Function spaces vs. Scaling functions: Some issues in image classification

Stéphane Jaffard, Patrice Abry, Stéphane Roux

Function spaces vs. scaling functions: Tools for image classification

S. Jaffard, P. Abry and S. Roux

S. Jaffard
Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050 du CNRS, Université Paris Est Créteil Val de Marne, 61 Avenue du Général de Gaulle, 94010 Créteil Cedex, France. e-mail: jaffard@u-pec.fr

P. Abry
CNRS UMR 5672 Laboratoire de Physique, ENS de Lyon, 46, allée d'Italie, F-69364 Lyon cedex, France e-mail: patrice.abry@ens-lyon.fr

S. Roux
CNRS UMR 5672 Laboratoire de Physique, ENS de Lyon, 46, allée d'Italie, F-69364 Lyon cedex, France e-mail: stephane.roux@ens-lyon.fr

Contents

Abstract We investigate the properties of several classes of new parameters issued from multifractal analysis and used in image analysis and classification. They share the following common characteristics: They are derived from local quantities based on wavelet coefficients; at each scale, l^p averages of these local quantities are performed and exponents are deduced form a regression through the scales on a log-log plot. This yields scaling functions, which depend on the parameter p, and are used

for model selection and classification. We expose possible variants, and their pros and cons. We relate the values taken by these scaling functions with the determination of the regularity of the image in some classes of function spaces, and we show that looking for richer criteria naturally leads to the introduction of new classes of function spaces. We will show which type of additional information this information yields for the initial image.

1 Introduction

Tools supplied by fractal geometry have been widely used in image processing in order to derive parameters of fractal nature, which can be used for classification, model selection and parameter fitting. Fractal objects often present two related aspects: One is analytic, and consists in scale invariance properties, and the other is geometric and is expressed by a fractional dimension: For instance, the one-dimensional Brownian motion is scale invariant: $B(ax)$ has the same law as $a^{1/2}B(x)$, and its sample paths have fractional dimension $3/2$.

Similarly, the *Weierstrass-Mandelbrot functions*

$$W_{a,b}(x) = \sum_{n=-\infty}^{+\infty} \frac{\sin(a^n x)}{b^n} \qquad (0 < 1/b < a < 1)$$

are selfsimilar of exponent $\alpha = -\log a / \log b$ since $W_{a,b}(bx) = b^{\alpha} W_{a,b}(x)$, and the box dimensions of their graphs (see definition below) is $2 - \alpha$.

Let us consider the geometric aspect, which is supplied by fractal dimensions. The simplest notion of dimension which can be used is the *box dimension*:

Definition 1 *Let A be a bounded subset of $\mathbb{R}^d$; if $\varepsilon > 0$, let $N_\varepsilon(A)$ denote the smallest number of balls of radius ε required to cover A.*

The upper and lower box dimension of A are respectively given by

$$\overline{\dim}_B(A) = \limsup_{\varepsilon \to 0} \frac{\log N_\varepsilon(A)}{-\log \varepsilon}, \quad \textit{and} \quad \underline{\dim}_B(A) = \liminf_{\varepsilon \to 0} \frac{\log N_\varepsilon(A)}{-\log \varepsilon}.$$

The following important feature makes this notion useful in practical applications: If both limits coincide, then the box dimension can be computed through a regression on a log-log plot ($\log N_\varepsilon(A)$ vs. $\log \varepsilon$):

$$\dim_B(A) = \lim_{\varepsilon \to 0} \frac{\log N_\varepsilon(A)}{-\log \varepsilon}. \tag{1}$$

As such, this tool has a rather narrow field of applications in image processing; indeed, it applies only when a particular "fractal" set can be isolated in the image, with a sufficient resolution. However, the fact that the limit exists in (1) points towards another possible feature, which is much more common: Some quantities display an approximate power-law behavior through the scales. Such a property may hold for many other quantities than $N_\varepsilon(A)$ and the associated power-law exponents can thus yield parameters which can be used in image classification; this points towards the "analytic aspect" that we mentioned. The study of such quantities is usually referred to as *multifractal analysis*; the purpose of this field is:

- To introduce new quantities which present such power-law behaviors,
- to study their mathematical properties, and relate them with scales of function spaces,
- to determine geometric implications, as regards the presence of "fractal features" in the image.

This paper is divided as follows: In Section 2 we review the different quantities which are related with function spaces, and have been used for signal or image classification. We explain their relationships with function space indicators, and in particular with Sobolev and Besov spaces. Section 3 deals with new classes of parameters, which are built using quantities related with different types of pointwise regularity criteria: The local L^q behavior of the function (Section 3.1), the regularization performed by fractional integration (Section 3.2), or a mixture of both approaches (Section 3.3). Finally Section 4 deals with some general results concerning grand canonical formalisms, which are required when one takes into account simultaneously two different pointwise criteria. We will develop the case where one of the exponents is a regularity index, and the other measures the local oscillations of f near a point, see Section 4.1.

2 Scaling functions and function spaces

Let us now briefly review the different quantities which have been used up to now as possible candidates for scale invariance features. First, these quantities usually depend on (at least) one auxiliary parameter p (see e.g. (2) below), and therefore the exponents which are derived are not given by one real number (or a few), but are functions of this parameter p, hence the term *scaling functions* used in order to characterize these collections of exponents. Note that the use of a whole function in order to perform classification yields a potentially much richer tool than the use of one single number.

Let us now be more specific and start with what was historically the first example of a scaling function. It was introduced by N. Kolmogorov in the context of fully developed turbulence, with a motivation which was quite similar to ours, see [43]: Take advantage of the (expected) scale invariance of fully developed turbulence in order to derive a collection of "universal" parameters which could be computed on experimental data, and used for the validation of turbulence models.

Let $f : \mathbb{R}^d \to \mathbb{R}$. The *Kolmogorov scaling function* of f is the function $\eta_f(p)$ which satisfies

$$\int |f(x+h)-f(x)|^p dx \quad \sim \quad |h|^{\eta_f(p)}. \tag{2}$$

This is mathematically formalized by

$$\eta_f(p) = \liminf_{|h|\to 0} \; -\frac{1}{\log|h|} \log\left(\int |f(x+h)-f(x)|^p dx \right). \tag{3}$$

As in the case of the box dimension, we have to draw a distinction between the mathematical definition, whose purpose is to make sense in a general setting (e.g. whenever f is locally bounded), and the numerical evaluation of $\eta_f(p)$ which requires that the liminf is a real limit and, in practice, that one can make a precise regression on the scales available in the data. Note that the mathematical hypotheses bearing on f such that such scaling invariance holds are far from being well understood.

This first scaling function has a function space interpretation which will serve several purposes. First, it allows to derive several of its mathematical properties, but its main advantage will be to point the way towards variants and extensions of this scaling function which yield sharp information on the singularities present in the signal. This last motivation had unexpected consequences: For instance such new scaling functions allow to show the presence of "oscillating singularities" in the data, which was an important open issue in several applications, see [5, 30].

The most straightforward function space interpretation of the scaling function is obtained through the use of the spaces $\mathrm{Lip}(s, L^p)$ defined as follows.

Definition 2 *Let $s \in (0,1)$, and $p \in [1,\infty]$; $f \in Lip(s, L^p(\mathbb{R}^d))$ if*

$$f \in L^p \quad and \quad \exists C > 0,\ \forall h > 0, \quad \int |f(x+h)-f(x)|^p dx \le C|h|^{sp}. \tag{4}$$

It follows from (2) and this definition that, if $\eta_f(p) < p$, then

$$\eta_f(p) = \sup\{s : f \in \mathrm{Lip}(s/p, L^p(\mathbb{R}^d))\}. \tag{5}$$

In other words, the scaling function allows to determine which spaces $\mathrm{Lip}(s, L^p)$ contain the signal for $s \in (0,1)$, and $p \in [1,\infty]$. This reformulation has several advantages:

- Using classical embeddings between function spaces, one can derive alternative formulations of the scaling function which, though they are mathematically equivalent to (2) or (5), allow a better numerical implementation,
- It has extensions outside of the range $s \in (0,1)$, and therefore applies to signals which are either smoother or rougher than allowed by this range,
- extensions of these function spaces to $p < 1$ lead to a scaling function defined for $p < 1$, and therefore supply a richer tool for classification.

The simplest setting for these extension is supplied by Besov spaces; they offer the additional advantage of yielding a wavelet reformulation of the scaling function, which is simple and robust, see [1, 3, 4, 34]. In order to define these spaces, we need to recall the notion of wavelet bases.

2.1 Orthonormal and biorthogonal wavelet bases

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 often be characterized by conditions bearing on the wavelet coefficients, see Section 2.2. We will only recall the properties of orthonormal and biorthogonal wavelet bases that will be useful in the sequel. We refer the reader for instance to [15, 16, 36, 37] for detailed expositions.

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 $2^{dj/2}\psi^{(i)}(2^jx-k)$ ($k\in\mathbb{Z}^d$, $j\in\mathbb{Z}$) form orthonormal wavelet 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 $\partial^\alpha\psi^{(i)}$, for $|\alpha|\le r$, have fast decay.

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

$$f(x)=\sum_{k\in\mathbb{Z}^d}c_k^{(0)}\varphi(x-k)+\sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}^d}\sum_i c^i_{j,k}\psi^{(i)}(2^jx-k); \tag{6}$$

the $c^i_{j,k}$ and $c_k^{(0)}$ are called the wavelet coefficients of f and given by

$$c^i_{j,k}=2^{dj}\int_{\mathbb{R}^d}f(x)\psi^{(i)}(2^jx-k)dx, \text{ and } c_k^{(0)}=\int_{\mathbb{R}^d}f(x)\varphi(x-k)dx.$$

Note that the computation of these coefficients makes sense with very little assumption on f (a wide mathematical setting is supplied by tempered distributions). A natural setting for functions is supplied by the space L^1 *with slow growth* which is defined as follows.

Definition 3 *Let f be a locally integrable function defined over $\mathbb{R}^d$; f belongs to $L^1_{SG}(\mathbb{R}^d)$ if*

$$\exists C,N>0 \quad \textit{such that} \qquad \int_{\mathbb{R}^d}|f(x)|(1+|x|)^{-N}dx\le C.$$

The wavelet expansion of a function $f\in L^1_{SG}(\mathbb{R}^d)$ converges a.e.; in particular at Lebesgue points, it converges towards the Lebesgue value

$$\lim_{r\to 0}\frac{1}{Vol(B(x_0,r))}\int_{B(x_0,r)}f(x)dx.$$

Furthermore, let $C_{SG}(\mathbb{R}^d)$ be the set of locally bounded and continuous functions which satisfy

$$\exists C, N > 0: \qquad |f(x)| \leq C(1+|x|)^N.$$

Then, if $f \in C_{SG}(\mathbb{R}^d)$, then its wavelet expansion converges uniformly on compact sets.

We will also need decompositions on *biorthogonal wavelet bases*, which are a useful extension of orthonormal wavelet bases. Recall that a *Riesz basis* of a separable Hilbert space H is a collection of vectors (e_n) such that the finite linear expansions $\sum_{n=1}^N a_n e_n$ are dense in H and

$$\exists C, C' > 0: \ \forall N, \quad \forall a_n, \quad C \sum_{n=1}^N |a_n|^2 \leq \left\| \sum_{n=1}^N a_n e_n \right\|_H^2 \leq C' \sum_{n=1}^N |a_n|^2. \tag{7}$$

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 this is the case, any element $f \in H$ can be written $f = \sum \langle f | f_n \rangle e_n$ (where the series converges in H). Biorthogonal wavelet bases are couples of Riesz bases of L^2 which are of the following form: On one side,

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

and, on the other side,

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

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

$$f(x) = \sum_{k \in \mathbb{Z}^d} c_k^{(0)} \varphi(x-k) + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}^d} \sum_i c_{j,k}^i \psi^{(i)}(2^j x - k); \tag{8}$$

where

$$c_{j,k}^i = 2^{dj} \int_{\mathbb{R}^d} f(x) \tilde{\psi}^{(i)}(2^j x - k) dx, \ \text{ and } \ c_k^{(0)} = \int_{\mathbb{R}^d} f(x) \tilde{\varphi}(x-k) dx. \tag{9}$$

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 $\{0,1\}^d - (0, \dots, 0)$; we introduce

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

and, accordingly: $c_\lambda = c_{j,k}^i$ and $\psi_\lambda(x) = \psi^{(i)}(2^j x - k)$. Indexing by dyadic cubes will be useful in the sequel because the cube λ indicates the localization of the corresponding wavelet. Note that this indexing is injective: if $(i, j, k) \neq (i', j', k')$, then $\lambda(i, j, k) \neq \lambda(i', j', k')$. Dyadic cubes have the remarkable property that two of them are either disjoint, or one is included in the other.

The wavelet ψ_λ is essentially localized near the cube λ; more precisely, when the wavelets are compactly supported, then, $\exists C > 0$ such that when $\psi^{(i)} \subset [-C/2, C/2]^d$ then $\psi_\lambda \subset 2^{-j}k + 2^{-j}[-C/2, C/2] \subset 2C\lambda$.

Remarks: Two classes of orthonormal wavelet bases play a prominent role: For theoretical purposes, the wavelets introduced by Y. Meyer and P.-G. Lemarié, which belong to the Schwartz class (but are not compactly supported), and Daubechies wavelets which are compactly supported, but of limited regularity (the size of the support essentially is a linear function of the regularity of the wavelet), see [36, 37, 17].

In order to have a common notation for wavelets and functions φ, when $j = 0$, we note ψ_λ the function $\varphi(x-k)$ (where λ is, in this case, the unit cube shifted by k), and by c_λ the corresponding coefficient.

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}), and Λ will denote the union of the Λ_j for $j \geq 0$.

2.2 *The wavelet scaling function*

A remarkable property of wavelet bases is that they supply bases for most function spaces that are used in analysis. The case of Besov spaces is typical:

Definition 4 *Let $s \in \mathbb{R}$ and $p, q \in (0, \infty]$. Let ψ_λ be an r-smooth wavelet basis with $r > \sup(|s|, |s| + d(\frac{1}{p} - 1))$. A distribution f belongs to the Besov space $B_p^{s,q}(\mathbb{R}^d)$ if and only if its wavelet coefficients satisfy*

$$\sum_{j \geq 0} \left(\sum_{\lambda \in \Lambda_j} \left[2^{(s-d/p)j} |c_\lambda| \right]^p \right)^{q/p} < \infty \tag{10}$$

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

Remarks: Historically, this was not the first definition of Besov spaces, since they were initially introduced as interpolation spaces between Sobolev spaces. However, this very simple characterization has opened the way to an extensive use of Besov spaces in many fields, including image processing and statistics, so that it is now rather taken as a definition, see [37]. Informally, it essentially means that f and its fractional derivative of order s belong to L^p (and, indeed, $B_2^{s,2}$ coincides with the Sobolev space H^s), see [37] for precise embeddings with Sobolev spaces. Let $s \in (0,1)$, and $p \in [1, \infty]$; then the following embeddings hold:

$$B_p^{s,1} \hookrightarrow \mathrm{Lip}(s, L^p(\mathbb{R}^d)) \hookrightarrow B_p^{s,\infty}. \tag{11}$$

Furthermore, if $s \geq 0$, $p > 0$ and $0 < q_1 < q_2$, then we have the embeddings

$$\forall \varepsilon > 0, \;\; B_p^{s+\varepsilon,\infty} \hookrightarrow B_p^{s,q_1} \hookrightarrow B_p^{s,q_2} \hookrightarrow B_p^{s,\infty}, \tag{12}$$

which allow to redefine Kolmogorov's scaling function for $p \geq 1$ by,

$$\eta_f(p) = \sup\{s : f \in B_p^{s/p,\infty}\}. \tag{13}$$

There are two advantages in doing so: on one hand, it extends the scaling function to all values of $p > 0$, on the other hand, it suggests an alternative way to compute it, through a regression based on wavelet coefficients. Indeed, it follows from (10) and (13) that the scaling function of f is

$$\forall p > 0, \qquad \eta_f(p) = \liminf_{j\to+\infty} \; -\frac{1}{j} \log_2 \left(2^{-dj} \sum_{\lambda\in\Lambda_j} |c_\lambda|^p \right). \tag{14}$$

We will use the same notation for the Kolmogorov scaling function and the wavelet scaling function since they coincide for a range of values; this slight abuse of notations will lead to no confusion. From now on, we take (14) for definition of the *wavelet scaling function*, and also, as a practical way to compute it through a regression on a log-log plot. Note that its interpretation in terms of function spaces implies that it is independent of the (smooth enough) wavelet basis. A similar formula was previously introduced by A. Arneodo, E. Bacry and J.-F. Muzy, using the continuous wavelet transform, see for instance [9]. An additional advantage of using (14) as a definition is that the scaling function is well defined even if f is not a function (in the most general case, it can be a tempered distribution); a large degree of flexibility may prove necessary: Indeed, a picture is a discretization of the light intensity, which is a non-negative quantity. Therefore the most general mathematical modeling which takes into account only this a priori information amounts to make the assumption that f is a measure (indeed non-negative distributions necessarily are measures, by a famous theorem of L. Schwartz); a posteriori estimation of the smoothness of images using the wavelet scaling function showed that, indeed, some types of natural images are not smoother than measures are, see [2, 48].

2.3 The uniform Hölder exponent

We consider now the particular case of Besov spaces with $p = q = +\infty$, which will play an important role in the following: f belongs to $B_\infty^{s,\infty}(\mathbb{R}^d)$ if and only if $(c_k^{(0)}) \in l^\infty$ and

$$\exists C, \, \forall \lambda \in \Lambda, \qquad |c_\lambda| \; \leq C 2^{-sj}. \tag{15}$$

The spaces $B_\infty^{s,\infty}$ coincide with the Lipschitz spaces $C^s(\mathbb{R}^d)$ when $s \notin \mathbb{N}$ (for instance, if $s = 1$, then (15) characterizes function in the Zygmund class, see [37]); it follows that, if $0 < s < 1$, $f \in B_\infty^{s,\infty}(\mathbb{R}^d)$ if f is bounded and $\exists C, \;\; \forall x,y, \; |f(x) - f(y)| \leq C|x-y|^s$.

The *uniform Hölder exponent* of f is

$$H_f^{min} = \sup\{s : \ f \in C^s(\mathbb{R}^d)\}. \tag{16}$$

It yields an additional parameter for image processing and classification that will prove important for additional reasons: Its determination will be a mandatory step for multifractal analysis; note that H_f^{min} is related with the scaling function according to $H_f^{min} = \lim_{p\to+\infty} \eta'_f(p)$, see [26]. In practice, it is derived directly from the wavelet coefficients of f through a simple regression on a log-log plot ; indeed, it follows from (16) and the wavelet characterization of the Besov spaces $B_\infty^{s,\infty}$ that

$$H_f^{min} = \liminf_{j\to+\infty} \ -\frac{1}{j}\log_2\left(\sup_{\lambda\in\Lambda_j}|c_\lambda|\right). \tag{17}$$

Note that the determination of H_f^{min} does not require any a priori assumption on f and that, in practice, it allows to settle if f is bounded or not. Indeed, it follows from (16) that

- if $H_f^{min} > 0$, then f is locally bounded,
- if $H_f^{min} < 0$, then $f \notin L^\infty_{loc}$.

We will meet several situations where this a priori information is needed.

Let us consider now two simple models of random fields. The first one will show no bias in the estimation of H_f^{min}, and the other will show one.

The first model is supplied by Lacunary Wavelet Series (LWS) $X_{\alpha,\gamma}$ of type (α,γ), see [24]; they are random fields defined on $[0,1]^d$ (for $\alpha > 0$ and $\gamma < d$) as follows: A biorthogonal wavelet basis in the Schwartz class is used for the construction. One draws at random (uniformly) $2^{\gamma j}$ locations λ among the 2^{dj} dyadic cubes of width 2^{-j} included in $[0,1]^d$, and the corresponding wavelet coefficients are set to the value $2^{-\alpha j}$, whereas the others are set to 0. In order to define LWS on the whole $\mathbb{R}^d$, one repeats this construction on all cubes of width 1. In this case it follows from (17) that $H_f^{min} = \alpha$ and (17) yields α exactly at each scale.

The second example falls in the general model of *Random Wavelet Series*. Let us describe the particular case that we consider now.

Definition 5 *Let ψ_λ be a biorthogonal wavelet basis in the Schwartz class. A Uniform Random Wavelet Series (URWS) of type (α,β) is a random field of the form*

$$X = \sum 2^{-\alpha j} X_\lambda \psi_\lambda,$$

where the X_λ are IID with common law, which is a non-vanishing random variable satisfying a tail estimate of the form

$$\mathbb{P}(|X| \geq A) \sim C\exp(-B|A|^\beta),$$

and C, B and β are positive constants.

Note that this model includes the Fractional Brownian Motion (FBM) of exponent $\gamma \in (0,1)$. Indeed, if $s \in \mathbb{R}$, let $\hat{\psi}_s(\xi) = |\xi|^{-s}\hat{\psi}(\xi)$, where ψ is generating an orthonormal wavelet basis (ψ_s is the fractional integral of ψ of order s). If $\psi \in \mathscr{S}(\mathbb{R})$, then ψ_s is a wavelet and the $2^{j/2}\psi_s(2^j x - k)$ and the $2^{j/2}\psi_{-s}(2^j x - k)$ form biorthogonal bases, see e.g. [26]. If $t \in [0,1]$ then

$$B_\gamma(t) = \sum_{j=0}^{\infty} \sum_{k\in\mathbb{Z}} 2^{-\gamma j} \xi_{j,k}\ \psi_{\gamma+1/2}(2^j t - k) + R(t) \tag{18}$$

where $R(t)$ is a C^∞ random process, and the $\xi_{j,k}$ are I.I.D. standard centered Gaussians, see [6, 40].

If $\beta > 1$, then the model supplied by URWS yields a random field, with a constant Hölder exponent α, see [11]. Since the X_λ are independent, one obtains that

$$\mathbb{P}\left(\sup_{\lambda\in\Lambda_j} |X_\lambda| \le A\right) \sim 1 - C\exp(j\log 2 - BA^\alpha);$$

therefore, the quantity $\sup_{\lambda\in\Lambda_j} |X_\lambda|$ is asymptotically equivalent to $Cj^{1/\alpha}$. It follows that

$$-\frac{1}{j}\log_2\left(\sup_{\lambda\in\Lambda_j} |c_\lambda|\right) = \alpha - \frac{\log_2 j}{\alpha j}(1+o(1));$$

therefore the statistical estimator of H_f^{min} supplied by $-\log_2(\sup|c_\lambda|)/j$ is biased by a term equivalent to $(\log_2 j)/j$.

Note that, from the point of view of statistical estimation, the case of the determination of H^{min} is quite different from the determination of the scaling function. Indeed, for URWS, if $C_p = \mathbb{E}(|X|^p)$, then

$$\mathbb{E}\left(2^{-dj}\sum |C_\lambda|^p\right) = C_p 2^{-\alpha j} \ \text{ and } \ Var\left(2^{-dj}\sum |C_\lambda|^p\right) = 2^{-dj} C_{2p} 2^{-2\alpha j}.$$

It follows that no such bias exists for the estimation of the scaling function of URWS: It is at most C/j, i.e., smaller than for the uniform Hölder exponent by a logarithmic term.

2.4 Scaling functions for image model validation

We start by recalling a general problem posed by function-space modeling when applied to real-life signals: Data are always available with a finite resolution; therefore, assuming that images are functions (or perhaps distributions) continuously defined on $\mathbb{R}^2$ (or a subset of $\mathbb{R}^2$ such as a square or a rectangle) is an idealization which may be convenient for mathematical modeling, but should not mask the fact that real-life images are sampled and given by a finite array of numbers. Therefore, the problem of finding which function spaces contain a particular image is ill-posed: Indeed, given any "classical" space of functions defined on a square, and such an array of numbers, one can usually find a function in this space that will have the pre-assigned values at the corresponding points of the grid. For instance, a commonly met pitfall is that an image is given by grey-levels, and thus takes values in $[0,1]$. Therefore, it may seem appropriate to use a modeling by bounded functions. We will see that wavelet techniques allow to discuss this assumption, and show that it is not satisfied for large classes of natural textures.

The resolution of these problems requires the use of *multiscale techniques* as supplied by wavelet analysis. Let us consider for instance the last example we mentioned: Starting with a discrete image, given by an array of 1024×1024 numbers all lying between 0 and 1, how can we decide that it can be modeled or not by a bounded function? It is clear that, if we consider the image at only one scale (the finest scale in order to lose no information), then the answer seems to be affirmative. One way to solve the difficulty is to consider the image simultaneously at all the scales available and inspect if certain quantities behave through this range of scales as bounded functions do. The practical recipe in this case is to use (17) in order to determine numerically the value of H_f^{min}, through a regression on a log-log plot; if $H_f^{min} < 0$, then the image is not bounded, and if $H_f^{min} > 0$, then the image is bounded. Of course, if the numerical value obtained for H_f^{min} is close to 0 (i.e., if 0 is contained in the confidence interval which can be obtained using statistical methods such as the bootstrap, see [46, 48]) then the issue remains unsettled.

The wavelet scaling function yields an extensive information on the function spaces which contain a particular signal. As an example of its use, we will discuss the assumption that real-life images can be modeled as a sum of a function $u \in BV$ and another term v which will model noise and texture parts. There is no consensus on which regularity should be assumed for the second term v. The first "$u+v$ model" (introduced by Rudin, Osher and Fatemi [44]) assumed that $v \in L^2$; however, the very strong oscillations displayed by some textures suggested that such components do not have a small L^2 norm, but might have a small norm in spaces of negative regularity index (i.e., spaces of distributions). Therefore the use of spaces such as divergences of L^∞ functions (or divergences of L^2 functions) were proposed by Y. Meyer, see [39], followed by several other authors, see [12, 42] and references therein. More sophisticated models also aim at separating noise from texture, and therefore split the image into three components ($u+v+w$ models, see [12]).

The Rudin-Osher-Fatemi algorithm consists in obtaining the regular component u of an image f by minimizing the functional $J(u) = \| u \|_{BV} + t \| f - u \|_2^2$, where t is a scale parameter which has to be tuned. In 2001, Y. Meyer proposed to minimize the alternative functional $J(u) = \| u \|_{BV} + t \| f - u \|_G$, where $\| f \|_G = \inf_{g:\, f = \nabla \cdot g} \| g \|_\infty$. More recently, in 2003, Osher, Solé and Vese proposed another model which uses for the texture and noise component a space of distributions easier to handle, the Sobolev space H^{-1}. The corresponding functional is $J(u) = \| u \|_{BV} + t \| f - u \|_{H^{-1}}^2$.

More recently, several variants have been proposed, based on the same fundamental ideas, but using alternative function spaces. The relevance of one particular function space is usually advocated using either theoretical arguments derived from functional analysis, or practical arguments motivated by the algorithmic implementation. The fundamental problem of determining to which function spaces a given image (or a part of a given image) belongs to has been rarely considered (see however [20] where the authors question the fact that natural images belong to BV). A first motivation for this question rises implicitly from the short review we just performed: The function spaces used in modeling should fit the data. Another motivation is that, if these function spaces depend strongly on the image that is considered, then this information might prove useful in image classification. The knowledge of the scaling function allows to settle the issues we raised concerning the function spaces which contain a given image. The following result is a direct consequence of the embeddings (11) and (12) between Besov spaces and other classical function spaces, and allows to settle, for a given image, if the models mentioned above fit the data.

Corollary 1 *Let f be a tempered distribution defined on $\mathbb{R}^2$. Then, the values taken by $\eta_f(p)$ and by H_f^{min} have the following interpretation:*

- *If $\eta_f(1) > 1$, then $f \in BV$, and if $\eta_f(1) < 1$, then $f \notin BV$*
- *If f is a measure, then $\eta_f(1) \geq 0$, and, if $\eta_f(1) > 0$, then f belongs to L^1.*
- *If $\eta_f(2) > 0$, then $f \in L^2$ and if $\eta_f(2) < 0$, then $f \notin L^2$.*
- *If $\eta_f(2) > -2$, then $f \in H^{-1}$ and if $\eta_f(2) < -2$, then $f \notin H^{-1}$.*
- *If $H_f^{min} > 0$, then f is bounded and continuous, and if $H_f^{min} < 0$, then $f \notin L^\infty$.*
- *If $H_f^{min} > -1$, then $f \in G$ and if $H_f^{min} < -1$, then $f \notin G$.*
- *If f is a measure, then $H_f^{min} \geq -2$.*

2.5 Multifractal formalism

Later refinements and extensions of the wavelet scaling function were an indirect consequence of its interpretation in terms of fractal dimensions of Hölder singularities, proposed by G. Parisi and U. Frisch in their seminal paper [43]. In order to explain their argumentation, we first recall the definition associated with pointwise regularity. The most widely used one is supplied by *Hölder regularity*.

Definition 6 *Let $f:\ \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function, $x_0 \in \mathbb{R}^d$ and let $\alpha \geq 0$; f belongs to $C^\alpha(x_0)$ if there exist $C > 0$, $R > 0$ and a polynomial P of degree less than α such that, if $|x - x_0| \leq R$, then $|f(x) - P(x - x_0)| \leq C|x - x_0|^\alpha$.*

The Hölder exponent of f at x_0 is: $h_f(x_0) = \sup\{\alpha :\ f \text{ is } C^\alpha(x_0)\}$.

The isohölder sets are: $E_H = \{x_0 :\ h_f(x_0) = H\}$.

Note that Hölder exponents met in signal processing often lie between 0 and 1, in which case the Taylor polynomial $P(x - x_0)$ boils down to $f(x_0)$ and the definition of the Hölder exponent means that, heuristically, $|f(x)-f(x_0)| \sim |x-x_0|^{h_f(x_0)}$. The loose idea which lies at the starting point of the derivation proposed in [43] is that, if f is not smooth on a large set of locations, then, at a given scale h, the increments $f(x+h)-f(x)$ will bring a large contribution to (2), and therefore the knowledge of the scaling function should yield some information of the size of the sets where f has a given Hölder regularity. A statistical physics argument leads to a precise statement, usually referred to as the *multifractal formalism* concerning the size of the sets of singularities of f. In order to recall it, we start by giving the notion of "size" which is adapted to this problem, see [18].

Definition 7 *Let $A \subset \mathbb{R}^d$. If $\varepsilon > 0$ and $\delta \in [0,d]$, we denote*

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

where R is an ε-covering of A, i.e., a covering of A by bounded sets $\{A_i\}_{i\in\mathbb{N}}$ of diameters $|A_i| \leq \varepsilon$. The infimum is therefore taken on all ε-coverings.

For any $\delta \in [0,d]$, the δ-dimensional Hausdorff measure of A is $mes_\delta(A) = \lim_{\varepsilon\to 0} M_\varepsilon^\delta$. There exists $\delta_0 \in [0,d]$ such that $\forall \delta < \delta_0$, $mes_\delta(A) = +\infty$ and $\forall \delta > \delta_0$, $mes_\delta(A) = 0$. This critical δ_0 is called the Hausdorff dimension of A, and is denoted by $\dim(A)$ (by convention, $\dim(\emptyset) = 0$).

If f is locally bounded, then the function $H \to \dim(E_H)$ is called the *spectrum of singularities* of f. Note that, in distinction with the box dimension, the Hausdorff dimension cannot be computed via a regression on a log-log plot; therefore it can be estimated on experimental data only through indirect methods. We will see that the use of the multifractal formalism is one of them.

The formula proposed by Paris and Frisch is

$$\dim(E_H) = \inf_{p\in\mathbb{R}} \left(d + Hp - \eta_f(p)\right), \tag{19}$$

see [43]. However, it meets severe limitations: Many natural processes used in signal or image modeling, are counterexamples, see [1]; additionally, the only result relating the spectrum of singularities and the scaling function in all generality is partial, and stated below, see [23, 26].

Theorem 1 *Let $f:\ \mathbb{R}^d \to \mathbb{R}$ be such that $H_f^{min} > 0$. Define p_0 by the condition:*

$$\eta_f(p_0) = dp_0,$$

then

$$\dim(E_H) \leq \inf_{p>p_0} \left(d + Hp - \eta_f(p)\right). \tag{20}$$

2.6 Wavelet leaders

A natural line of research is to look for an "improved" scaling function, i.e., one such that (19) would have a wider range of validity, and for which the upper bound supplied by Theorem 1 would be sharper. This led to the construction of the *wavelet leader scaling function*, which we now recall. The "basic ingredients" in this formula are no more wavelet coefficients, but wavelet leaders, i.e., local suprema of wavelet coefficients. The reason is that pointwise smoothness can be expressed much more simply in terms of wavelet leaders than of wavelet coefficients.

Let λ be a dyadic cube; 3λ will denote the cube of same center and three times wider. If f is a bounded function, the *wavelet leaders* of f are the quantities

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

It is important to require f to be bounded; otherwise, the wavelet leaders of f can be infinite; therefore checking that $H_f^{min} > 0$ is a prerequisite of the method.

Remark: Similar quantities were previously introduced by S. Mallat, in the *Wavelet Transform maxima method*, which can be seen as counterpart of the wavelet leader technique for the continuous wavelet transform; it was used in multifractal analysis by A. Arneodo, E. Bacry and J.-F. Muzy:, see [9, 36] and references therein; however it does not enjoy the same simplicity of implementation and very few mathematical results apply to it.

The reason for introducing wavelet leaders is that they give an information on the pointwise Hölder regularity of the function. Indeed, let $x_0 \in \mathbb{R}^d$, and denote by $\lambda_j(x_0)$ is the dyadic cube of width 2^{-j} which contains x_0. If $H_f^{min} > 0$, then

$$h_f(x_0) = \liminf_{j \to +\infty} \; -\frac{1}{j} \; \log_2 \left(d_{\lambda_j(x_0)} \right) \tag{21}$$

(see [26] and references therein). Therefore, constructing a scaling function with the help of wavelet leaders is a way to incorporate pointwise smoothness information. It is therefore natural to expect that (20) will be improved when using such a scaling function instead of $\eta_f(p)$.

The *leader scaling function* is defined by

$$\forall p \in \mathbb{R}, \qquad \zeta_f(p) = \liminf_{j \to +\infty} \; -\frac{1}{j} \log_2 \left(2^{-dj} \sum_{\lambda \in \Lambda_j} |d_\lambda|^p \right). \tag{22}$$

An important property of the leader scaling function is that it is "well defined" for $p < 0$, although it can no more be subject to a function space interpretation. By "well defined", we mean that it has the following robustness properties if the wavelets belong to the Schwartz class (partial results still hold otherwise), see [31, 26]:

- ζ_f is independent of the wavelet basis.
- ζ_f is invariant under the addition of a C^∞ perturbation.

- ζ_f is invariant under a C^∞ change of variable.

Note that the wavelet scaling function does not possess these properties when p is negative. The leader scaling function can also be given a function-space interpretation for $p > 0$: Let $p \in (0,\infty)$; a function f belongs to the *Oscillation space* $\mathcal{O}_p^s(\mathbb{R}^d)$ if and only if

$$\forall j \geq 0, \qquad \sum_{\lambda \in \Lambda_j} \left[2^{(s-d/p)j} d_\lambda \right]^p < \infty.$$

Then

$$\zeta_f(p) = \sup\{s : f \in \mathcal{O}_p^{s/p}\}.$$

Properties of oscillation spaces are investigated in [31, 26]. Note that their definition mixes two quantities of different natures: a local l^∞ norm (the leader, which is a supremum of wavelet coefficients), and a global l^p norm. Thus they are reminiscent of *amalgam spaces*, which were introduced by N. Wiener, and thoroughly studied by H. Feichtinger, see [19]: These spaces are defined by taking local L^p norms of the function, and then averaging them through a discrete l^q sum. Note however the strong difference with oscillation spaces, where the l^∞ local norms are taken at each scale. Thus oscillation spaces do not favor a particular scale, in contradistinction with amalgam spaces, where a specific scale for the amalgam is chosen.

Let us now prove the following result, which shows that the wavelet leader scaling function can be alternatively defined through the "restricted leaders"

$$e_\lambda = \sup_{\lambda' \subset \lambda} |c_{\lambda'}|.$$

Lemma 1 *Let f be a function satisfying $H_f^{min} > 0$; then*

$$\forall p \in \mathbb{R}, \qquad \zeta_f(p) = \liminf_{j \to +\infty} \; -\frac{1}{j} \log_2 \left(2^{-dj} \sum_{\lambda \in \Lambda_j} |e_\lambda|^p \right).$$

Remark: This equivalence is important because the formulation of the scaling function using "extended leaders" (where the supremum is taken on 3λ) is required in order to prove upper bounds for spectra, see [26]; on the other hand, the formulation using "restricted leaders" is more suited to numerical implementation, and in order to prove properties of the scaling function (since the supremums at a given scale are taken on non-overlapping cubes).

Proof. Let

$$S_f(p,j) = 2^{-dj} \sum_{\lambda \in \Lambda_j} |c_\lambda|^p \quad \text{and} \quad T_f(p,j) = 2^{-dj} \sum_{\lambda \in \Lambda_j} |e_\lambda|^p.$$

Since

$$\sup_{\lambda' \subset \lambda} |c_{\lambda'}| \le \sup_{\lambda' \subset 3\lambda} |c_{\lambda'}|$$

it follows that, if $p \ge 0$, then $S_f(p,j) \le T_f(p,j)$, and, if $p \le 0$, then $T_f(p,j) \le S_f(p,j)$.

On the other hand, denote by μ the "father" of the cube λ (i.e., the cube twice as wide which contains λ), and denote by $N(\mu)$ the 3^d "neighbours" of λ (i.e., the cubes of same width, whose boundary intersects the boundary of λ). Then

$$\sup_{\lambda' \subset 3\lambda} |c_{\lambda'}| \le \sup_{\nu \in N(\mu)} \sup_{\lambda' \subset \nu} |c_{\lambda'}|.$$

It follows that, if $p \ge 0$, then $T_f(p,j) \le 3^d S_f(p,j-1)$.

Finally, for any dyadic cube λ, there exists a "grandson" λ'' of λ such that $3\lambda'' \subset \lambda$. Therefore

$$\sup_{\lambda' \subset 3\lambda''} |c_{\lambda'}| \le \sup_{\lambda' \subset \lambda} |c_{\lambda'}|;$$

therefore, if $p \le 0$, then $S_f(p,j) \le T_f(p,j+2)$. The result follows from these four estimates. □

The *leader spectrum* of f is defined through a Legendre transform of the leader scaling function as follows

$$L_f(H) = \inf_{p \in \mathbb{R}} \left(d + Hp - \zeta_f(p) \right). \tag{23}$$

The following result of [26] shows the improvement obtained when using wavelet leaders: The upper bound is sharpened since, on one hand $\eta_f(p) = \zeta_f(p)$ if $p > p_0$, and, on the other hand, the infimum in (23) is taken for all $p \in \mathbb{R}$.

Theorem 2 *If $H_f^{min} > 0$, then,* $\forall H, \quad \dim(E_H) \le L_f(H)$.

Furthermore, equality holds for large classes of models used in signal and image processing, such as Fractional Brownian Motions, lacunary and random wavelet series, cascade models,.... see [1, 31] and references therein.

2.7 Scaling function for measures

Since nonnegative measures supply a natural setting for the modeling of images, we now expose the tools related with the multifractal analysis of nonnegative measures (see e.g. [2, 13, 21]), and we will show that they can also be related with function space interpretations.

Definition 8 *Let $x_0 \in \mathbb{R}^d$ and let $\alpha \ge 0$. A nonnegative measure μ defined on $\mathbb{R}^d$ belongs to $C^\alpha(x_0)$ if there exists a constant $C > 0$ such that, in a neighbourhood of x_0, $\mu(B(x_0,r)) \le Cr^\alpha$, where $B(x_0,r)$ denotes the open ball of center x_0 and radius* *r. Let x_0 belong to the support of μ; then the Hölder exponent of μ at x_0 (also called the "local dimension") is*

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

Let $E_\mu(H) = \{x_0 : h_\mu(x_0) = H\}$. The *spectrum of singularities* of μ (denoted by $d_\mu(H)$) is the Hausdorff dimension of $E_\mu(H)$.

We will need to deduce the Hölder exponent at every point from a "discretized version" of μ, i.e., from the values of μ on a countable collection of sets. A possible choice for this collection of sets is supplied by the dyadic cubes. The following lemma is a key ingredient in the derivation of the multifractal formalism for measures.

Lemma 2 *Let μ be a nonnegative measure defined on $\mathbb{R}^d$. Then*

$$h_\mu(x_0) = -\frac{1}{j} \liminf_{j\to+\infty} (\log_2 (\mu[3\lambda_j(x_0)])) . \tag{24}$$

Proof. By definition of the Hölder exponent,

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

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

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

and it follows that $h_\mu(x_0)$ is bounded by the right hand-side of (24).

On the other hand, if $h_\mu(x_0) = H$, then there exist balls $B_n = B(x_0, r_n)$ and $\varepsilon_n > 0$ such that $r_n \to 0$, $\varepsilon_n \to 0$ and $r_n^{H+\varepsilon_n} \le \mu(B_n) \le r_n^{H-\varepsilon_n}$. Let j_n be such that

$$\frac{1}{2} 2^{-j_n} < r_n \le 2^{-j_n};$$

then $B_n \subset 3\lambda_{j_n}(x_0)$ so that $\mu(B_n) \le \mu(3\lambda_{j_n}(x_0))$, which implies the lower bound. □

Let μ be a probability measure on $\mathbb{R}$. It follows from Lemma 2 that it is natural to define a scaling function associated with μ by

$$\tau_\mu(p) = \limsup_{n\to\infty} -\frac{1}{j} \log_2 \left(\sum_{\lambda\in\Lambda_j} (\mu(3\lambda))^p \right).$$

Let us now show why the spectrum of singularities is expected to be recovered from the scaling function. The definition of the scaling function roughly means that $\sum \mu(3\lambda) \sim 2^{-\eta_\mu(p)j}$. Let us estimate the contribution to this sum of the cubes λ that cover the points of $E_\mu(H)$. Lemma 2 asserts that they satisfy $\mu(3\lambda) \sim 2^{-Hj}$. By definition of the fractal dimension, we need about $2^{-d_\mu(H)j}$ such cubes to cover $E_\mu(H)$; thus, the corresponding contribution can be estimated by $2^{-dj}2^{d_\mu(H)j}2^{-Hpj} = 2^{-(d-d_\mu(H)j+Hp)j}$. When $j \to +\infty$, the dominant contribution comes from the smallest exponent, so that $\tau_\mu(p) = \inf_H(d - d_\mu(H) + Hp)$. Inverting this Legendre duality relationship, and assuming that $d_\mu(H)$ is concave, we obtain

$$d_\mu(H) = \inf_H (d - \tau_\mu(p) + Hp). \tag{25}$$

Let us now show the relationship between the scaling function for measures, and the function spaces which contain the distribution function of this measure. Recall that the distribution function F_μ of μ is defined by $F_\mu(x) = \mu((-\infty, x]$.

Let $s \in (0,1)$ and $p > 1$. A function F belongs to the Sobolev spaces $W^{s,p}(\mathbb{R})$ if it satisfies

$$\exists C > 0 \text{ such that} \qquad \int\int \frac{|F(x)-F(y)|^p}{|x-y|^{sp+1}} dxdy \le C. \tag{26}$$

We take $F = F_\mu$, and we split the integral (26) as a sum on the subdomains

$$A_j = \{(x,y) : \ 2^{-j} \le |x-y| < 2 \cdot 2^{-j}\}.$$

The lefthand side of (26) is equivalent to

$$\sum_j 2^{j(sp+1)} \int\int_{A_j} |F_\mu(x) - F_\mu(y)|^p dxdy,$$

which, because F_μ is increasing, is equivalent to

$$\sum_j 2^{j(sp+1)} 2^{-2j} \sum_{\lambda \in \Lambda_j} |F_\mu(\lambda^+) - F_\mu(\lambda^-)|^p,$$

where $\lambda^+ = \frac{k+2}{2^j}$ and $\lambda^- = \frac{k-1}{2^j}$. But $F_\mu(\lambda^+) - F_\mu(\lambda^-) = \mu(3\lambda)$, and therefore the double integral in (26) is equivalent to

$$\sum_j 2^{j(sp-1)} \sum_{\lambda \in \Lambda_j} \mu(3\lambda)^p.$$

Coming back to the definition of $\tau_\mu(p)$ yields the function space interpretation

$$\tau_\mu(p) = \sup\{s : \ F_\mu \in W^{(s+1)/p,p}(\mathbb{R})\}, \tag{27}$$

which, using the embeddings between Sobolev and Besov spaces, implies that

$$\tau_\mu(p) = \eta_\mu(p) + 1.$$

3 Extensions of the scaling function: q-leaders and fractional integration

Advances concerning the construction of new scaling functions beyond (22) were motivated by the following restriction: In order to be used, the wavelet leader method requires the data to be locally bounded. We saw a practical procedure in order to decide if this assumption is valid, namely the determination of the uniform Hölder exponent H_f^{min}. Experimental investigations showed that H_f^{min} is negative for large classes of natural "texture type" images, see [46, 47, 48], and therefore the method cannot be used as such.

In order to circumvent this problem, one can either change the data or change the method. By changing the data, we will mean "smoothing " them, in order to obtain a new signal, with a positive index H_{min}, on which the previous analysis may be safely performed. On the other hand, by changing the method, we will mean replacing the wavelet leaders by alternative quantities, which measure pointwise regularity (for another definition of regularity) and make sense even if the data are no more locally bounded. Let us start by exploring this second possibility.

3.1 q-leaders

We will use the following extension of pointwise smoothness, which was introduced by Calderón and Zygmund in 1961, see [14].

Definition 9 *Let $B(x_0,r)$ denote the open ball centered at x_0 and of radius r; let $q \in [1,+\infty)$ and $\alpha > -d/q$. Let f be function which locally belongs to $L^q(\mathbb{R}^d)$. Then f belongs to $T^q_\alpha(x_0)$ if there exist $C,R > 0$ and a polynomial P such that*

$$\forall r \leq R, \quad \left(\frac{1}{r^d} \int_{B(x_0,r)} |f(x) - P(x-x_0)|^q dx \right)^{1/q} \leq C r^\alpha. \tag{28}$$

The q-exponent of f at x_0 is

$$h^q_f(x_0) = \sup\{\alpha : f \in T^q_\alpha(x_0)\}.$$

Note that the Hölder exponent corresponds to the case $q = +\infty$. This definition is a natural substitute for pointwise Hölder regularity when dealing with functions which are not locally bounded, but locally belong to L^q. In particular, the q-exponent can take values down to $-d/q$, and therefore it allows to model behaviors which locally are of the form $1/|x-x_0|^\alpha$ for $\alpha < d/q$, i.e., to deal with negative regularity exponents.

Let us now show how the notion of T^q_α regularity can be related to local l^q norms of wavelet coefficients. This will be done with the help of the following theorem, which yields a characterization of this pointwise smoothness. First, recall that the *local square functions at* x_0 are

$$S_f(j,x_0)(x) = \left(\sum_{\lambda \subset 3\lambda_j(x_0)} |c_\lambda|^2 1_\lambda(x) \right)^{1/2}.$$

The following theorem is proved in [28].

Theorem 3 *Let $q \in (0,\infty)$, $\alpha > -d/q$ and assume that the wavelet basis used is r-smooth with $r > \sup(2\alpha, 2\alpha + 2d(\frac{1}{q} - 1))$; if $f \in T^q_\alpha(x_0)$, then $\exists C \geq 0$ such that $\forall j \geq 0$,*

$$\| S_f(j,x_0) \|_q \leq C 2^{-j(\alpha + d/q)} \tag{29}$$

(if $q < 1$, then $\| \, . \, \|_q$ denotes the L^q quasi-norm).
Conversely, if (29) holds and if $\alpha \notin \mathbb{N}$, then $f \in T^q_\alpha(x_0)$.

Remark: When α is an integer, then the wavelet condition (29) characterizes a slightly different space, which implies a $T^q_\alpha(x_0)$ with a logarithmic loss on the modulus of continuity. This is reminiscent of the case of uniform Hölder spaces for $s = 1$, in which case the wavelet condition (15) characterizes the Zygmund class instead of the usual C^1 space.

We now relate local square functions and local l^q norms of wavelet coefficients. (The derivation that we propose slightly differs form the one of [32], since it is in the spirit of "wavelet leaders", whereas the one performed in [32] relies on extensions of two-microlocal spaces which were proposed by Y. Meyer and H. Xu in [41].)

Note that the condition $\| S_f(j,x_0) \|_q \leq C 2^{-j(\alpha + d/q)}$ implies that, at each scale $j' \geq j$,

$$\left\| \left(\sum_{\lambda' \subset 3\lambda_j(x_0), |\lambda'| = j'} |c_{\lambda'}|^2 1_{\lambda'}(x) \right)^{1/2} \right\|_q \leq C 2^{-j(\alpha + d/q)},$$

where, by definition, $|\lambda'|$ denotes the scale j' of the cube λ'. Since dyadic cubes at a given scale do not overlap, the left-hand side can be computed exactly and yields

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

It follows that the pointwise q-exponent can be expressed by a regression on a log-log plot of the *q-leaders*

$$d^q_\lambda = \left(\sum_{\lambda' \subset 3\lambda_j(x_0)} |c_{\lambda'}|^q 2^{-d(j'-j)} \right)^{1/q}. \tag{30}$$

Note that q-leaders are nothing but local Besov norms. Therefore, besides the fact that they are much easier to estimate than L^q norms of the local square function (as proposed in [29, 31]), they also offer the additional possibility of taking $q < 1$. In that case, the relationship with pointwise regularity is more intricate, see [29]. However, for classification purposes, one does not necessarily require an interpretation in terms of pointwise exponents.

As in the case of the usual wavelet leaders, one can associate function spaces with q-leaders: Let $p,q \in (0,\infty)$; a function f belongs to the *Oscillation space* $\mathscr{O}^s_{p,q}(\mathbb{R}^d)$ if and only if

$$\exists C,\ \forall j \geq 0, \qquad \sum_{\lambda \in \Lambda_j} \left[2^{(s-d/p)j} d^q_\lambda \right]^p \leq C.$$

These function spaces are closely related with the spaces $S^\alpha_{q,r,m}$ and $\mathscr{D}^\alpha_{q,r,m}$ introduced by A. Seeger in [45]. Indeed these spaces are also defined through local L^q norms or the function. However, the motivation of Seeger was different from ours: He wished to obtain a new characterization of Triebel-Lizorkin spaces, and therefore only considered the cases where the parameters α, q, r, m lead to these spaces, whereas our motivation is at the opposite, since we are interested in these spaces when they strongly differ from the classical ones, and lead to new scaling functions.

The q-scaling function is defined by

$$\forall p \in \mathbb{R}, \qquad \zeta_f(p,q) = \liminf_{j \to +\infty} \ -\frac{1}{j} \log_2 \left(2^{-dj} \sum_{\lambda \in \Lambda_j} |d^q_\lambda|^p \right). \tag{31}$$

In that case too, the scaling function is "well defined" for $p < 0$ (it satisfies the same robustness properties as the leader scaling function). When p is positive, the q-scaling function has the following function-space interpretation:

$$\zeta_f(p,q) = \sup\{s : f \in \mathscr{O}^{s/p}_{p,q}\}.$$

Note that, for a given q, this analysis is possible only if the data locally belong to L^q. This can be checked on the wavelet scaling function, since this condition will be satisfied if $\eta_f(q) > 0$. On the opposite, if $\eta_f(q) < 0$, this analysis is not relevant. We see here another use of the wavelet scaling function, as a preliminary quantity which is required to be computed before estimating the q-scaling function. Therefore it plays a similar role as the computation of H^{min}_f when dealing with the multifractal analysis based on wavelet leaders. Note that, as above, a spectrum can be attached to the q-exponent, and a multifractal formalism worked out, using the usual procedure; this spectrum is obtained as a Legendre transform of the q-scaling function, see [28, 29, 32]. An important open question is to understand, for a given function f, the relationships that exist between these q-spectra, as a function of the parameter q.

3.2 Fractional integration

Let us come back to the problem raised at the beginning of Section 3, namely the introduction of tools adapted to situations where wavelet leaders cannot be used because the data analyzed are not locally bounded. We mentioned that an alternative solution is to “smooth” the data, so that the exponent H_f^{min} becomes positive. This can be done through the use of a *fractional integration*. We now expose this procedure, both from a theoretical and practical point of view.

Definition 10 *Let f be a tempered distribution; the fractional integral of order s of f is the operator $(Id-\Delta)^{-s/2}$ defined as the convolution operator which amounts to multiply the Fourier transform of f with $(1+|\xi|^2)^{-s/2}$.*

Let ϕ be a C^∞ compactly supported function satisfying $\phi(x)=1$ for x in a neighbourhood of x_0. If $f \in L^1_{SG}$, its local fractional integral of order s is

$$f^{(-s)} = (Id-\Delta)^{-s/2}(\phi f). \tag{32}$$

Note that the function spaces which can be locally defined through the notion of local fractional integral do not depend on the function ϕ which is chosen. Indeed, denote by $f_1^{(-s)}$ and $f_2^{(-s)}$ the local fractional integrals of f corresponding to two different functions ϕ_1 and ϕ_2; $f\phi_1 - f\phi_2$ vanishes in a neighbourhood of x_0, and therefore is C^∞ in a neighbourhood of x_0. The local regularity properties of the operator $(Id-\Delta)^{-s/2}$ imply that $(Id-\Delta)^{-s/2}(f\phi_1 - f\phi_2)$ also is C^∞ in a neighbourhood of x_0; therefore $f_1^{(-s)}$ and $f_2^{(-s)}$ differ by a C^∞ function in a neighbourhood of x_0.

Numerically, the fractional integration can be replaced by a much easier procedure which shares the same properties in terms of regularity, a *pseudo-fractional integration* defined as follows, see [37].

Definition 11 *Let $s>0$, let ψ_λ be an r-smooth wavelet basis with $r>s+1$ and let f be a function, or a distribution, with wavelet coefficients c_λ. The pseudo-fractional integral (in the basis ψ_λ) of f of order s, denoted by $I^s(f)$, is the function whose wavelet coefficients (on the same wavelet basis) are*

$$c^s_\lambda = 2^{-sj} c_\lambda.$$

The following result shows that this numerically straightforward operation retains the same properties as the fractional integration.

Proposition 1 *Let f be a function satisfying $H_f^{min} > 0$. Then, for any $s \in \mathbb{R}$, the wavelet scaling functions of $I^s(f)$ and $f^{(-s)}$ coincide. It is also the case for Hölder exponents or q-exponents:*

$$\forall s>0,\ \ \forall x_0, \forall q \in (0,+\infty], \qquad h^q_{I^s(f)}(x_0) = h^q_{f^{(-s)}}(x_0).$$

The proof of this result requires to introduce several tools; the first one is the algebras $\mathscr{M}^\gamma$, which are defined as follows.

Definition 12 *An infinite matrix $A(\lambda,\lambda')$ indexed by the dyadic cubes of $\mathbb{R}^d$ belongs to $\mathscr{M}^\gamma$ if*

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

Matrices of operators which map an r-smooth wavelet basis onto another one belong to these algebras, as soon as $\gamma > r$, and more generally matrices (on wavelet bases) of pseudodifferential operators of order 0, such as the Hilbert transform in dimension 1, or the Riesz transforms in higher dimensions, belong to these algebras, see [37]. We denote by $\mathscr{O}p(\mathscr{M}^\gamma)$ the space of operators whose matrix on a r-smooth wavelet basis (for $r > \gamma$) belongs to $\mathscr{M}^\gamma$. Note that this space does not depend on the (smooth enough) wavelet basis which is chosen; indeed, the matrix of coordinate change from one r-smooth wavelet basis to another belongs to $\mathscr{M}^\gamma$ if $r > \gamma$, so that the independence is a consequence of the fact that the spaces $\mathscr{M}^\gamma$ are algebras.

The second tool that we will need is the notion of *vaguelette* system.

Definition 13 *A set of functions (θ_λ) indexed by the dyadic cubes of scale $j \geq 0$, forms a vaguelette system of order s if*

- *for any $j \geq 1$ the vaguelettes θ_λ of scale j have vanishing moment up to order $s+1$, i.e., if, for any multi-index α satisfying $|\alpha| \leq s+1$, then*

$$\int \theta_\lambda(x)x^\alpha dx = 0,$$

- *the θ_λ satisfy the following uniform decay estimates: For any multi-index α satisfying $|\alpha| \leq s+1$, then*

$$\forall N \in \mathbb{N}, \qquad \left|\frac{\partial^\alpha \theta_\lambda}{\partial x^\alpha}\right| \leq \frac{C_N 2^{(\alpha+d/2)j}}{(1+|2^j x-k|)^N}.$$

Biorthogonal vaguelette bases are couples of Riesz bases θ^1_λ and θ^2_λ which are both vaguelette systems and form biorthogonal bases. Therefore, $\forall f \in L^2$,

$$f(x) = \sum_\lambda c^1_\lambda \theta^2_\lambda, \quad \text{where} \quad c^1_\lambda = \int_{\mathbb{R}^d} f(x)\theta^1_\lambda dx.$$

The notions we introduced are related by the following key property, see [37]:

Proposition 2 *Let $\mathscr{M}$ be an operator which maps an r-smooth wavelet basis to a vaguelette system of order r; then, for any for $\gamma < r$, $\mathscr{M}$ belongs to $\mathscr{O}p(\mathscr{M}^\gamma)$.*

We now turn to the proof of Proposition 1

Proof. It is performed using the wavelet techniques developed in [37], such as the function spaces characterizations; therefore, we won't give a complete detailed proof, but only mention the main lines.

The first point consists in noticing that the systems

$$\psi^1_\lambda = 2^{sj}(Id-\Delta)^{-s/2}\psi_\lambda \qquad \text{and} \qquad \psi^2_\lambda = 2^{-sj}(Id-\Delta)^{s/2}\psi_\lambda$$

are biorthogonal vaguelette systems. This property is straightforward to check on the Fourier transform ψ^1_λ and ψ^2_λ, which, in this case, are completely explicit.

Note that

$$I^s(f) = \sum 2^{-sj} c_\lambda \psi_\lambda \qquad \text{and} \qquad f^{(-s)} = \sum 2^{-sj} c_\lambda \psi^1_\lambda.$$

If f is locally bounded, then $I^s(f)$ belongs locally to C^s (because of the wavelet characterization of C^s);

Assume now that $I^s(f)$ belongs to $C^\alpha(x_0)$; the operator that maps ψ_λ to ψ^1_λ belongs to $\mathcal{O}p(\mathcal{M}^\gamma)$, and therefore preserves the pointwise wavelet regularity criterium, see [22], and it also preserves the $T^q_\alpha(x_0)$ regularity, see [14]. Therefore, it is satisfied by $f^{(-s)}$ and, since the uniform Hölder exponent of $f^{(-s)}$ is positive, the converse part of the wavelet pointwise regularity criterium implies that it belongs to $C^\beta(x_0)$ for any $\beta < \alpha$, see [22, 26]. The proof of the converse part is similar, using the biorthogonality of the ψ^1_λ and ψ^2_λ. □

As regards signals and images for which $H^{min}_f < 0$, a possibility for applying to them a multifractal analysis based on the Hölder exponent is to perform on the data a fractional integration of sufficiently large order t; indeed, the uniform regularity exponent H^{min}_f is always shifted exactly by t, see [30]. This simple property shows a possible strategy in order to perform the multifractal analysis of a signal which is not locally bounded: First determine its exponent H^{min}_f; then, if $H^{min}_f < 0$, perform a fractional integration of order $t > -H^{min}_f$; it follows that the uniform regularity exponent of $I^t(f)$ is positive, and therefore it is a bounded function. The *t-leaders* are

$$d^t_\lambda = \sup_{\lambda' \subset 3\lambda} |c_{\lambda'} 2^{-tj'}| \tag{33}$$

The strategy we sketched leads to the consideration of the following *fractional leader scaling function* :

$$\forall p \in \mathbb{R}, \forall t > 0, \qquad \zeta_f(t,p) = \liminf_{j\to+\infty} \; -\frac{1}{j}\log_2\left(2^{-dj}\sum_{\lambda\in\Lambda_j} |d^t_\lambda|^p\right). \tag{34}$$

It can also be given a function-space interpretation: Let $p \in (0,\infty)$; a function f belongs to the *Oscillation space* $\mathcal{O}^{s,t}_p(\mathbb{R}^d)$ if and only if $(c^{(0)}_k) \in l^p$ and

$$\exists C,\ \forall j, \qquad \sum_{\lambda\in\Lambda_j} \left[2^{(s-d/p)j} d^t_\lambda\right]^p \le C,$$

see [25, 27] for properties of these spaces. Then

$$\zeta_f(p,t) = \sup\{s : f \in \mathscr{O}_p^{s/p,t}\}.$$

The Legendre transform of this scaling function is expected to yield, for each t, the spectrum of singularities of $I^t(f)$, see [25].

The strategy developed in this paragraph raises an open question: There is no canonical choice for the order of fractional integration, (the only condition is that it has to be larger than $-H_f^{min}$) and the quantities considered may depend on this order. This is one of the motivations for understanding how multifractal properties are modified under fractional integration. We will give some clues on these questions in Section 4 (see also [5, 30]).

3.3 The q-fractional scaling function

The two strategies developed in the last two subsections correspond to different answers of the following problem: How to perform a multifractal analysis of non locally bounded data. They led to the consideration of either q-exponents, or of fractional integrals. A third option consists in not making a choice but taking advantage of both strategies, and considering q-exponents of fractional integrals. It follows from Proposition 1 and the definition of q-leaders (30) that the q-exponent of a fractional integral of order t is given by a regression on a log-log plot of the following *grand leaders*

$$d_\lambda^{t,q} = \left(\sum_{\lambda' \subset 3\lambda_j(x_0)} \left| 2^{-tj'} c_{\lambda'} 2^{-d(j'-j)} \right|^q 2^{-dj'} \right)^{1/q}. \tag{35}$$

Not surprisingly, one can associate function spaces to grand leaders: Let $p, q \in (0, \infty)$; a function f belongs to the *Oscillation space* $\mathscr{O}_{p,q}^{s,t}(\mathbb{R}^d)$ if and only if

$$\exists C,\ \forall j \geq 0, \qquad \sum_{\lambda \in \Lambda_j} \left[2^{(s-d/p)j} d_\lambda^{t,q} \right]^p \leq C.$$

The grand scaling function is defined by

$$\forall p \in \mathbb{R}, \qquad \zeta_f(t,p,q) = \liminf_{j \to +\infty} \ -\frac{1}{j} \log_2 \left(2^{-dj} \sum_{\lambda \in \Lambda_j} |d_\lambda^{t,q}|^p \right).$$

Note that the computation of the grand leaders require that f^{-t} locally belongs to L^q. Once again, this can be checked on the wavelet scaling function of f: This requirement will be fulfilled as soon as

$$\eta_f(q) > -qt. \tag{36}$$

Note that, in that case too, $\zeta_f(t,p,q)$ is "well defined" for $p<0$ (it satisfies the same robustness properties as $\zeta_f(p)$). When p is positive, the grand scaling function has the following function space interpretation:

$$\zeta_f(t,p,q) = \sup\{s : f \in \mathcal{O}^{s/p,t}_{p,q}\}.$$

The derivation proposed by Parisi and Frisch can be adapted to the previous settings; indeed, as soon as a scaling function is supplied, one can derive a multifractal formalism, with corresponding upper bounds, following a general abstract procedure which was developed in [31]. We mention what it yields only in the most general setting supplied by the grand scaling function (actually, the other settings developed in Sections 3.1 and 3.2 can be considered as subcases). Let

$$E_f(H,q,t) = \{x_0 : h^q_{f^{-t}}(x_0) = H\}.$$

The corresponding *grand multifractal formalism* asserts that the *grand spectrum* $d_f(H,q,t) := \dim(E_f(H,q,t))$ can be deduced from the grand scaling function by a Legendre transform

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

Using either the scaling function $\zeta_f(t,p,q)$ or its Legendre transform $d_f(H,q,t)$ gives a very complete information on the nature of the pointwise singularities of f, and a rich tool for classification. However, by construction, it cannot yield information on the dimensions of the sets where two different exponents take given values. Such information requires the construction of *grand canonical formalisms*. Let us now explain in a general setting how they can be derived.

4 An introduction to the grand canonical formalism

A common feature shared by (21), (24), and (29) is that all the pointwise exponents we considered can be deduced from a countable number of quantities e_λ 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. In the measure case, $e_\lambda = \mu(3\lambda)$, in the pointwise Hölder case, $e_\lambda = d_\lambda$, and in the $T^p_\alpha(x_0)$ case, $e_\lambda = d^p_\lambda$, so that the general setting in which we work in this section will cover all the previous cases we already considered. These examples all fit in the following general framework for collections of positive quantities e_λ indexed by the dyadic cubes:

- The e_λ are increasing i.e., if $\lambda' \subset \lambda$ then $e_{\lambda'} \leq C\cdot e_\lambda$.
- A pointwise exponent is associated with the (e_λ):

$$h(x_0) = \liminf_{j\to+\infty} -\frac{1}{j}\log_2\left(e_{\lambda_j(x_0)}\right) \tag{37}$$

Grand canonical formalisms are fitted to situations where a couple of such exponents is considered: For example, they may correspond to the wavelet leaders of two functions f and g, in which case we will be interested in elaborating a multifractal formalism which yields the dimensions of the sets of points where each exponent takes a given value, or they may correspond to two different exponents associated with the same function. One important example is the couple formed by the Hölder and the oscillation exponents. Before deriving the grand canonical formalism, we recall the motivation for studying oscillation exponents.

4.1 Oscillation exponents: Discussion of possible definitions

All notions of pointwise singularity which have been considered are variants on the notion of "regularity exponent", which, roughly speaking, associates the exponent γ to the singularity $|x-x_0|^\gamma$ at x_0, at least if $\gamma > 0$ (and $\gamma \notin 2\mathbb{N}$), see [29, 38] for explicit general definitions of the notion of regularity exponent); such exponents include the (most widely used) Hölder exponent (see definition 6), the q-exponent (see definition 9), and the weak-scaling exponent, see [38]. However, one can wish to have information on how the function considered oscillates near the singularity at x_0: Consider for instance the "chirps"

$$F_{\gamma,\beta}(x) = |x-x_0|^\gamma \sin\left(\frac{1}{|x-x_0|^\beta}\right), \tag{38}$$

for a given regularity γ, their oscillatory behavior in the neighbourhod of x_0 increases with β; in this example, β parametrizes a "degree of oscillation". (Note that the case $\beta = 0$, i.e., $F_{\gamma,0}(x) = |x-x_0|^\gamma$ is usually referred to as a "cusp"). We will use a refined description of singularities by introducing an additional "oscillation" parameter, that allows to draw distinctions between functions which share the same Hölder exponent. Measuring such an additional exponent rises additional difficulties, one of them being that several mathematical definitions have been proposed, yielding different types of information.

One motivation for the detection of singularities such as (38) is that the existence of such behaviors has been conjectured in some physical data, such as fully developed turbulence. Another motivation is an internal mathematical requirement in multifractal analysis: We saw that one is often obliged to compute a fractional integral of the signal before performing its multifractal analysis. In that case, the singularity sets can be modified is a way which is difficult to predict if singularities such as (38) are present in the signal. Therefore understanding what multifractal analysis yields in this case requires the consideration of such behaviors.

We wish to describe strong local oscillations which display the same qualitative feature as in (38). A clue is supplied by the following remark. Let $\gamma > 0$, be given and let us estimate the primitive of (38). Since

$$\int_0^x |t-x_0|^\gamma \sin\left(\frac{1}{|t-x_0|^\beta}\right) dt = \int_0^x \frac{|t-x_0|^{\gamma+\beta+1}}{\beta}\left(\cos\left(\frac{1}{|t-x_0|^\beta}\right)\right)' dt,$$

it follows that, in the neighbourhood of x_0, this integral is the sum of

$$\frac{|x-x_0|^{\gamma+\beta+1}}{\beta}\cos\left(\frac{1}{|x-x_0|^\beta}\right)$$

and higher order terms; thus the Hölder exponent of $F_{\gamma,\beta}$ is γ, but the Hölder exponent of its primitive is $\gamma+\beta+1$; thus one can reasonably expect that the oscillation exponent can be recovered by comparing the regularity exponents of f and of its primitive.

Before proposing a precise mathematical procedure which allows to recover β, let us mention a natural requirements that a notion of "oscillation exponent" should satisfy in order to be of practical use. The definition proposed should allow for possible superpositions and "mixtures"; indeed, in the spirit of multifractal analysis, we do not expect these local behaviors to appear only in an isolated, "perfect" form as in (38), but rather for a dense set of values of x_0, and with possible corruptions by noise. Therefore one should find a key feature of (38) that characterizes the exponent β, and use it as a general definition of oscillating singularity. We noticed that β should be recovered by comparing the regularity exponents of f and of its primitive. We will need a slight extension of this remark in order to obtain a definition of oscillation exponent which fulfills this requirement. For that purpose, we will use the notion of *local fractional integral* already considered. The Hölder exponent of the local fractional integral of f of order s at x_0 is called the ***fractional Hölder exponent of*** f ***at*** x_0 and denoted by

$$h^s_f(x_0) = h_{f^{(-s)}}(x_0).$$

Lemma 3 *The definition of $h^s_f(x_0)$ does not depend on the function ϕ which is chosen.*

The proof is similar to the proof of Proposition 1. Properties of the fractional Hölder exponent have been investigated in [8, 35]. In particular one can show that it is a concave function.

It follows from the concavity of the fractional Hölder exponent that

$$\text{either} \qquad \forall s>0,\ h^s_f(x_0)=+\infty \qquad \text{or} \qquad \forall s>0,\ h^s_f(x_0)<+\infty;$$

an example of the first case is given by

$$|x-x_0|^\alpha \sin\left(\exp\left(\frac{1}{|x-x_0|}\right)\right),$$

where the Hölder exponent at x_0 is α, but, as soon as $s>0$, then $h^s_f(x_0)=+\infty$. When this first occurrence happens, the oscillation exponent should clearly be set to $+\infty$. Therefore, in the sequel, we suppose that the second case occurs. The definition of the oscillation exponent that we will choose is motivated by the simple but important remark that follows.

Lemma 4 *Let $f : \mathbb{R} \to \mathbb{R}$ be a locally bounded function, and let s be a positive integer. Denote by I_f^s an iterated primitive of order s of f; then*

$$h_f^s(x_0) = h_{I_f^s}(x_0).$$

Proof. Since the result is clearly local, we can assume that f is supported in a neighbourhood of x_0, and therefore belongs to L^2, which allows to use the Fourier transform without any restriction, and also allows to assume that $\phi = 1$ in the definition of the local fractional integral. Up to a polynomial term, one derives I_f^s from f by multiplying $\hat{f}$ by $(i\xi)^{-s}$; this iterated primitive has the same Hölder exponent as the one obtained using $|\xi|^{-s}$ instead. Indeed, the corresponding operators are either the same (up to a multiplicative constant), or deduced from each other by an Hilbert transform (the Fourier multiplier by *sign* (ξ)). In addition, applying the Hilbert transform does not modify the pointwise Hölder exponent since $f^{(-s)}$ has an exponent H^{min} which is positive, see [22]. The result follows by noticing that

$$\frac{1}{(1+|\xi|^2)^{s/2}} - \frac{1}{|\xi|^s} \sim \frac{C}{|\xi|^{s+2}} \qquad \text{when} \qquad |\xi| \to +\infty,$$

and therefore the corresponding operator is a uniformly smoothing operator which, for any $\alpha \in \mathbb{R}$, maps $C^{\alpha}(\mathbb{R})$ to $C^{\alpha+s+2}(\mathbb{R})$. Therefore,

$$h_{f^{(-s)} - I_f^s}(x_0) \geq 2 + \sup(h_{f^{(-s)}}(x_0), h_{I_f^s}(x_0)),$$

and we get $h_{f^{(-s)}}(x_0) = h_{I_f^s}(x_0)$. □

It follows from this lemma that results in dimension 1 which are checked by hand through the computation of primitives can then be extended to any dimension and to the non-integer case by using fractional integrals; for instance, let us check that

$$\forall s \geq 0, \qquad h^s_{F_{\gamma,\beta}}(x_0) = \gamma + (1+\beta)s, \tag{39}$$

where $F_{\gamma,\beta}$ was defined by (38). Indeed, a straightforward integration by parts shows that this result holds for iterated primitives, and thus, using Lemma 4, it also holds for fractional integrals of integer order; the result for all s follows from the concavity of the fractional Hölder exponent. Therefore, a natural definition for the oscillation exponent of an arbitrary function at x_0 should use the slope of the function

$$s \to h_f^s(x_0). \tag{40}$$

In the case of $F_{\gamma,\beta}$, this function is linear, and the definition is unambiguous; however, it is not always the case, and one can show (see [8, 35]) that, in general, functions defined by (40) only satisfy the following properties:

Proposition 3 *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function, and let $x_0 \in \mathbb{R}^d$. The function $s \to h^s_f(x_0)$ is concave; therefore it has everywhere a left and a right derivative. These derivatives satisfy*

$$\forall s \geq 0, \qquad \frac{\partial (h^s_f(x_0))}{\partial s} \geq 1.$$

Furthermore, these properties characterize fractional Hölder exponents.

It follows from this characterization that $h^s_f(x_0)$ is, in general, not a linear function of s, and therefore many choices are possible for its slope. In practice, only two choices have been used up to now, leading to two different exponents:

- The chirp exponent (choice of the slope "at infinity")

$$\gamma_f(x_0) = \lim_{s\to+\infty} \frac{\partial}{\partial s}(h^s_f(x_0)) - 1, \tag{41}$$

- the oscillation exponent (choice of the slope "at the origin")

$$\beta_f(x_0) = \lim_{s\to 0} \frac{\partial}{\partial s}(h^s_f(x_0)) - 1, \tag{42}$$

(see [33] for properties associated with chirp exponents, and [8] for properties associated with oscillation exponents). The drawback of using (41) is that this notion is very unstable: If g is an arbitrary smooth (but not C^∞) function, one can have

$$\gamma_f(x_0) = \Gamma > 0 \quad \text{but} \quad \gamma_{f+g}(x_0) = 0.$$

A simple example of this phenomenon is given by the functions: $f(x) = x\sin(1/x)$ and $g = |x|^a$ for an $a \notin 2\mathbb{N}$ and large enough. Thus, the chirp exponent does not satisfy the natural requirement of stability under the addition of smooth noise.

4.2 Properties and wavelet estimation of the oscillation exponent

The instabilities mentioned above cannot occur when using the oscillation exponent as shown by the following result, which states that, as soon as one imposes this stability requirement, then the choice of $\beta_f(x_0)$ is canonical.

Proposition 4 *Let f be a locally bounded function satisfying: $\forall s > 0$, $h^s_f(x_0) < +\infty$. Then $\beta_f(x_0)$ is the only quantity which satisfies the following properties:*

- *It is deduced from the function $s \to h^s_f(x_0)$.*
- *If $h_g(x_0) > h_f(x_0)$, then $\beta_{f+g}(x_0) = \beta_f(x_0)$ (the oscillation exponent of f is not altered under the addition of a function which is smoother than f).*
- *It yields the exponent β for the functions $F_{\gamma,\beta}$.*

Proof. Let us first check that the oscillation exponent satisfies these properties. We only have to check the second one; indeed, the first one follows from the definition, and the last one has already been proved. Because of the finiteness assumption for h^s_f, and its concavity, it is a continuous function; therefore, it remains strictly smaller than $h_g(x_0)$ in a small neighbourhood of x_0, and thus

$$\exists \eta > 0 : \quad \forall s < \eta, \qquad h^s_f(x_0) < h^s_g(x_0),$$

so that $h^s_{f+g}(x_0) = h^s_f(x_0)$.

Let us now prove the converse result. The stability requirement implies that the quantity considered cannot be a function of $h^s_f(x_0)$ for an $s > 0$; therefore it is a "germ property" at $s = 0$, and therefore, a function of the value at 0 of the function $s \to h^s_f(x_0)$ and its derivatives. Since $h^s_f(x_0)$ can be an arbitrary concave function, higher order derivatives do not exist in general; therefore, only $h_f(x_0)$ and the first derivative can be involved. Finally, the fact that the exponent takes the value β for the functions $F_{\gamma,\beta}$ implies that it is given precisely by (42). $\square$

Remarks: In practice, one cannot directly measure the oscillation exponent since it involves the estimation of how the Hölder exponent evolves under a fractional integration of "infinitesimal" order, and one rather estimates the evolution under a fractional integration of given fixed order s, thus obtaining the *s-oscillation exponent*:

$$\beta_f(s,x_0) = \frac{h^s_f(x_0) - h_f(x_0)}{s} - 1. \tag{43}$$

Because of the concavity of $h^s_f(x_0)$,

$$\forall s > 0, \qquad \gamma_f(x_0) \leq \beta_f(s,x_0) \leq \beta_f(x_0)$$

and

$$\beta_f(s,x_0) = 0 \iff \beta_f(x_0) = 0. \tag{44}$$

Note that the general definition of a cusp singularity at x_0 is given by the condition: $\beta_f(x_0) = 0$.

Note that (44) is of practical importance for the following reason: If one is interested in a qualitative information such as the existence or absence of oscillating singularities (and not the precise values taken by β_f), then, (44) shows that, in practice, one may as well work with the s-oscillation exponent rather than the oscillation exponent in order to obtain the required information.

We now recall the characterization of oscillating singularities, that was discovered by J.-M. Aubry, cf [10]. It makes use of the following notion introduced in [33], which is a natural generalization (in particular to higher dimensions) of the oscillating behavior of the sine function.

Definition 14 *A function $g \in L^\infty(\mathbb{R}^d)$ is called indefinitely oscillating if and only if $\exists \omega \in \mathscr{S}(\mathbb{R}^d)$ such that such $\forall N \in \mathbb{N}$, g can be written*

$$g = \omega + \sum_{|\alpha|=N+1} \partial^\alpha g_\alpha$$

with each $g_\alpha \in L^\infty(\mathbb{R}^d)$.

J.-M. Aubry proved that, if $h < h_f(x_0)$ and if f has oscillating exponent β, then f can be written, in the neighbourhood of x_0 as

$$f(x) = |x-x_0|^h g\left(\frac{x-x_0}{|x-x_0|^{\beta+1}}\right) + r(x),$$

where g is indefinitely oscillating, and r is smoother than the first term at x_0 (see [10]) for a more precise statement).

4.3 Derivation of β-leaders

Let us now show how, heuristically, one can derive quantities similar to leaders, that would yield the oscillation exponent on a log-log plot. Let us work out the general heuristic that backs the derivation of a multifractal formalism in the specific setting of the oscillation exponent (for the derivation in the initial setting of the Hölder exponent, see the seminal paper [43], and for a derivation in a dyadic setting adapted to the wavelet framework, see [26, 31]). We wish to express the oscillation exponent by a condition bearing on wavelet leaders. Recall that the leaders associated with the fractional integral of order t are the t-leaders (33).

It follows from Proposition 1 that

$$h^s_f(x_0) = \liminf_{j\to+\infty} \left(\frac{\log\left(d^s_j(x_0)\right)}{\log(2^{-j})} \right), \tag{45}$$

therefore, the oscillation exponent can be derived from the d^s_λ according to (42).

We will need the following notations. The first one is a weak form of the $\mathscr{O}$ notation of Landau: If F and G are two functions which tend to 0, $F = \overline{\mathscr{O}}(G)$ if

$$\liminf \frac{\log|F|}{\log|G|} \geq 1,$$

and the second one expresses the fact that two functions are of the same order of magnitude, disregarding "logarithmic corrections":

$$F \sim G \quad \text{if} \quad \lim \frac{\log|F|}{\log|G|} = 1.$$

The following proposition is a consequence of the previous remarks.

Proposition 5 *Let f be such that $H_f^{min} > 0$. The oscillating singularity exponents of f at x_0 are (h,β) if and only if its wavelet leaders satisfy the following conditions:*

1. $d_j(x_0) = \overline{\mathscr{O}}\left(2^{-hj}\right)$,
2. *there exists a sequence $j_n \to \infty$ such that*

$$d_{j_n}(x_0) \sim 2^{-hj_n}, \tag{46}$$

3. *there exists a sequence $\lambda'_n \subset 3\lambda_{j_n}(x_0)$ such that:*

$$\begin{cases} j'_n = (1+\beta)j_n + o(j_n) \\ |c_{\lambda'_n}| \sim d_{j_n}(x_0), \end{cases} \tag{47}$$

4. *β is the smallest number such that (47) holds.*

Remarks: The last condition means that the supremum in the definition of the wavelet leader $d_{j_n}(x_0)$ is "almost" attained at a scale close to $(1+\beta)j$.

Let $s > 0$ small enough be given. If f has an oscillating singularity with exponents (H,β) at x_0 then its wavelet leaders satisfy $d_j(x_0) \sim 2^{-Hj}$ and its integrated wavelet leaders satisfy $d_j^s(x_0) \sim 2^{-(H+s(1+\beta))j}$ By elimination, this allows to construct quantities that will scale like $2^{-\beta j}$ in the limit of small scales: the β*-leaders*

$$B_\lambda^s = 2^j \left(\frac{d_j^s(x_0)}{d_j(x_0)}\right)^{1/s},$$

and we expect that, if β is the oscillation exponent at x_0, then, for s small enough,

$$B_\lambda^s \sim 2^{-\beta j}. \tag{48}$$

4.4 Derivation of grand canonical formalisms

We now come back to the general setting supplied by two pointwise exponents given by the rate of decay, on a log-log plot, of two quantities e_λ and f_λ indexed by dyadic cubes. In order to derive a multifractal formalism for this couple of exponents, following the path suggested by the grand canonical formalism in statistical physics, we consider the following structure function

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

The associated scaling function is

$$\eta(p,q) = \liminf_{j \to +\infty} \left(\frac{\log(S_j(p,q))}{\log(2^{-j})}\right) \tag{49}$$

We will reformulate the fundamental idea due to G. Parisi and U. Frisch in this general setting. These authors gave an interpretation of the nonlinearity of the scaling function as the signature of the presence of different pointwise exponents (see [43] and also [21] for applications, particularly in the setting of invariant measures of dynamical systems).

The following notion of spectrum is adapted to the simultaneous consideration of two exponents.

Definition 15 *Let $(e_\lambda)_{\lambda\in\Lambda}$ and $(f_\lambda)_{\lambda\in\Lambda}$ be increasing dyadic functions, and let*

$$E_{H_1,H_2} = \{x: \quad h_e(x) = H_1 \text{ and } h_f(x) = H_2\}.$$

The spectrum of singularities associated with the $(e_\lambda)_{\lambda\in\Lambda}$ is the function $d(H_1,H_2)$ defined by

$$d(H_1,H_2) = \dim(E_{H_1,H_2}).$$

The support of the spectrum is the set of values (H_1,H_2) for which $E_{H_1,H_2} \neq \emptyset$.

Let us now show heuristically how the spectrum of singularities can be recovered from the scaling function. The definition of the scaling function (49) means that,

$$\text{when } j \to +\infty, \qquad S_j(p,q) \sim 2^{-\eta(p,q)j}.$$

Let us estimate the contribution to $S_j(p,q)$ of the dyadic cubes λ that cover the points of E_{H_1,H_2}. By definition of E_{H_1,H_2}, they satisfy

$$e_\lambda \sim 2^{-H_1 j} \quad \text{and} \quad f_\lambda \sim 2^{-H_2 j};$$

by definition of $d(H_1,H_2)$, since we use cubes of the same width 2^{-j} to cover E_{H_1,H_2}, we need about $2^{-d(H_1,H_2)j}$ such cubes; therefore the corresponding contribution is of the order of magnitude of

$$2^{-dj}2^{d(H_1,H_2)j}2^{-(H_1p+H_2q)j} = 2^{-(d-d(H_1,H_2)+H_1p+H_2q)j}.$$

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

$$\eta(p,q) = \inf_H (d - d(H_1,H_2) + H_1 p + H_2 q). \tag{50}$$

In the next section, we will show that the scaling function $\eta(p,q)$ is a concave function on $\mathbb{R}$, which is in agreement with the fact that the right-hand side of (50) necessarily is a concave function (as an infimum of a family of linear functions) no matter whether $d(H_1,H_2)$ is concave or not. However, if $d(H_1,H_2)$ also is a concave function, then the Legendre transform in (50) can be inverted (as a consequence of the duality of convex functions), which justifies the following assertion:

Definition 16 *A sequence (e_λ, f_λ) follows the multifractal formalism if its spectrum of singularities satisfies*

$$d(H_1,H_2) = \inf_{(p,q)\in\mathbb{R}^2} (d - \eta(p,q) + H_1 p + H_2 q). \tag{51}$$

The derivation exposed above is not a mathematical proof, and we do not expect (51) to hold in complete generality. However, in applications, it often happens that the spectrum of singularities itself has no direct interpretation and multifractal analysis is only used as a classification tool in order to discriminate between several types of signals; when such is the case, one is no more concerned with the validity of (51) but only with having its right-hand side defined in a numerically precise way.

Note that, in the specific case where the couple of exponents considered is the Hölder exponent and the oscillation exponent, one applies the above general derivation using wavelet leaders and β-leaders. This justifies the grand-canonical multifractal formalism proposed in [5, 30].

4.5 Concavity of the scaling function

We will now prove that the function η is concave on $\mathbb{R}^2$, a property that was used in the derivation of the multifractal formalism

Proposition 6 *The function η defined by (49) is concave on $\mathbb{R}^2$.*

In order to prove Proposition 6, we will need the following lemma.

Lemma 5 *Let $(a_i)_{i\in\mathbb{N}}$ and $(b_i)_{i\in\mathbb{N}}$ be sequences of positive real numbers. The function $\omega : \mathbb{R}^2 \longrightarrow \overline{\mathbb{R}}$ $(= \mathbb{R}\cup\{+\infty,-\infty\})$ defined by*

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

is convex on $\mathbb{R}^2$.

Proof. We need to check that $\forall (p,q),(p',q')\in\mathbb{R}^2,\ \forall\alpha\in]0,1[$,

$$\omega\left(\alpha(p,q)+(1-\alpha)(p',q')\right) \le \alpha\omega(p,q)+(1-\alpha)\omega(p',q'). \tag{52}$$

Consider the sequences

$$A = \left((a_1^p b_1^q)^\alpha, \ldots (a_N^p b_N^q)^\alpha, \cdots\right) \quad \text{and} \quad B = \left((a_1^{p'} b_1^{q'})^{1-\alpha}, \ldots (a_N^{p'} b_N^{q'})^{1-\alpha}, \cdots\right);$$

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

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

Taking logarithms on both sides of this inequality yields (52). □

We will now show that η is concave on $\mathbb{R}^2$. For each j, one applies Lemma 5 to the sequences (e_λ) and (f_λ). We obtain that, for any $j \geq 0$, the function

$$(p,q) \to \log \left(2^{-dj} \sum_{\lambda \in \Lambda_j} d_\lambda^p f_\lambda^q \right)$$

is convex; therefore, after dividing by $\log(2^{-j}) < 0$, we obtain a concave function; since concavity is preserved by taking infimums and pointwise limits, the concavity of the scaling function follows.

5 Concluding remarks

Let us summarize the whole strategy that we proposed in order to find the relevant multifractal parameters attached to an image. The first step consists in computing the wavelet scaling function $\eta_f(q)$ defined for $q > 0$ by (14), and the additional parameter supplied by the uniform Hölder exponent H_f^{min} defined by (17); note that these parameters can be computed in all cases, without any regularity assumption on the data. They can be used for classification: For instance preliminary studies made on a collection of natural textures show that their exponents H_f^{min} may widely differ, and therefore that it is a pertinent classification parameter, see [46, 47, 48].

These parameters also serve another purpose as a preliminary step in order to go further: H_f^{min} allows to determine if the leader scaling function (defined by (22)) is well defined: It is the case if $H_f^{min} > 0$. This scaling function has already proved useful in many situations: A particularly striking example is supplied by deciding which models for fully developed turbulence are relevant, see [34] : This is a typical situation where H_f^{min} and $\eta_f(q)$ are not discriminatory (for $p > 0$, several models fit the experimental values measured on turbulent data); on the other hand, the fact that $\zeta_f(p)$ is defined for negative values of p yields a range of parameters on which discrimination can be efficiently performed (note that this is in sharp distinction with the wavelet scaling function, which is defined only for $p > 0$). This example also shows the importance of using the corresponding parameters despite the fact that they do not have a function space interpretation, which is the case of scaling functions associated with negative ps, see [3, 4].

If $H_f^{min} < 0$, then we have the choice between two strategies. One consists in performing a pseudo-fractional integration of order larger than $-H_f^{min}$, so that the new signal has a positive uniform Hölder exponent. This strategy has been successfully applied to images, and yields a set of parameters which are well defined, and also proved useful for classification, see [46, 47, 48].

On the other hand, this method is not universal, and one can easily check on the following toy examples why this strategy may have for consequence a loss of information on the nature of the singularities present in the signal. Consider for instance the function

$$F(x) = |x-x_1|^{\alpha} \sin\left(\frac{1}{|x-x_1|^{\beta}}\right) + |x-x_2|^{\delta} + B_{\gamma}(x),$$

where B_γ is a Fractional Brownian Motion (FBM) of order $\gamma \in (0,1)$. We assume that $0 < \alpha < \gamma$, and $-1 < \delta < 0$ so that $F \in L^p_{loc}$ as soon as $\alpha > -1/p$. Since $\delta < 0$, then, the computation of wavelet leaders cannot be performed, and a fractional integration of order $t > -\delta$ is required so that $H^{min}_{I^t(f)} > 0$. But if $\alpha + t(1+\beta) > \gamma$, then this fractional integration make the oscillating singularity disappear, since its exponent becomes larger that the exponent of the integrated FBM. It is easy to check that such situations do happen: If $\alpha > \gamma + \delta\beta$, then either the order of fractional integration is too small, and the wavelet leaders are not well defined, or it is too large, and the oscillating singularity has been too much smoothed, and cannot be detected. This example shows that performing a fractional (or pseudo-fractional) integration has the effect of smoothing the data, and can make the oscillating singularities disappear in the presence of noise. Therefore, a better strategy in order to detect the presence of oscillating singularities can be to deal with a q-scaling function of the signal; in the more favorable cases, we won't have to perform any fractional integration at all, because, for some values of q, $\eta_f(q) > 0$. It is the case in the example we picked, since $\delta > -1$; see also [46, 47, 48] for examples taken from natural textures where H^{min}_f is negative, but nonetheless, the wavelet scaling function takes positive values. In the worst cases where $\eta_f(q)$ is negative for all values of q, then one has to perform a fractional integration; however, one can use the grand scaling function for values of the fractional integration parameter t which are smaller than when dealing with leaders, and thus less information concerning the presence of oscillating singularities will be lost. The previous example showed the relevance of the computation of the wavelet scaling function of f: If, for a given q, $\eta_f(q) > 0$, then the q-scaling function (defined by (31)) is well defined for this value of q. The determination of this scaling function has also been advocated as a useful tools in order to perform a multifractal analysis of fractal boundaries, see [32]. Therefore, this discussion shows situations where the q scaling function can be required in order to perform a multifractal analysis, and even situations where the grand scaling function proves necessary.

We are particularly thankful to the anonymous referee for many useful comments, and to Maïtine Bergounioux for for her constant help and encouragements during the redaction of this text.

References

1. P. Abry, S. Jaffard, B. Lashermes (2006) Wavelet leaders in multifractal analysis, “Wavelet Analysis and Applications”, T. Qian et al. eds., pp. 201–246, “Applied and Numerical Harmonic Analysis”.
2. P. Abry, S. Jaffard, S. Roux, B. Vedel, H. Wendt (2008) Wavelet decomposition of measures: Application to multifractal analysis of images Unexploded ordnance detection and mitigation,

J. Byrnes ed. Springer, NATO Science for peace and security, Series B 1–20 .
3. P. Abry, B. Lashermes, S. Jaffard (2004) Revisiting scaling, multifractal and multiplicative cascades with the wavelet leader lens, Optic East, Wavelet applications in Industrial applications II **5607**, pp. 103-117 Philadelphia, USA.
4. P. Abry, B. Lashermes, . Jaffard (2005) Wavelet leader based multifractal analysis, Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing
5. P. Abry, S. Roux, S. Jaffard (2010) Detecting oscillating singularities in multifractal analyis: Application to hydrodynamic turbulence, Preprint.
6. P. Abry, F. Sellan (1996) The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation, Appl. Comput. Harmon. Anal. **3** No.4, pp. 377–383.
7. A. Arneodo, B. Audit, N. Decoster, J.-F. Muzy, C. Vaillant (2002) 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.
8. A. Arneodo, E. Bacry, S. Jaffard, J.-F. Muzy (1998) Singularity spectrum of multifractal functions involving oscillating singularities J. Four. Anal. Appli., **4**, pp. 159-174.
9. A. Arneodo, E. Bacry, J.-F. Muzy (1995) The thermodynamics of fractals revisited with wavelets, Phys. A, vol. **213**, pp. 232–275.
10. J.-M. Aubry (1999) Representation of the singularities of a function, Appl. Comp. Harmon. Anal., **6**, pp. 282–286.
11. J.-M. Aubry, S. Jaffard (2002) Random Wavelet Series, Comm. Math. Phys. , **227** , pp. 483–514.
12. J.-F. Aujol, A. Chambolle (2005) Dual norms and image decomposition models, Int. J. Comput. Vis. , **63**, pp. 85–104.
13. G. Brown, G. Michon, J. Peyrière (1992) On the multifractal analysis of measures, J. Statist. Phys., **66**, pp. 775–790.
14. A. P. Calderòn, A. Zygmund (1957) Singular integral operators and differential equations, Amer. J. Math., **79, pp.** 901–921.
15. A. Cohen, I. Daubechies, .-C. Fauveau (1992) Biorthogonal bases of compactly supported wavelets, Comm. Pure Appl. Math., **44**, pp. 485–560.
16. A. Cohen, R. Ryan (1995) Wavelets and Multiscale Signal Processing, Chapman and Hall
17. I. Daubechies (1992) Ten lectures on wavelets, S.I.A.M..
18. K. Falconer (1990) Fractal Geometry, Wiley.
19. H. Feichtinger (1992) Wiener amalgams over Euclidean spaces and some of their applications. Lect. Notes Pure Appl. Math., **136**, pp. 123-137 , Springer Verlag.
20. Y. Gousseau, J.-M. Morel (2001) Are natural images of bounded variation?, SIAM J. Math. Anal. **33**, pp. 634–648.
21. T. Halsey, M. Jensen, L. Kadanoff, I. Procaccia , B. Shraiman (1986) Fractal measures and their singularities: The characterization of strange sets, Phys. Rev. A, **33**, pp. 1141–1151.
22. S. Jaffard (1991) Pointwise smoothness, two-microlocalization and wavelet coefficients, Publ. Matem., **35**, pp. 155–168.
23. S. Jaffard (1997) Multifractal formalism for functions, SIAM J. Math. Anal., **28**, pp. 944–998.
24. S. Jaffard (2000) Lacunary wavelet series, Ann. App. Prob., **10**, no. 1, pp. 313–329.
25. S. Jaffard (2005) Beyond Besov spaces Part 2: Oscillation spaces, Const. Approx., **21**, no 1, pp. 29–61.
26. S. Jaffard (2004) 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, **72**, Part 2, pp. 91–152.
27. S. Jaffard (1998) Oscillation spaces: Properties and applications to fractal and multifractal functions, J. Math. Phys., **39**, pp. 4129–4141.
28. S. Jaffard(2006) Pointwise regularity associated with function spaces and multifractal analysis, Banach Center Pub., Approximation and Probability, T. Figiel and A. Kamont Eds., **72**, pp. 93–110.

29. S. Jaffard (2006) Wavelet Techniques for pointwise regularity, Ann. Fac. Sci. Toul., **15**, no. 1, pp. 3–33.
30. S. Jaffard, P. Abry, S. Roux (2009) Singularités oscillantes et coefficients d'ondelettes dominants, Proc. of the GRETSI Conference at Dijon .
31. S. Jaffard, P. Abry, S. Roux, B. Vedel, H. Wendt (2010) The contribution of wavelets in multifractal analysis, in "Wavelet methods in Mathematical Analysis and Engineering , Higher Education Press, " Series in Contemporay Applied Mathematics" CAM **14**, pp. 51–98.
32. S. Jaffard, C. Melot (2005) Wavelet analysis of fractal Boundaries, Part 1: Local regularity and Part 2: Multifractal formalism, Comm. Math. Phys. **258**, no 3, pp. 513–539.
33. S. Jaffard, Y. Meyer (1996) Wavelet methods for pointwise regularity and local oscillations of functions, Mem. Amer. Math. Soc., vol. 1**23**, no. 587.
34. B. Lashermes, S. Roux, P. Abry ,S. Jaffard (2008) Comprehensive multifractal analysis of turbulent velocity using wavelet leaders, Eur. Phys. J. B., Vol. **61**, (2), pp. 201–215.
35. J. Lévy-Véhel, S. Seuret (2002) The local Hölder function of a continuous function, Appl. Comp. Harm. Anal. Vol. **13**, no3 , pp. 263–276.
36. S. Mallat (1998) A Wavelet Tour of Signal Processing, Academic Press.
37. Y. Meyer (1990) Ondelettes et Opérateurs, Hermann.
38. Y. Meyer (1998) Wavelets, Vibrations and Scalings, CRM Ser. AMS Vol. **9**, Presses de l'Université de Montréal .
39. Y. Meyer (2001) Oscillating patterns in image processing and nonlinear evolution equations, University Lecture Series **22**, AMS.
40. Y. Meyer, F. Sellan, M. Taqqu (1999) Wavelets, generalized white noise and fractional integration: The synthesis of fractional Brownian motion, J. Fourier Anal. Appl., **5**, No.5, pp. 465–494.
41. Y. Meyer, H. Xu(1997) Wavelet analysis and chirps, Appl. Comput. Harmon. Anal., **4**, pp. 366–379.
42. S. Osher, A. Solé, L. Vese (2003) Image decomposition and restoration using total variation minimization and the L^1 norm. Multiscale Model Simul., **1**, pp. 349–370
43. G. Parisi, U. Frisch(1985) 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, North Holland, pp. 84–88 .
44. L.I. Rudin, S. Osher, E. Fatemi, (1992) Nonlinear total variation based noise removal algorithms. Physica D, **60**, pp. 259–268.
45. A. Seeger (1989) A note on Triebel-Lizorkin spaces, Approximation and Function spaces, Banach Center Publications, **22**, pp. 391–400.
46. H. Wendt, P. Abry, S. Jaffard (2007)Bootstrap for Empirical Multifractal Analysis IEEE Signal Proc. Mag., **24**, no. 4, pp. 38–48 .
47. H. Wendt, P. Abry S. Roux, S. Jaffard, B. Vedel (2009) Analyse multifractale d'images: l'apport des coefficients dominants Traitement du Signal, vol. **25,** no. 4-5, pp. 47–65.
48. H. Wendt, S. Roux P. Abry, S. Jaffard (2009) Wavelet leaders and Bootstrap for multifractal analysis of images Signal processing, **89**, pp. 1100–1114.