 3\lambda} |c_{\lambda'}|. \tag{39}$$

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

(Figure 1 gives an illustration of the construction of the wavelet leaders in dimension 1).
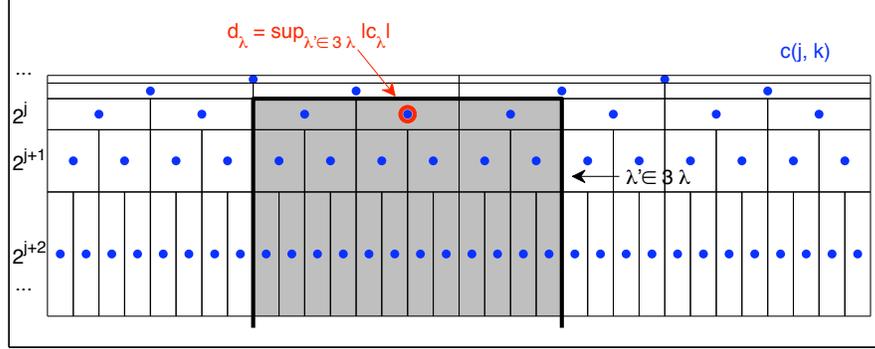

Figure 1: **Definition of wavelet Leaders.** The wavelet Leader d_λ (red circle) is defined as the largest wavelet coefficient $c_{\lambda'}$ (blue dots) within the time neighborhood 3λ (grey area) over all finer scales.

If $f \in L^\infty$, then

$$|c_\lambda| \leq 2^{dj} \int_{\mathbb{R}^d} |f(x)||\psi_\lambda(x)|dx \leq C \sup_{x \in \mathbb{R}^d} |f(x)|,$$

so that the wavelet leaders are finite. The wavelet characterization of the Hölder exponent requires the following regularity hypothesis, which is slightly stronger than continuity.

Definition 16 *A function f is uniform Hölder if there exists $\epsilon > 0$ such that $f \in C^\epsilon(\mathbb{R}^d)$.*

The following theorem allows to characterize the pointwise regularity by a decay condition of the $d_j(x_0)$ when $j \to +\infty$, see [18] for the first statement of this result, and [20] for its reformulation in terms of wavelet leaders; see also [25] for similar results in the setting of general moduli of continuity.

Theorem 2 *Let $\alpha > 0$ and let ψ_λ be an orthonormal basis (or let $(\psi_\lambda, \tilde{\psi}_\lambda)$ be a couple of biorthogonal wavelet bases) with regularity $r > \alpha$. If f is $C^\alpha(x_0)$, then there exists $C > 0$ such that*

$$\forall j \geq 0, \qquad d_j(x_0) \leq C2^{-\alpha j}. \tag{40}$$

Conversely, if (40) holds and if f is uniform Hölder, then there exist $C > 0$, $\delta > 0$ and a polynomial P satisfying $\deg(P) < \alpha$ such that

$$if \ |x - x_0| \leq \delta, \qquad |f(x) - P(x - x_0)| \leq C|x - x_0|^\alpha \log\left(1/|x - x_0|\right). \tag{41}$$

Remarks: Some uniform regularity is a necessary assumption in the converse part of Theorem 2: One can show that there exist bounded functions satisfying (40) for arbitrary large values of α, and whose Hölder exponent at x_0 vanishes. The reason of this phenomenon will be clarified in Section 4.3. Let f be a uniform Hölder function; since

(41) only involves a logarithmic correction of the modulus of continuity, the regularity of f at x_0 is therefore determined by the decay rate of the $d_j(x_0)$ on a log-log plot. Hence a formula similar to (11) holds:

If f is a uniform Hölder function, then

$$h_f(x_0) = \liminf_{j \to +\infty} \left(\frac{\log(d_j(x_0))}{\log(2^{-j})} \right). \tag{42}$$

The Hölder exponent supplies a definition of pointwise regularity which can be difficult to handle or irrelevant; here are a few reasons:

- Mathematical results concerning multifractal analysis based on the Hölder exponent necessarily make the assumption that f is continuous, see Theorem 2; in many situations, one wishes to analyze discontinuous functions; an important case is natural images which, because of the *occlusion* phenomenon (one object is partially hidden by another), present discontinuities. There exists other fields where one even has to consider non-locally bounded functions (for instance in the study of fully developed turbulence, see [4]).

- A standard function space setting used for the mathematical study of images is supplied by the space of functions of bounded variations $BV(\mathbb{R}^2)$. Recall that $f \in BV$ if ∇f (defined in the sense of distributions) is a bounded measure. In dimension 1, BV functions are bounded (but may be discontinuous); in dimension 2, they can be nowhere locally bounded; however $BV(\mathbb{R}^2) \subset L^2$ and therefore the natural setting to perform the multifractal analysis of BV functions in dimension 2 is to use $T_\alpha^2(x_0)$ regularity instead of $C^\alpha(x_0)$ regularity, see [15].

- Let $\Omega \subset \mathbb{R}^d$ be a domain of $\mathbb{R}^d$ with a fractal boundary. A possible way to perform a multifractal analysis of Ω consists in associating to its characteristic function 1_Ω a pointwise regularity exponent. The Hölder exponent is clearly not the right tool since, in this case, it can take only two values: 0 on the boundary $\partial\Omega$ and $+\infty$ elsewhere.

Another reason, based on "stability" requirements, will be detailed in the following; they are related to the fact that the condition $f \in C^\alpha(x_0)$ is not invariant under simple pseudodifferential operators of order 0, and equivalently, cannot be characterized by conditions on the moduli of the wavelet coefficients of f.

These considerations motivated the use of $T_\alpha^p(x_0)$ regularity, introduced in Definition 6 as an alternative criterium of pointwise regularity (already considered in Section 3). One immediately checks that the $T_\alpha^p(x_0)$ regularity condition is weaker than Hölder regularity: If $f \in C^\alpha(x_0)$, then, $\forall p$, $f \in T_\alpha^p(x_0)$. The drawbacks of the $C^\alpha(x_0)$ criterium of smoothness that we listed above disappear when one considers this notion of regularity; for instance, the Hölder exponent of a characteristic function 1_Ω only takes the value 0 along the boundary of Ω; on the opposite, consider the "cusp domain" $\Omega \subset \mathbb{R}^2$ defined by the conditions

$$(x,y) \in \Omega \quad \text{if} \quad 0 \le |x| \le |y|^\beta, \qquad \text{for a } \beta \ge 1;$$

at the origin, the p-exponent of 1_Ω is $(\beta - 1)/p$ which can take any nonnegative value. Therefore, the p-exponent of a characteristic function 1_Ω can freely vary along the boundary of Ω, thus opening the way to a multifractal analysis of domains, see [24]. Furthermore, mathematical results concerning a multifractal analysis based on the p-exponent do not require any uniform regularity assumption, see [22].

Let us come back to the initial problem we mentioned, i.e. the instability of the $C^\alpha(x_0)$ condition. Indeed, the initial motivation of Calderón et Zygmund was to understand how pointwise regularity conditions are transformed in the resolution of elliptic PDEs, and they introduced the $T_u^p(x)$ spaces because the standard pseudodifferential operators of order 0 are not continuous on $C^\alpha(x_0)$, whereas, it is the case for the $T_u^p(x)$ spaces. We point out how this deficiency of the $C^\alpha(x_0)$ condition can be put into light. We consider the Hilbert transform, which is the simplest possible singular integral operator in dimension 1, and also plays a key-role in signal processing; it is the convolution with the principal value of $1/x$, i.e. is defined by

$$\mathcal{H}f(x) = \lim_{\epsilon \to 0} \ \frac{1}{\pi} \int_{I_\epsilon(x)} \frac{f(y)}{x - y} dy,$$

where $I_\epsilon(x) = (-\infty, x - \epsilon] \cup [x + \epsilon, +\infty)$. An immediate computation shows that

$$\mathcal{H}(1_{[a,b]})(x) = \log \left| \frac{x - b}{x - a} \right|. \tag{43}$$

Let now $(x_n)_{n \in \mathbb{N}}$ be a strictly decreasing sequence such that $\lim_{n \to \infty} x_n = 0$. We can pick a positive, strictly decreasing sequence a_n such that

$$f = \sum_{n=1}^{\infty} a_n 1_{[x_{n+1}, x_n]} \tag{44}$$

is arbitrarily smooth at 0. Nonetheless, (43) implies that

$$\mathcal{H}f(x) = \sum_{n=1}^{\infty} a_n \log \left| \frac{x - x_{n+1}}{x - x_n} \right| = -a_1 \log |x - x_1| + \sum_{n=1}^{\infty} (a_n - a_{n+1}) \log |x - x_{n+1}|,$$

which is not locally bounded near the origin, and therefore cannot have any Hölder regularity there. Note that what we really used here is the fact that the Hilbert transform is not continuous on L^∞.

We can actually reinterpret the previous counterexample in a way that sheds some light on the necessity of the uniform Hölder assumption in Theorem 2. We will need the following lemma.

Lemma 3 *Let ψ be a wavelet generating an r-smooth orthonormal wavelet basis. Then the $\tilde{\psi}_{j,k} = \mathcal{H}(\psi_{j,k})$ also form an r-smooth orthonormal wavelet basis.*

We just sketch the proof: First, we recall that, if $f \in L^2$, then

$$\widehat{\mathcal{H}(f)}(\xi) = sgn(\xi)\hat{f}(\xi); \tag{45}$$

21

It follows that $\tilde{\psi} = \mathcal{H}(\psi)$ has the same number of vanishing moments as ψ. The smoothness of $\tilde{\psi}$ follows from the continuity of the Hilbert transform on the Hölder spaces $\dot{C}^\alpha$. The vanishing moments of ψ up to order r imply that $\tilde{\psi}$ and its derivatives up to order r decay like $(1 + |x|)^{-r-1}$. Note that (45) implies that the Hilbert transform is an L^2 isometry; therefore the functions $\mathcal{H}(\psi_{j,k})$ form an orthonormal basis of $L^2(\mathbb{R})$. Since the Hilbert transform is given by a convolution kernel, it commutes with translations, so that $\mathcal{H}(\psi_{j,k})$ is deduced from $\mathcal{H}(\psi_{j,0})$ by a translation of $k2^{-j}$. Because of the homogeneity of degree 0 of the function $sgn(\xi)$, it follows that the $\mathcal{H}(\psi_{j,0})$ are deduced from $\mathcal{H}(\psi)$ by a dyadic dilation. Therefore $(\mathcal{H}(\psi))_{j,k} = \mathcal{H}(\psi_{j,k})$; hence Lemma 3 holds.

Suppose now that a wavelet characterization of $C^\alpha(x_0)$ did exist; consider the function f defined by (44); its coefficients on the $\tilde{\psi}_{j,k}$ would satisfy this characterization. But

$$\langle f | \tilde{\psi}_{j,k} \rangle = \langle f | \mathcal{H}(\psi_{j,k}) \rangle = \langle \mathcal{H}(f) | \psi_{j,k} \rangle,$$

so that the criterium would be satisfied by the coefficients of $\mathcal{H}(f)$ on the $\psi_{j,k}$, which is absurd, since $\mathcal{H}(f)$ is not locally bounded in a neighbourhood of x_0.

We now turn to the wavelet characterization of $T_\alpha^p(x_0)$. It is derived from the wavelet characterization of L^p, which is a particular case of (38), where $s = 0$: We obtain that $f \in L^p$ if

$$\left(\sum_{\lambda \in \Lambda} |2^{sj} c_\lambda|^2 1_\lambda(x) \right)^{1/2} \in L^p(\mathbb{R}^d). \tag{46}$$

It is natural to expect that $T_\alpha^p(x_0)$ regularity will be characterzed by a local condition bearing on the quantities involved in (46).

Definition 17 *Let ψ_λ be a given wavelet basis on $\mathbb{R}^d$. The local square function is*

$$S_{f,\lambda}(x) = \left(\sum_{\lambda' \subset 3\lambda} |c_{\lambda'}|^2 1_{\lambda'}(x) \right)^{1/2}.$$

The following theorem of [22] gives the wavelet characterization of $T_\alpha^p(x_0)$ (see also [24]).

Theorem 3 *Let $p \in (1,\infty)$ and $f \in L^p$. Let $\alpha > -d/p$ and assume that the wavelet basis used is r-smooth with $r > \sup(2\alpha, 2\alpha + 2d(\frac{1}{p} - 1))$; let*

$$d_\lambda^p = 2^{dj/p} \parallel S_{f,\lambda} \parallel_p.$$

If $f \in T_\alpha^p(x_0)$, then $\exists C \geq 0$ such that $\forall j \geq 0$,

$$d_{\lambda_j(x_0)}^p \leq C 2^{-\alpha j}. \tag{47}$$

Conversely, if (47) holds and if $\alpha \notin \mathbb{N}$, then $f \in T_\alpha^p(x_0)$.

22

Remarks: In contradistinction with Theorem 2, this result does not require a uniform regularity assumption. If $p = 2$, this characterization boils down to a local l^2 condition on the wavelet coefficients, since

$$d_\lambda^2 = \left(2^{dj} \sum_{\lambda' \subset 3\lambda_j(x_0)} 2^{-dj'} |c_{\lambda'}|^2 \right)^{1/2}. \tag{48}$$

Theorem 3 can be given the following interpretation which is similar to (11) for Hölder exponents of measures and to (42) for Hölder exponents of functions:

Lemma 4 *Let $p \in (1, \infty)$ and $f \in L^p$. Then*

$$h_f^p(x_0) = \liminf_{j \to +\infty} \left(\frac{\log \left(d_{\lambda_j(x_0)}^p \right)}{\log(2^{-j})} \right). \tag{49}$$

4.4   Application to decentered Fractional Brownian Motions

Fractional Brownian Motion (denoted by FBM) of index γ ($0 < \gamma < 1$) is the only centered Gaussian random process $B^\gamma(x)$ satisfying

$$E(|B^\gamma(x) - B^\gamma(y)|^2) = |x - y|^{2\gamma}.$$

We will use the following important feature: FBM of index γ can be deduced from the Gaussian white noise by a fractional integration of order $\gamma + 1/2$. With probability 1, a sample path of FBM of order γ has everywhere the Hölder exponent γ.

Since the Gaussian white noise $N(x)$ has standard Gaussian I.I.D. coefficients on any orthonormal basis, using a wavelet basis, we obtain

$$N(x) = \sum_{j,k} \chi_{j,k} 2^{j/2} \psi(2^j x - k).$$

Let ψ be in the Schwartz class and

$$\hat{\psi}_\alpha(\xi) = \frac{1}{|\xi|^\alpha} \hat{\psi}(\xi) \tag{50}$$

(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 one easily checks that the $2^{j/2}\psi_\alpha(2^j x - k)$ and the $2^{j/2}\psi_{-\alpha}(2^j x - k)$ form biorthogonal bases. 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,

$$B_\gamma(x) = \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}} 2^{-\gamma j} \chi_{j,k} \, \psi_{\gamma+1/2}(2^j x - k) + R(x) \tag{51}$$

where R is a C^∞ random process, and the $\xi_{j,k}$ are I.I.D. standard centered Gaussians. The following result extends Theorem 1 to the case of decentered FBMs.

Theorem 4 *Let f be an arbitrary L^2 function. Let*

$$X(x) = f(x) + B^\gamma(x).$$

With probability 1, the sample paths of X satisfy

$$\forall x_0 \in \mathbb{R}, \quad \limsup_{x \to x_0} \frac{|X(x) - X(x_0)|}{|x - x_0|^\gamma} > 0.$$

Therefore the Hölder exponent of X satisfies

$$a.s. \qquad \forall x \qquad h_X(x) = \inf\left(h_f(x), \gamma\right).$$

An illustration of this result is shown in Figure 2 for the theoretical and estimated spectrum of $X = f + B^\gamma$ with f a multifractal random walk (see [32] for its definition) and $\gamma = 0.7$.

Remark: The prevalent implication of this theorem is that the Hölder exponent of almost every continuous function is everywhere at most γ. Since this is true for any $\gamma > 0$ we recover that the Hölder exponent of almost every continuous function vanishes everywhere, see [16].

Proof of Theorem 4: Let us denote by $C_{j,k}$ the coefficients of X on the wavelet basis generated by $\psi_{\gamma+1/2}$. We call a (C, j_0)-*slow point* a point x_0 where the sample path of $X(x)$ satisfies

$$|X(x) - X(x_0)| \le C|x - x_0|^\gamma$$

for any x such that $|x - x_0| \le 2^{-j_0}$. If x_0 is such a point, using Theorem 2 for all couples $(j, k_j(x_0))$ such that $j \ge j_0$, we have $|C_{j,k_j(x_0)}| \le C2^{-\gamma j}$, which, using (51), can be rewritten

$$\forall j \ge 0, \qquad |2^{-\gamma j}\chi_{j,k_j(x_0)} + f_{j,k_j(x_0)}| \le C2^{-\gamma j},$$

where $f_{j,k}$ denotes the wavelet coefficients of f. Let

$$p_{j,k} = \mathbb{P}(|C_{j,k}| \le C2^{-\gamma j}).$$

Then

$$p_{j,k} = \sqrt{\frac{2}{\pi}} \int_{-C-2^{\gamma j}f_{j,k}}^{C-2^{\gamma j}f_{j,k}} e^{-x^2/2}dx \le 2C\sqrt{\frac{2}{\pi}} := p_C.$$

The end of the proof follows just as in the Brownian case.

5 The multifractal formalism

A common feature shared by (11), (12), (42) and (49) is that all the pointwise exponents we considered can be deduced from a countable number of quantities indexed by the dyadic cubes, and the derivation is performed on a log-log plot bearing on the cubes that contain the point x_0. Therefore, these examples all fit in the following general framework for sets of positive quantities e_λ indexed by a subset of the dyadic cubes, for which we introduce the following definition:
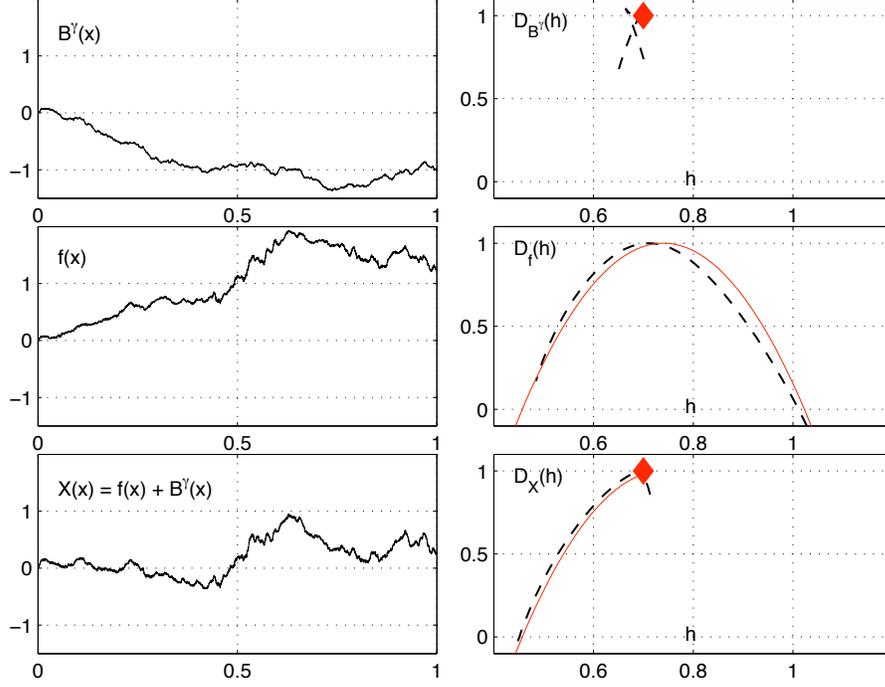
Figure 2: **Superimposition of functions.** Right column:Theoretical multifractal spectrum (red solid line) together with its Leader based estimation (dashed black line) for three different processes (left column). Top, fractional Brownian motion (FBM) with $\gamma = 0.7$, ; Middle, multifractal random walk (MRW) with scaling function $\zeta(p) = c_1 p + c_2 p^2/2$ ($c_1 = 0.72$ and $c_2 = -0.04$) [5] ; Bottom additive superimposition of the two previous processes. The bottom right plot clearly shows that for additive superimposition the actual Hölder exponent at each x_0 actually corresponds to the minimum of each function, hence the Hölder exponents of MRW that are larger than $\gamma = 0.7$ are no longer present in FBM + MRW, as detected by the leader based estimation.

- The e_λ are *hierarchical* if there exists $C > 0$ such that, if λ' is a "child" of λ (i.e. $\lambda' \subset \lambda$ and $j' = j + 1$) then $e_{\lambda'} \leq C \cdot e_\lambda$;

- The e_λ are *strictly hierarchical* if $\lambda' \subset \lambda \Rightarrow e_{\lambda'} \leq e_\lambda$;

- The (e_λ) belong to $C^\alpha(x_0)$ if, for j large enough, $e_{\lambda_j(x_0)} \leq 2^{-\alpha j}$.

- The (e_λ) belong to $I^\alpha(x_0)$ if, for j large enough, $e_{\lambda_j(x_0)} \geq 2^{-\alpha j}$.

- The *pointwise exponent* associated with the e_λ is

$$h(x_0) = \sup\{\alpha : (e_\lambda) \in C^\alpha(x_0)\} = \liminf_{j \to +\infty} \left(\frac{\log\left(e_{\lambda_j(x_0)}\right)}{\log(2^{-j})} \right) \qquad (52)$$

25

- The *upper pointwise exponent* is

$$\tilde{h}(x_0) = \inf\{\alpha : \ (e_\lambda) \in I^\alpha(x_0)\} = \limsup_{j \to +\infty} \left(\frac{\log\left(e_{\lambda_j(x_0)}\right)}{\log(2^{-j})} \right). \tag{53}$$

- The structure function is

$$S_j(p) = 2^{-dj} \sum_{\lambda \in \Lambda_j} (e_\lambda)^p$$

- The scaling function and the upper scaling function are respectively

$$\eta(p) = \liminf_{j \to +\infty} \left(\frac{\log\left(S_j(p)\right)}{\log(2^{-j})} \right) \quad \text{and} \quad \tilde{\eta}(p) = \limsup_{j \to +\infty} \left(\frac{\log\left(S_j(p)\right)}{\log(2^{-j})} \right). \tag{54}$$

- The discrete Besov spaces $\mathcal{B}_p^s$ are defined (for $s, p \in \mathbb{R}$) by

$$(e_\lambda)_{\lambda \in \Lambda} \in \mathcal{B}_p^s \iff \exists C \ \forall j : \quad 2^{-dj} \sum_{\lambda \in \Lambda_j} (e_\lambda)^p \leq C \cdot 2^{-spj}, \tag{55}$$

- The upper Besov spaces $\tilde{\mathcal{B}}_p^s$ are defined (for $s, p \in \mathbb{R}$) by

$$(e_\lambda)_{\lambda \in \Lambda} \in \tilde{\mathcal{B}}_p^s \iff \exists C \ \exists j_n \to +\infty : \quad 2^{-dj_n} \sum_{\lambda \in \Lambda_{j_n}} (e_\lambda)^p \leq C \cdot 2^{-spj_n}. \tag{56}$$

Here Λ_j is a subset of the set of dyadic cubes of width 2^{-j}. In practice, it will usually consist in all the dyadic cubes included in a given bounded domain, and therefore this set will be finite, of cardinality $\sim 2^{dj}$.

The irregularity condition $I^\alpha(x_0)$ has been considered since the 70's in the Hölder setting in order to study irregularity properties of stochastic processes, and more recently by M. Clausel in [9]. In the measure setting, it has been considered by Brown, Michon and Peyrière in [6] and by Tricot in [39, 40].

The purpose of introducing strictly hierarchical sequences is that it is a condition satisfied by sequences of wavelet leaders. It will be particularly relevant in Section 5.5 where we construct counterexamples to the multifractal formalism.

Note that the definition of η that we give here is in agreement with (6). Among the above definitions of discrete Besov spaces, the only one which defines a vector space is (55) when p is positive. In that case, $\mathcal{B}_p^s$ is a Banach space if $p \geq 1$ and, if $0 < p < 1$, it is a quasi-Banach space endowed with a metric defined as follows: Let $e = (e_\lambda)$ and $f = (f_\lambda)$ be two sequences in $\mathcal{B}_p^s$; then

$$dist(e, f) = \sup_{j \geq 0} \left(2^{(sp-d)j} \sum_k (|e_\lambda - f_\lambda|)^p \right). \tag{57}$$

26

The spaces $\mathcal{B}_p^s$ are closely related with Besov spaces; indeed, if $q = +\infty$, a function f belongs to $B_p^{s;\infty}$ if $(c_k) \in l^p$ and if its wavelet coefficients c_λ satisfy (55). Therefore the wavelet decomposition establishes an isomorphism between the discrete and the continuous Besov spaces.

If the (e_λ) are the wavelet leaders of f, (55) yields the wavelet characterization of the oscillation spaces considered in [20].

Note that

$$\left.\begin{array}{ll} \text{if} \quad p > 0, \quad \eta(p) = \sup\{s: \ f \in \mathcal{B}_p^{s/p}\} \quad \text{and} \quad \tilde{\eta}(p) = \sup\{s: \ f \in \tilde{\mathcal{B}}_p^{s/p}\}; \\ \text{if} \quad p < 0, \quad \eta(p) = \inf\{s: \ f \in \mathcal{B}_p^{s/p}\} \quad \text{and} \quad \tilde{\eta}(p) = \inf\{s: \ f \in \tilde{\mathcal{B}}_p^{s/p}\}. \end{array}\right\} \qquad (58)$$

In the measure case, $e_\lambda = \mu(3\lambda)$, in the pointwise Hölder case, $e_\lambda = d_\lambda$, and in the $T_\alpha^p(x_0)$ case, $e_\lambda = d_\lambda^p$, so that the general setting in which we work in this section covers all the previous cases we already considered. Therefore, the results we will obtain in this general setting will be valid for all the settings we previously considered: measures, Hölder exponent, p-exponent. However, more precise properties can hold in these particular settings since the corresponding sequences satisfy additional properties (for instance, the $\mu(3\lambda)$ and the d_λ are strictly hierarchical).

We will reformulate the fundamental idea due to G. Parisi and U. Frisch in this general setting. They gave an interpretation of the nonlinearity of the scaling function as the signature of the presence of different pointwise exponents (see [37] and also [14] for applications, particularly in the setting of invariant measures of dynamical systems).

In cases where a whole range of pointwise exponents are present, we will see that the scaling function gives an information about the size of the set of points that display exactly this exponent. A first question is to determine what precise mathematical meaning should be given to the word "size" in this context. Assume for instance that all exponents in the range $[h_{min}, h_{max}]$ are obtained, with $h_{min} < h_{max}$, and let E_H denote the corresponding isohölder sets, i.e.

$$E_H = \{x: \ h(x) = H\}.$$

The first idea is to try the most usual mathematical meaning of "size" in analysis, i.e. the Lebesgue measure; however, this cannot be the right notion here; indeed, because of the countable additivity of the Lebesgue measure, and since bounded domains have a finite measure, it follows that $meas(E_H) = 0$ for almost all values of H. A standard way in order to compare the sizes of different sets that have a vanishing Lebesgue measure is to compute their *fractal dimension*. The idea of associating fractal sets to measures or functions can be traced back to the works of B. Mandelbrot in the 70s and 80s [30, 31]. Independently, Orey and Taylor in [35] were the first to consider dimensions of the sets of points where the modulus of continuity of the Brownian motion has a particular order of magnitude (slow and fast points of Brownian motion), which is very close to the idea of multifractal analysis.

5.1 Fractal dimensions and spectrums of singularities

We recall the different notions of dimension which are used. The simplest one is the *box dimensions*.

Definition 18 *Let $A \subset \mathbb{R}^d$; if $\epsilon > 0$, let $N_\epsilon(A)$ be the smallest number of sets of radius ϵ required to cover A.*
   The upper box dimension of A is

$$\overline{dim}_B(A) = \limsup_{\epsilon \to 0} \frac{\log N_\epsilon(A)}{-\log \epsilon}.$$

The lower box dimension of A is

$$\underline{dim}_B(A) = \liminf_{\epsilon \to 0} \frac{\log N_\epsilon(A)}{-\log \epsilon}.$$

One important drawback of the box dimensions is that, if A is dense, then the box dimensions take invariably the value d. Since most multifractal functions of interest have dense sets of Hölder singularities, box dimensions are unable to draw any distinction between the sizes of these sets. This explains why box dimensions are not used in the definition of the spectrum of singularities (see Definition 20). However, they are an intermediate step in the definition of the packing dimension that we give below.
   Two alternative definitions of fractal dimension have been introduced and are used in multifractal analysis. In order to define the Hausdorff dimensions, we need to recall the notion of δ-dimensional Hausdorff measure.

Definition 19 *Let $A \subset \mathbb{R}^d$. If $\epsilon > 0$ and $\delta \in [0, d]$, we denote*

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

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

$$mes_\delta(A) = \lim_{\epsilon \to 0} M_\epsilon^\delta.$$

There exists $\delta_0 \in [0, d]$ such that

$$\forall \delta < \delta_0, \quad mes_\delta(A) = +\infty$$
$$\forall \delta > \delta_0, \quad mes_\delta(A) = 0.$$

This critical δ_0 is called the Hausdorff dimension of A, and is denoted by $dim(A)$.

The other notion of dimension we will use is the *packing dimension* which was introduced by C. Tricot, see [39, 40] (see also Chap. 5 of [33]): The *lower packing dimenson* is

$$Dim(A) = \inf \left\{ \sup_{i \in \mathbb{N}} \left(\underline{dim}_B A_i : A \subset \bigcup_{i=1}^{\infty} A_i \right) \right\} \tag{59}$$

(the infimum is taken over all possible partitions of A into a countable collection A_i). We will use this alternative notion in order to bound the dimensions of some sets of singularities. The dimensions we introduced can be compared as follows, see [33, 39, 40],

$$\forall A \subset \mathbb{R}^d, \quad dim(A) \leq Dim(A) \leq \underline{dim}_B(A) \leq \overline{dim}_B(A). \tag{60}$$

A usual way to classify fractal sets in mathematics is to consider their Hausdorff dimensions; this motivates the following definition.

Definition 20 *Let $(e_\lambda)_{\lambda \in \Lambda}$ be a hierarchical dyadic function, and let*

$$E_H = \{x : \quad h(x) = H\}.$$

The spectrum of singularities associated with the $(e_\lambda)_{\lambda \in \Lambda}$ is the function d defined by

$$d(H) = dim(E_H)$$

(we use the convention $d_f(H) = -\infty$ if H is not a Hölder exponent of f). The support of the spectrum is the set of values of H for which $E_H \neq \emptyset$.

Remarks: We can consider many alternative definitions which are variants of this one and are obtained by considering the set

$$F_H = \{x : \quad \tilde{h}(x) = H\}.$$

We can also consider the sets of points where these exponents are larger or smaller than a given value, and finally the corresponding Hausdorff or packing dimensions. This leads to a large collection of possible spectra. They will be useful in Section 5.4 in order to obtain optimal upper bounds for the dimensions of the spectra.

The spectrum of singularities of many mathematical functions or measures can be determined directly from its definition. On the opposite, for many real-life signals, whose Hölder exponent is expected to be everywhere discontinuous, the numerical determination of their Hölder regularity is not feasible, and therefore, one cannot expect to have direct access to their spectrum of singularities. In such cases, one has to find an indirect way to compute $d(H)$; the multifractal formalism is a formula which is expected to yield the spectrum of singularities of f from the scaling function, which is numerically computable. (We will derive such a formula in Section 5.2.) Mathematically, these quantities are interpreted as indicating that f belongs to a certain subset of a family of function spaces.

The pointwise exponent of a sequence often takes all values on a whole interval $[h_{min}, h_{max}]$, where $h_{min} < h_{max}$; in such cases, the computation of the spectrum of

singularities requires the study of an infinite number of fractal sets E_H; this explains the introduction of the term "multifractal" by G. Parisi and U. Frisch in [37]. Note that speaking of the "multifractal analysis" of either a mathematical function or a signal derived from real-life data does not imply that it is assumed to be multifratal: Such an analysis often concludes that it is monohölder for instance.

5.2 Derivation of the multifractal formalism

Let us now show how the spectrum of singularities is expected to be recovered from the scaling function. The definition of the scaling function (54) roughly means that, for j large,

$$S_j(p) \sim 2^{-\eta(p)j}.$$

Let us estimate the contribution to $S_j(p)$ of the dyadic cubes λ that cover the points of E_H. By definition of E_H, they satisfy

$$e_\lambda \sim 2^{-Hj};$$

by definition of $d(H)$, since we use cubes of the same width 2^{-j} to cover E_H, we need about $2^{-d(H)j}$ such cubes; therefore the corresponding contribution is of the order of magnitude of

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

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

$$\eta(p) = \inf_H (d - d(H) + Hp). \tag{61}$$

We will show that the scaling function $\eta(p)$ is a concave function on $\mathbb{R}$, which is in agreement with the fact that the right-hand side of (61) necessarily is a concave function (as an infimum of a family of linear functions) no matter whether $d(H)$ is concave or not. However, if $d(H)$ also is a concave function, then the Legendre transform in (61) can be inverted (as a consequence of the duality of convex functions), which justifies the following assertion:

Definition 21 *A sequence (e_λ) follows the multifractal formalism if its spectrum of singularities satisfies*

$$d(H) = \inf_{p \in \mathbb{R}} (d - \eta(p) + Hp). \tag{62}$$

Note that an alternative method in order to recover the spectrum is proposed in [8]; it is commonly used in practice since it allows to get rid of the Legendre transform.

The derivation exposed above is not a mathematical proof, and the determination of the range of validity of (62) (and of its variants) is one of the main mathematical problems concerning multifractal analysis. It does not hold in complete generality. However, three types of verifications can be performed:

- The multifractal formalism is proved under additional assumptions on the e_λ (usually, assuming that it is derived from a self-similar function or measure).

- It is proved for a "large" subset of the function space considered: We refer to [13, 20] and references therein for "generic results" of multifractality, either in the sense either of Baire categories or of prevalence.

- The multifractal formalism is shown to yield an upper bound of the spectrum of singularities, see Section 5.4.

Note that, in applications, it often 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 (62) but only with having its right-hand side defined in a numerically precise way.

5.3 Properties of the scaling function

We start by proving that the function η is concave on $\mathbb{R}$, a property that was used in the derivation of the multifractal formalism

Proposition 4 *The function η defined by (54) is concave on $\mathbb{R}$.*

Note that $\tilde{\eta}$ (also defined by (54)) has no reason to be concave in general. However, in practical applicatons, the liminf and limsup in (54) often coincide (and are therefore true limits), in which case $\tilde{\eta} = \eta$ and is therefore concave. In order to prove Proposition 4, we will need the following lemma.

Lemma 5 *Let $(a_i)_{i \in \mathbb{N}}$ be a sequence of positive real numbers . The function $\omega : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ $(= \mathbb{R} \cup \{+\infty, -\infty\})$ defined by*

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

is convex on $\mathbb{R}$.

Proof of Lemma 5: We need to check that

$$\forall p, q \in \mathbb{R}, \ \ \forall \alpha \in]0, 1[, \ \ \ \omega \left(\alpha p + (1 - \alpha)q \right) \leq \alpha \omega(p) + (1 - \alpha)\omega(q). \tag{63}$$

Consider the sequences

$$A = (a_1^{\alpha p}, ... a_N^{\alpha p}, \cdots) \ \ \text{and} \ \ B = (a_1^{(1-\alpha)q}, ... a_N^{(1-\alpha)q}, \cdots);$$

Hölder's inequality applied with the conjugate exponents $p' = 1/\alpha$ and $q' = 1/(1 - \alpha)$ yields

$$\sum_{i=1}^{\infty} a_i^{\alpha p + (1-\alpha)q} \leq \left(\sum_{i=1}^{\infty} a_i^p \right)^{\alpha} \left(\sum_{i=1}^{\infty} a_i^q \right)^{1-\alpha}.$$

Taking logarithms on both sides of this inequality yields (63).

We will now show that η is concave on $\mathbb{R}$. For each j, one applies Lemma 5 to the e_λ. We obtain that, $\forall j$, the function

$$p \to \log \left(\sum_{\lambda \in \Lambda_j} d_\lambda^p \right)$$

is convex; therefore, after dividing by $\log(2^{-j})$, we obtain a concave function; since concavity is preserved by taking infimums and pointwise limits, the concavity of the scaling function follows.

We will make the following uniform regularity and irregularity assumptions:

$$\exists C_1, C_2, A, B, \quad \text{such that} \qquad \forall \lambda \quad C_1 2^{-Bj} \le e_\lambda \le C_2 2^{-Aj}. \tag{64}$$

Note that in the measure case and in Hölder exponent case, one can pick $A = 0$. In the Hölder case, the uniform regularity assumption means that $A > 0$. When the e_λ are wavelet leaders, the assumption on the lower bound implies that the function f considered has no C^∞ components. In the measure case, it implies that μ does not vanish on a set of nonempty interior.

In order to better understand the range of validity of the multifractal formalism, we have to explore the properties of the Legendre transform of the scaling function. Therefore, we introduce the *Legendre spectrum* associated with the sequence (e_λ):

$$L(H) = \inf_{p \in \mathbb{R}} (d + Hp - \eta(p)). \tag{65}$$

The validity of the multifractal formalism therefore states that the Legendre spectrum coincides with the spectrum of singularities.

Lemma 6 *Let (e_λ) be a sequence satisfying (64);*

$$if \quad p \ge 0, \quad then \quad Ap \le \eta(p) \le Bp,$$

$$if \quad p \le 0, \quad then \quad Bp \le \eta(p) \le Ap.$$

Proof of Lemma 6: The assumption (64) implies that, if $p > 0$,

$$(C_1)^p 2^{-Bpj} \le S_j(p) \le (C_2)^p 2^{-Apj},$$

and the inequalities are reversed if $p < 0$. The result follows from the definition of the scaling function (54).

It follows from Lemma 6 that $\eta(0) = 0$, $L(H) = -\infty$ if $H \le A$ or $H \ge B$, and

$$\forall p \in \mathbb{R}, \quad A \le \eta'(p) \le B.$$

Proposition 5 *Let (e_λ) be a sequence satisfying (64). Let*

$$H_{min} = \sup\{A : \ (64) \ \ holds\} \ \ and \ \ H_{max} = \inf\{B : \ (64) \ \ holds\}.$$

The Legendre spectrum $L(H)$ of (e_λ) is a concave function satisfying

$$if \ \ H \notin [H_{min}, H_{max}], \quad L(H) = -\infty. \tag{66}$$

$$if \ \ H \in [H_{min}, H_{max}], \quad 0 \leq L(H) \leq d. \tag{67}$$

There exist H^1_{med} and $H^2_{med} \in [H_{min}, H_{max}]$ such that

- *$L(H)$ is strictly increasing on $[H_{min}, H^1_{med}]$,*
- *$\forall H \in [H^1_{med}, H^2_{med}]$, $L(H) = d$,*
- *$L(H)$ is strictly decreasing on $[H^2_{med}, H_{max}]$.*

Remark: It follows immediately from the Legendre transform formula (65) that $H^1_{med} = \eta'_r(0)$ and $H^2_{med} = \eta'_l(0)$ (where η'_r and η'_l respectively denote the right and left derivatives of η); therefore, if η is differentable at 0, then H^1_{med} and H^2_{med} coincide; in that case, we will denote their common value by H_{med}. Formula (67) means that the multifractal formalism cannot yield "negative" dimensions, when applied to one given set of data. Such a phenomenon can only happen after an averaging over a large number of realizations, in a probabilistic setting, i.e. when defining the scaling function as $\mathbb{E}((e_\lambda)^p)$ and not sample path by sample path.

We willl call *admissible spectral function* any function $L(H)$ which satisfies the conditions listed in Proposition 5. An *admissible scaling function* is any function $\eta(p)$ which is the Legendre transform of an admissible spectral function. Theorem 6 will imply that these conditions characterize the functions $L(H)$ which are the Legendre spectra of a given sequence (e_λ); and, equivalently, they also characterize the functions $\eta(p)$ which are scaling functions.

Proof of Proposition 5: The fact that $L(H)$ is concave follows directly from its definition as a Legendre transform. Lemma 6 imples that, if $p \geq 0$, then $\eta(p) \geq H_{min}p$ and, if $p \leq 0$, then $\eta(p) \geq H_{max}p$; (66) follows from these lower bounds.

Since $\eta(0) = 0$, it follows that $\forall H$, $L(H) \leq d$, and, the fact that $\forall H \in [H^1_{med}, H^2_{med}]$, $L(H) = d$ directly follows from (65). The increasing and then decreasing property follows from this observation together with the concavity property.

Let us now prove that $L(H)$ is nonnegative on $[H_{min}, H_{max}]$. From the definition of $L(H)$, it suffices to prove that

$$\forall H \in (H_{min}, H_{max}), \ \forall p \in \mathbb{R}, \qquad \eta(p) \leq d + Hp \tag{68}$$

(indeed, by continuity of L, the result will also be true for the extreme points H_{min} and H_{max}, if L differs from $-\infty$ at these points). By definition of H_{min} and H_{max}, there exist sequences $l_n, j_n \in \mathbb{N}$, and dyadic cubes μ_n and λ_n such that

$$|\mu_n| = 2^{-l_n}, \quad |\lambda_n| = 2^{-j_n}, \quad e_{\mu_n} \leq 2^{-Hl_n}, \quad \text{and} \quad e_{\lambda_n} \geq 2^{-Hj_n}.$$

Therefore, if $p < 0$,

$$2^{-dl_n} \sum_{\lambda \in \Lambda_{l_n}} |e_\lambda|^p \geq 2^{-(d+Hp)l_n}$$

and if $p > 0$,

$$2^{-dj_n} \sum_{\lambda \in \Lambda_{j_n}} |e_\lambda|^p \geq 2^{-(d+Hp)j_n},$$

and (68) follows.

5.4 Upper bound of the spectrums

Corollary 2 will state that $L(H)$ yields upper bound for the spectrum of singularities of any sequence. It will be a straightforward consequence of the following sharper bounds that involve the different types of singularities we introduced.

Theorem 5 *Let* (e_λ) *be a sequence and*

$$J^H = \{x_0 : (e_\lambda) \in C^H(x_0)\}, \ G^H = (J^H)^c, \ F^H = \{x_0 : (e_\lambda) \in I^H(x_0)\}, \ K^H = (F^H)^c.$$

- *If* $(e_\lambda)_{\lambda \in \Lambda} \in \mathcal{B}_p^s$ *with* $p > 0$, *then* $dim(G^H) \leq d - sp + Hp$.

- *If* $(e_\lambda)_{\lambda \in \Lambda} \in \tilde{\mathcal{B}}_p^s$ *with* $p > 0$, *then* $Dim(F^H) \leq d - sp + Hp$.

- *If* $(e_\lambda)_{\lambda \in \Lambda} \in \mathcal{B}_p^s$ *with* $p < 0$, *then* $dim(K^H) \leq d - sp + Hp$.

- *If* $(e_\lambda)_{\lambda \in \Lambda} \in \tilde{\mathcal{B}}_p^s$ *with* $p < 0$, *then* $Dim(J^H) \leq d - sp + Hp$.

Remarks: Using (11) and (12), this theorem, if applied to the sequence $e_\lambda = \mu(3\lambda)$, allows to recover the bounds for either packing or Hausdorff dimensions of the sets of points where the exponents h_μ and $\overline{h}_\mu$ of a measure μ are larger or smaller than H, see [6]. Similarly, if f is a uniform Hölder function, (42) shows that Theorem 5, if applied to the sequence $e_\lambda = d_\lambda$ yields the same bounds for the Hölder exponent. Finally, if f belongs to L^p, Theorem 3 implies a similar result for the p-exponent, using the sequence $e_\lambda = d_\lambda^p$.

In order to prove this theorem, we start by establishing simple estimates on the number of small and large coefficients e_λ.

Lemma 7 *Let*

$$\begin{cases} G_{j,H} = \{\lambda : e_\lambda \geq 2^{-Hj}\}, & N_{j,H} = Card(G_{j,H}) \\ F_{j,H} = \{\lambda : e_\lambda \leq 2^{-Hj}\}, & M_{j,H} = Card(F_{j,H}). \end{cases}$$

- *If* $(e_\lambda)_{\lambda \in \Lambda} \in \mathcal{B}_p^s$ *with* $p > 0$, *then* $\forall j, \ N_{j,H} \leq 2^{(d-sp+Hp)j}$.

- *If* $(e_\lambda)_{\lambda \in \Lambda} \in \tilde{\mathcal{B}}_p^s$ *with* $p > 0$, *then* $\exists j_n \to \infty : N_{j_n,H} \leq 2^{(d-sp+Hp)j_n}$.

- If $(e_\lambda)_{\lambda \in \Lambda} \in \mathcal{B}_p^s$ with $p < 0$, then $\forall j$, $M_{j,H} \leq 2^{(d-sp+Hp)j}$.

- If $(e_\lambda)_{\lambda \in \Lambda} \in \tilde{\mathcal{B}}_p^s$ with $p < 0$, then $\exists j_n \to \infty$: $M_{j_n,H} \leq 2^{(d-sp+Hp)j_n}$.

Proof of Lemma 7: Let us consider the first assertion. Since, $(e_\lambda)_{\lambda \in \Lambda} \in \mathcal{B}_p^s$ with $p > 0$,

$$\exists C, \ \forall j, \quad 2^{(sp-d)j} \sum |e_\lambda|^p \leq C;$$

therefore, by restricting the sum to the elements of $G_{j,H}$, we obtain

$$\exists C, \ \forall j, \quad 2^{(sp-d)j} N_{j,H} 2^{-Hpj} \leq C,$$

so that $N_{j,H} \leq C 2^{(d-sp+Hp)j}$. The proof of the third case is similar because, if p is negative, the condition $e_\lambda \leq 2^{-Hj}$ becomes $(e_\lambda)^p \geq 2^{-Hpj}$.

The second and the fourth case also follow by the same arguments applied to a subsequence j_n.

Let (A_j) be a sequence of sets; we recall that $\overline{\lim}(A_j)$ denotes the set of points that belong to an infinite number of the A_j and $\underline{\lim}(A_j)$ denotes the set of points that belong to all A_j for j large enough. It follows from the definitions of G_H and K_H that

$$G_H = \overline{\lim}(G_{j,H}) \quad \text{and} \quad K_H = \overline{\lim}(F_{j,H}).$$

Furthermore, if j_n is the sequence that appears in the definition of $\tilde{\mathcal{B}}_p^s$, then it follows from the definitions of J_H and F_H that

$$F_H = \underline{\lim}(G_{j,H}) \subset \underline{\lim}(G_{j_n,H}) \quad \text{and} \quad J_H = \underline{\lim}(F_{j,H}) \subset \underline{\lim}(F_{j_n,H}).$$

Therefore, in order to prove Theorem 5, there only remains to prove the following estimates on Hausdorff and Packing dimensions.

Lemma 8 *Let $j_n \to +\infty$ and A_{j_n} be a union of at most $2^{\omega j_n}$ dyadic cubes $\lambda_{j_n,k}$ of width 2^{-j_n}. Then*

$$dim\left(\overline{\lim}(A_{j_n})\right) \leq \omega \quad \text{and} \quad Dim\left(\underline{\lim}(A_{j_n})\right) \leq \omega.$$

Proof: We use $\bigcup_{l \geq n} A_{j_l}$ for covering of $\overline{\lim}(A_{j_n})$; it is an ϵ-covering as soon as $\sqrt{d} 2^{-j_n} \leq \epsilon$. It follows that, for this covering,

$$\sum_{l \geq n} \sum_k (diam(\lambda_{j_l,k})^\delta \leq \sum_{l \geq n} 2^{\omega j_l} 2^{-\delta j_l}$$

which is finite as soon as $\delta > \omega$.

Note that $\underline{\lim}(A_{j_n})$ is the countable union of the sets

$$B_{j_n} = \bigcap_{l \geq n} A_{j_l}.$$

For any $l \geq n$, B_{j_n} is included in A_{j_l}, i.e. in the union of $2^{\omega j_l}$ dyadic cubes of width 2^{-j_l}, so that the lower box dimension of B_{j_n} is bounded by ω, and the lower packing dimension of $\underline{\lim}(A_{j_n})$ is bounded by ω.

Corollary 2 *Let (e_λ) be a sequence. Then*

$$d(H) \leq \inf_{p \in \mathbb{R}} \left(Hp - \eta(p) + d \right). \tag{69}$$

Proof: Clearly, $\forall \epsilon > 0$, $E_h \subset G^{H+\epsilon}$. Furthermore, using (58),

$$\forall p > 0, \ \forall \delta > 0, \quad (e_\lambda) \in \mathcal{B}_p^{\frac{\eta(p)}{p} - \delta},$$

so that, by Theorem 5,

$$dim(E_H) \leq d - \left(\frac{\eta(p)}{p} - \delta \right) p + (H + \epsilon)p.$$

Since this is true $\forall \epsilon, \delta > 0$, we obtain the expected bound for $p > 0$.

Similarly, we have $\forall \epsilon > 0$, $E_h \subset J^{H-\epsilon}$, and using (58),

$$\forall p < 0, \ \forall \delta > 0, \quad (e_\lambda) \in \mathcal{B}_p^{\frac{\eta(p)}{p} + \delta},$$

so that, by Theorem 5,

$$dim(E_H) \leq d - \left(\frac{\eta(p)}{p} + \delta \right) p + (H - \epsilon)p.$$

Since this is true $\forall \epsilon, \delta > 0$, we obtain that

$$\forall p < 0, Dim(E_H) \leq Hp - \eta(p) + d;$$

the required bound for $p < 0$ follows from (60).

5.5 Validity of the multifractal formalism

The following theorem states that, except for the very particular case of a linear scaling function, the multifractal formalism is not valid unless additional properties are assumed for the e_λ.

Theorem 6 *Let (e_λ) be a sequence. If their scaling function η is linear, i.e. if*

$$\exists \alpha > 0 \quad \text{such that} \quad \forall p \in \mathbb{R}, \quad \eta(p) = \alpha p, \tag{70}$$

then the multifractal formalism holds for the (e_λ), and their spectrum of singularities is given by

$$\begin{cases} d(\alpha) = 1 \\ d(H) = -\infty \quad \text{if} \quad H \neq \alpha. \end{cases}$$

If η is a nonlinear scaling function, then there exists a sequence (e_λ) whose scaling function is η, and whose spectrum of singularities is given by:

$$\begin{cases} d(H_{min}) = d(H_{max}) = 0 \\ d(H^1_{med}) = d(H^2_{med}) = 1 \\ d(H) = -\infty \quad \text{if} \quad H \notin \{H_{min}, H^1_{med}, H^2_{med}, H_{max}\}, \end{cases} \tag{71}$$

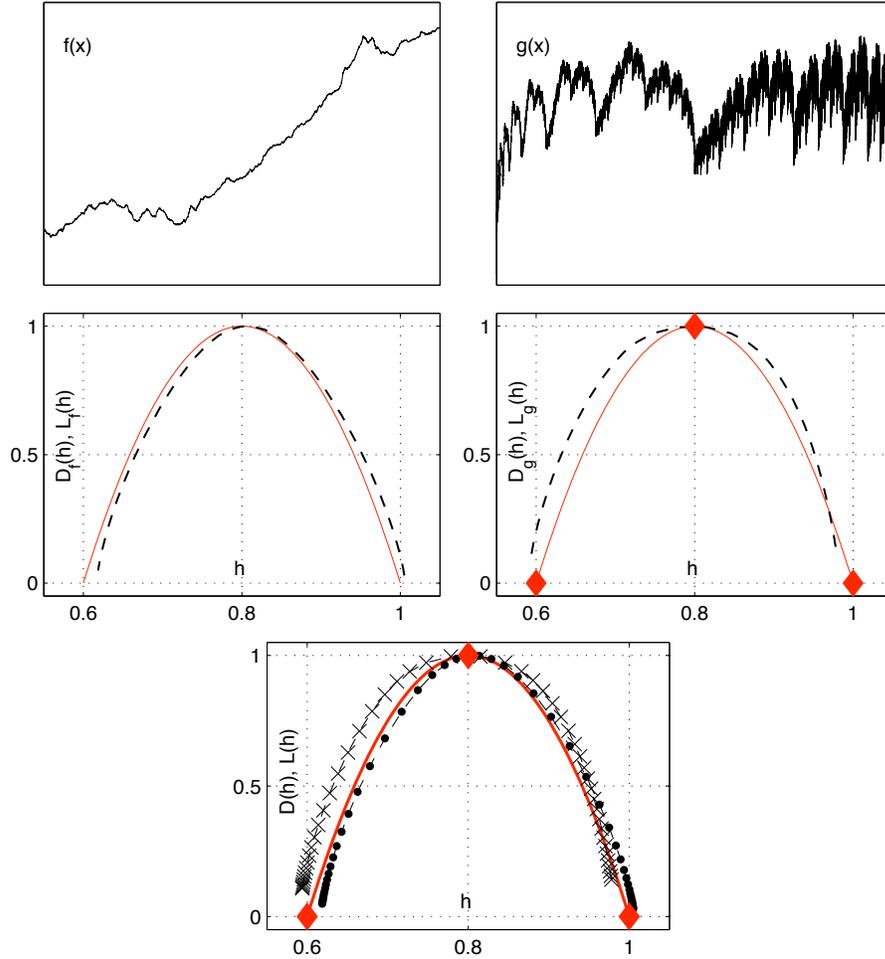*where $H_{min}, H^1_{med}, H^2_{med}$, and H_{max} are defined in Proposition 5.*

36

Figure 3: **Validity of the multifractal formalism.** Sample of MRW (top left) and process contructed according to the counter example for Theorem 6 (top right). Both functions have identical Legendre spectra but different singularity spectra. Middle left: Singularity and Legendre spectra (red solid line) are equal for MRW, their leader based estimate is superimposed (black dashed line with dots). Middle right: Singularity spectrum (red diamonds) and Legendre spectrum (red solid line) together with the leader based estimate (black dashed line with crosses). Bottom: All spectra are superimposed: Leader based estimates for MRW an for the counter example are very close and satisfactorily match the common Legendre spectrum of both processes. While for MRW, the leader based multifractal formalism enables to estimate the singularity spectrum, this is no longer the case for the counter example tailored in purpose. *(The sample size is $N = 2^{18}$. The counter example is synthesized using Daubechies4 wavelets. Estimations are performed with Daubechies3 wavelets and ordinary linear regressions over scales $j = [2, 10]$. Results for MRW $(c_1 = 0.8, c_2 = -0.02)$ are obtained as means over 50 realizations.)*

Remarks: A sequence whose spectrum is given by (71) clearly is a counterexample to the multifractal formalism. Note that, though, the linear case given by (70) can seem exceptional, it covers important examples such as the Weierstrass functions, or the FBM, see [1]. We state and prove Theorem 6 in dimension one; one easily deduce counterexamples in several dimensions.

Figure 3 shows the function whose wavelet coefficients (and wavelet leaders) are the (e_λ) constructed in Proof of Theorem 6 for the counterexample with $H_{min} = 0.6$, $H_{med}^1 = H_{med}^2 = 0.8$ and $H_{max} = 1$. This function has the same Legendre spectrum that a MRW - for which the multifractal formalism is valid - but not the same spectrum of singularity.

Proof of Theorem 6: Assume that the scaling function is linear and given by

$$\eta(p) = \alpha p.$$

Then $L(H) = -\infty$ except for $H = \alpha$; Corollary 2 implies in this case that $d(H) \leq -\infty$ for $H \neq \alpha$; therefore only one Hölder exponent is present, so that $\forall x$, $h(x) = \alpha$; it follows that $d(\alpha) = 1$, and the multifractal formalism therefore holds.

In order to construct a counterexample in the nonlinear case, we will define explicitely the e_λ at each scale for $\lambda \subset [0, 1]$. We fix a $j \geq 0$. The $e_\lambda = e_{j,k}$ are defined from $k = 0$ to $2^j - 1$ and are increasing when k increases. For $k = 0$ to $[2^{L(H_{max})j}]$ we define $e_{j,k} = 2^{-H_{max}j}$. When H decreases form H_{max} to H_{med}^2 the function $2^{L(H)j}$ is continuous and increases up to 2^j. Each time it reaches a new integer value k we define H_k as the only $H \in (H_{med}^2, H_{max}]$ such that

$$2^{L(H_k)j} = k;$$

the value of the k-th coefficient is

$$e_{j,k} = 2^{-H_k j}.$$

We do this until we have reached the value $k = 2^{j-1} - 1$. After this, we jump to the value of $H \in [H_{min}, H_{med}^1)$ such that $L(H) = 2^{j-1}$. We continue similarly: When H decreases, each time the function $2^{L(H)j}$ reaches a new integer value k we define $H_{2^j - k}$ as the only $H \in [H_{min}, H_{med}^1]$ such that

$$2^{L(H_{2^j - k})j} = k.$$

The value of the corresponding coefficient is

$$e_{j, 2^j - k} = 2^{-H_{2^j - k} j}.$$

When we reach the value H_{min} we set the remaining coefficients at the value $2^{-H_{min}j}$.

First, note that all coefficients lie between $2^{-H_{max}j}$ and $2^{-H_{min}j}$, so that (64) is satisfied. Let us now determine the Hölder exponent of this sequence. By construction, at each scale, there is (at least) one coefficient of size $2^{-H_{max}j}$ located above $x_0 = 0$ and

another coefficient of size $2^{-H_{min}j}$ located above $x_0 = 1$; therefore, the Hölder exponent at the two extreme points takes the values

$$h(0) = H_{max} \quad \text{and} \quad h(1) = H_{min}.$$

For any $\epsilon > 0$, the $2^{L(H_{med}^2+\epsilon)j}$ first coefficients are smaller than $2^{-(H_{med}^2+\epsilon)j}$ and the last $2^{L(H_{med}^1-\epsilon)j}$ coefficients are larger than $2^{-(H_{med}^1+\epsilon)j}$. As regards the others, either $k \leq 2^{j-1}$ and they lie between $2^{-(H_{med}^2+\epsilon)j}$ and $2^{-H_{med}^2 j}$, or $k > 2^{j-1}$ and they lie between $2^{-(H_{med}^1-\epsilon)j}$ and $2^{-H_{med}^1 j}$. The length of both intervals tends to $1/2$ because $\forall H < H_{med}^1$, $L(H) < 1$ and $\forall H > H_{med}^2$, $L(H) < 1$. Therefore

$$\forall x \in (0, 1/2), \qquad h(x) = H_{med}^2,$$

$$\forall x \in (1/2, 1), \qquad h(x) = H_{med}^1,$$

and (71) is proved.

Let us now show why the Legendre spectrum of the (e_λ) is indeed $L(H)$. First, we recall the definition of the *large deviation spectrum* $\mathcal{L}(H)$ of a sequence. Let

$$N_j(\epsilon, \alpha) = Card\left(\{\lambda \in \Lambda_j : 2^{-(\alpha+\epsilon)j} \leq |e_\lambda| \leq 2^{-(\alpha-\epsilon)j}\}\right);$$

then

$$\mathcal{L}(H) = \lim_{\epsilon \to 0} \limsup_{j \to +\infty} \frac{\log(N_j(\epsilon, \alpha))}{\log(2^j)}$$

By construction,

$$N_j(\epsilon, \alpha) = \sup\left(2^{L(\alpha+\epsilon)j}, 2^{L(\alpha-\epsilon)j}\right) + O(1),$$

and therefore a standard large deviation argument yields that there are $\sim 2^{L(h)j}$ coefficients of size $\sim 2^{-Hj}$, and therefore the Legendre spectrum of this sequence is indeed $L(H)$.

Let us now check that the sequence (e_λ) is strictly hierarchical. Let $H > H_{med}^2$. If $2^{L(H)j} \in \mathbb{N}$, then the coefficient $e_\lambda = e_{j,k}$ of size 2^{-Hj} is such that $k = 2^{L(H)j}$; therefore its distance to the origin is $\delta_\lambda = 2^{-j}2^{L(H)j}$. Since, for each j, the sequence $k \to e_{j,k}$ is increasing, it suffices to prove that, for the same size of coefficient (i.e. if $Hj = H'j'$), and $j' > j$, then $\delta_{\lambda'} \leq \delta_\lambda$, which is equivalent to

$$2^{(1-L(H'))j'} \geq 2^{(1-L(H))j}.$$

This is, in turn equavalent to prove that, if $H > H'$ then,

$$\frac{1 - L(H')}{H'} \leq \frac{1 - L(H)}{H}$$

which follows from the concavity of $L(H)$, and the fact that it reaches its maximum at H_{med}. We don't treat the case $H < H_{med}^1$ since it is easier to deal with: indeed it is a straightforward consequence of the fact that $L(H)$ is increasing on $[H_{min}, H_{med}^1]$. This last assertion implies that the counterexample we constructed also works for the Hölder exponent of functions. Indeed, if we pick the e_λ for wavelet coefficients of the function, the the wavelet leaders will also be the e_λ.

5.6   Some open questions

One can meet several types of multifractal functions. A first one is supplied by functions which present inhomogeneities: They are smoother in some regions than in others. This case is often met in image analysis. Indeed, a natural image is a patchwork of textures with different characteristics. Their spectrums of singularities reflects the multifractal nature of each component, and also of the boundaries (which may also be fractal) where discontinuities appear. In such situations, the determination of a local spectrum of singularities for each 'component' is more relevant. It is defined as follows: If $\Omega \subset \mathbb{R}^d$ is a nonempty open set,

$$d^\Omega(H) = dim(E_H \cap \Omega);$$

clearly, $d(H) = \sup_\Omega d^\Omega(H)$.

By contrast, *homogeneous* multifractal sequences present the same characteristics everywhere. The same remark applies to the function spaces to which the sequence belongs.

The scaling function $\eta^\Omega(p)$ of a sequence (e_λ) restricted to Ω is defined using (54), but starting with the structure function

$$S_j^\Omega(p) = \sum_{\lambda \in \Lambda_j, \, \lambda \subset \Omega} (e_\lambda)^p.$$

For function spaces, one may require a smooth cutoff, so that the scaling function of f restricted to Ω is

$$\eta^\Omega(p) = \inf_{\varphi \in \mathcal{D}, supp(\varphi) \subset \Omega} \eta_{f\varphi}^\Omega(p).$$

Definition 22 *A sequence is Hölder-homogeneous if the function $d^\Omega(H)$ is independent of Ω. It is scaling-homogeneous if the function $\eta^\Omega(p)$ is independent of Ω.*

Note that the same definition holds for functions and measures. The notion of homogeneous function is important for modelling; indeed, with regards to fully developed turbulence, Kolmogorov assumed that the scaling function is *universal*, i.e. that it does not depend on the particular fluid considered, on the limit conditions, on the particular region of he flud considered,... Parisi and Frisch also assumed the same property for the spectrum of singularities. Therefore, the velocity of a turbulent fluid is expected to be an homogeneous multifractal function. Usually, the mathematical functions and stochastic processes which are known to be multifractal are actually homogeneous multifractals. An important open question is to determine which functions $d(H)$ can be the spectrum of singularites of an homogeneous multifractal function. Note also that the counterexamples to the multifractal formalism that we gave are not homogeneous. Therefore, another open question is to determine if one can construct counterexamples that are also homogeneous with the same level of generality.

In real-life data obtained one usually observes two properties: The signal is scaling homogeneous, and the scaling function η is actually given in (54) by a true limit, which means that $\forall p, \eta(p) = \tilde{\eta}(p)$. We do not know what these properties imply on the sequence

(e_λ), and in particular, if these additional properties improve the range of validity of the multifractal formalism for such sequences.

References

[1] P. Abry, S. Jaffard and B. Lashermes, *Wavelet leaders in multifractal analysis,* "Wavelet Analysis and Applications", T. Qian et al. eds., pp. 201–246, "Applied and Numerical Harmonic Analysis", Springer (2006)

[2] P. Abry, B. Lashermes and S. Jaffard, *Revisiting scaling, multifractal and multiplicative cascades with the wavelet leader lens,* Optic East, Wavelet applications in Industrial applications II Vol. 5607 pp. 103-117 Philadelphia, USA (2004)

[3] P. Abry, B. Lashermes and S. Jaffard, *Wavelet leader based multifractal analysis,* Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing.

[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] E. Bacry, J. Delour and J.F. Muzy *Multifractal random walk*, Phys. Rev. E, vol. 64, 026103-026106 (2001)

[6] G. Brown, G. Michon and J. Peyrière, *On the multifractal analysis of measures,* J. Statist. Phys., vol. 66 pp. 775-790 (1992).

[7] A. P. Calderòn and A. Zygmund, *Singular integral operators and differential equations,* Amer. J. Math. vol. 79 pp. 901–921 (1957).

[8] A. B. Chhabra, C. Meneveau, R. V. Jensen and K. Sreenivasan, *Direct determination of the $f(\alpha)$ singularity spectrum and its applications to fully developed turbulence,* Phys. Rev. A, vol. 40 pp. 5284–5294, (1989)

[9] M. Clausel, *Etude de quelques notions d'irrégularité : le point de vue ondelettes,* Ph.D. Thesis, Université Paris 12 (2008)

[10] A. Cohen, I. Daubechies and J.-C. Fauveau, *Biorthogonal bases of compactly supported wavelets,* Comm. Pure Appl. Math., vol. 44 pp. 485–560 (1992).

[11] A. Cohen and R. Ryan, *Wavelets and Multiscale Signal Processing,* Chapman and Hall (1995)

[12] J. P. R. Christensen, *On sets of Haar measure zero in Abelian Polish groups*, Israel J. Math. **13** (1972), 255–260.

[13] A. Fraysse and S. Jaffard *How smooth is almost every function in a Sobolev space?* Rev. Matem. Iberoamer. Vol. 22, N. 2 pp. 663–682 (2006).

[14] 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)

[15] Y. Heurteaux and S. Jaffard *Multifractal analysis of images: New connexions between analysis and geometry,* Proceedings of the NATO-ASI Conference on Imaging for Detection and Identification, Springer (2007).

[16] B. Hunt, *The prevalence of continuous nowhere differentiable functions.*, Proceed. A.M.S **122** (1994), no. 3, 711–717.

[17] B. Hunt and T. Sauer and J. Yorke, *Prevalence: A translation invariant "almost ever" on infinite dimensional spaces*, Bull. A.M.S **27** (1992), 217–238.

[18] 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).

[19] S. Jaffard, *Multifractal formalism for functions,* SIAM J. Math. Anal., vol. 28 pp. 944–998 (1997).

[20] S. JAFFARD, *Wavelet techniques in multifractal analysis,* Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus et M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics, AMS, Vol. 72 Part 2 pp. 91–152 (2004).

[21] S. JAFFARD, *Wavelet Techniques for pointwise regularity,* Ann. Fac. Sci. Toul., Vol. 15 n. 1 pp. 3–33 (2006).

[22] S. JAFFARD, *Pointwise regularity associated with function spaces and multifractal analysis,* Banach Center Pub. Vol. 72 Approximation and Probability, T. Figiel and A. Kamont Eds. pp. 93–110 (2006)

[23] S. Jaffard, *Hölder width and directional smoothness of nonharmonic Fourier series,* preprint (2007)

[24] 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. 513-539 (2005).

[25] S. Jaffard and Y. Meyer, *Wavelet methods for pointwise regularity and local oscillations of functions,* Mem. Amer. Math. Soc., vol. 123 No. 587 (1996)

[26] S. Jaffard, Y. Meyer and R. Ryan, *Wavelets: Tools for Science and Technology,* SIAM, (2001)

[27] B. Lashermes, S. Roux et P. Abry and S. Jaffard   *Comprehensive multifractal analysis of turbulent velocity using wavelet leaders,* preprint (2007).

[28] J. Lévy-Véhel and S. Seuret,   *The local Hölder function of a continuous function,* Preprint (2001).

[29] S. Mallat,   *A Wavelet Tour of Signal Processing,* Academic Press (1998)

[30] 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).

[31] B. Mandelbrot,   *The Fractal Geometry of Nature,* W. H. Freeman (1982).

[32] B. Mandelbrot, *A multifractal walk down Wall Street,* Scienttific American, vol. 280, pp. 70–73, (1999).

[33] P. Mattila,   *Geometry of sets and measures in Euclidean Spaces,* Cambridge Univ. Press (1995).

[34] Y. Meyer,   *Ondelettes et Opérateurs,* Hermann (1990).

[35] S. Orey and S. J. Taylor   *How often on a Brownian path does the law of iterated logarithm fail?* Proc. London Math. Soc. (3) 28, 174–192 , 174–192. .

[36] K. I. Oskolkov   *The Schrödinger density and the Talbot effect,* Approximation and Probability, Banach Center Pub. Vol. 72 pp. 19–219 (2006)

[37] 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).

[38] I. M. Singer,   *Bases in Banach spaces, Vol. 1,* Springer (1970).

[39] C. Tricot,   *Two definitions of fractional dimension,* Math. Proc. Cambridge Philos. Soc, 91 (1) pp. 57-74 (1982)

[40] C. Tricot,   *Function norms and franctal dimensions,* SIAM J. Math. Anal. Vol 28 n. 1 pp. 189-212 (1997)

[41] H. Wendt, P. Abry and S. Jaffard, *Bootstrap for Emperical Multifractal Analysis* IEEE Signal Proc. Mag., vol. 24, no. 4, pp. 38-48 (2007)

[42] H. Wendt, S. Roux P. Abry and S. Jaffard,   *Bootstrapped wavelet leaders for multifractal analysis of images* preprint, (2008).