2-Microlocal Analysis and Application in Signal Processing

Bertrand Guiheneuf, Jacques Lévy Véhel

2-MICROLOCAL ANALYSIS AND APPLICATIONS IN SIGNAL PROCESSING

Bertrand Guiheneuf, Jacques Lévy Véhel

Projet FRACTALES, INRIA Rocquencourt
78153 Le Chesnay - FRANCE
Bertrand.Guiheneuf@inria.fr, Jacques.Levy_Vehel@inria.fr
http://www-rocq.inria.fr/fractales/

1. INTRODUCTION

Hölder regularity has numerous applications in signal processing [1, 2, 3, 4, 5, 6]: classification (for instance, in medical images such as scintographies, or for computer traffic traces), detection, segmentation (e.g. satellite images) or synthesis (for instance of speech signals). A lot is known on the theoretical properties of such a characterization [7, 8], and many papers have also been devoted to the important problem of obtaining reliable numerical estimates of the Hölder exponent [9, 10].

Nevertheless, there are well known situations where the Hölder exponent gives totally irrelevant and sometimes misleading information. The simplest case is that of the chirp $|x|^\gamma sin(\frac{1}{|x|^\delta})$, but more complex signals arise for instance in the frame of multifractal analysis. In such cases, several authors [11, 8, 12] have proposed to refine the regularity analysis by adjoining to the Hölder exponent α at point x a second exponent β which characterizes , in some sense, the "oscillation" of the signal around x. This is particularly clear in the the case of chirp, since one simply has $\alpha = \gamma$ and $\beta = \delta$. Though several variants exist, in most cases, one replaces a single exponent by a couple of values, that give a more complete information concerning the local regularity. While this enhanced description has many potential applications (the most celebrated one being the detection of gravitational waves), a major problem is the estimation of (α, β) on real data, which seems to be a difficult task.

On the other hand, the couple (α, β) is far from exhausting the local regularity information. To obtain a complete picture, one indeed needs to define a "regularity function" at each point, which, loosely speaking, describes how the Hölder exponent behaves under differentiation/integration of the signal. The technical name of such an tool is 2-microlocal analysis, and it was introduced by Bony [13] in the frame of EDP. Though it may seem a formidable task to a whole regularity function instead of a single number, such an information is needed in some applications, as image denoising, which is described in section 4. A careful study of 2-microlocal analysis shows however that this hierarchy of exponents is very much constrained, so that its determination is greatly simplified. Moreover, theoretical results such as the one described in theorem 3.2 tend to indicate that taking into account the whole available information might indeed in some cases ease the computation of the Hölder and chirp exponents.

This paper is organized as follows: Section 2 recalls the definitions of the different Hölder exponents and gives the basics of 2-microlocal analysis. Section 3 focuses on the properties of the 2-microlocal frontier which is defined and thoroughly studied. Section 4 presents an application to a topic in image analysis, namely image denoising. A general method is described and experiments on SAR images are presented, that support the idea that 2-microlocal analysis might indeed become an important tool in signal processing.

2. INTRODUCTION TO 2-MICROLOCAL ANALYSIS

2.1. Hölder regularity

There are two main paths of measuring the regularity of a function $f : K \to \mathbb{R}$ where K is a bounded subset of $\mathbb{R}$. The first one is geometrical in nature, and consists in evaluating a (fractional) dimension of its graph Γ. Roughly speaking, one seeks to determine how much Γ "fills the space" at small scales. The precise way of measuring this property depends on the definition of the dimension one uses : the most frequent choices are the Hausdorff, the box, and the Tricot dimension ([14])

The second, analytical, way of evaluating the regularity of f is to consider a family of (nested) functional spaces, and to determine the ones it actually belongs to. A popular choice is to consider Hölder spaces, either in their global, local or pointwise version (see below for precise definitions).

Of course, the selection of one or another approach, and the subsequent choice of the dimension or of the functional spaces used eventually depends on the kind of properties one wishes to investigate.

We will be interested in this paper in refinements of the Hölder characterization of regularity. Let us first recall some basic definitions.

Definition 2.1. *Pointwise Hölder exponent*

Let α be a positive real number, $\alpha \notin \mathbb{N}$, and $x_0 \in \mathbb{R}$. A function $f : \mathbb{R} \to \mathbb{R}$ is in $C^{\alpha}_{x_0}$ if there exists a polynomial P of degree less than α such that:

$$|f(x) - P(x - x_0)| < c|x - x_0|^{\alpha}. \quad (1)$$

When $\alpha \in]0, 1[$, this reduces to:

$$|f(x) - f(x_0)| < c|x - x_0|^{\alpha} \quad (2)$$

The pointwise Hölder exponent of f at x_0, denoted $\alpha(x_0)$, is the supremum of the α for which (1) holds.

The pointwise Hölder function of f is defined as:

$$\alpha(x) = \sup \{\alpha : f \in C^{\alpha}_{x}\}.$$

This regularity characterization is widely used in fractal analysis because it has direct interpretations both mathematically and in applications. It has been shown for instance [1] that α_f indeed corresponds to the auditive perception of smoothness for voice signals. Similarly, simply computing the Hölder exponent at each point of an image already gives a good idea of its structure, as for instance its edges [15]. More generally, in many applications, it is desirable to model, synthetize or process signals which are highly irregular, and for which the relevant information lies in the singularities more than in the amplitude (this is indeed the case for image segmentation). In such cases, the study of the Hölder function is of obvious interest. From a theoretical point of view, the structure of α_f is known [7]: the class of Hölder functions of continuous signals is exactly the set of lower limits of continuous functions. This allows to handle a large variety of situations. However, the pointwise Hölder characterization has a number of drawbacks, a major one being that it is not stable under the action of (pseudo) differential operators. This precludes, for instance, the use of the Hilbert transform, commonly used in signal processing. In the same way, knowing the pointwise Hölder exponent of a function at a point x_0 is not sufficient to predict the Hölder exponent of its derivative at the same point. Finally, this exponent is generally hard to compute numerically, in particular in the case of multifractal signals.

An alternative solution is to consider the local Hölder exponent: Let $\alpha \in]0, 1[$, $\Omega \subset \mathbb{R}$. One says that $f \in C^{\alpha}_{l}(\Omega)$ if:

$$\exists C : \forall x, y \in \Omega : \frac{|f(x) - f(y)|}{|x - y|^{\alpha}} < C$$

Then, let:

$$\alpha_l(f, x_0, \rho) = \sup \{\alpha : f \in C^{\alpha}_{l}(B(x_0, \rho))\}$$

$\alpha_l(f, x_0, \rho)$ is non increasing as a function of ρ.
We are now in position to give the definition of the local Hölder exponent :

Definition 2.2. *Let f be a continuous function. The* local Hölder exponent *of f at x_0 is the real number:*

$$\alpha_l(f, x_0) = \lim_{\rho \to 0} \alpha_l(f, x_0, \rho)$$

This exponent is stable under differentiation or integration. Moreover, it is easier to estimate than the pointwise Hölder exponent. Its main drawback is that the *local Hölder function*, $\alpha_l(f, .)$, is a lower semi continuous function (l.s.c), i.e. :

$$\forall x_0 \in \mathbb{R}, \forall \epsilon, \exists \eta : \\ y \in B(x_0, \eta) \Rightarrow \alpha_l(f, y) > \alpha_l(f, x_0) - \epsilon \quad (3)$$

Recall the following result:

Lemma 2.1. *[16] Two l.s.c. functions which coincide on an everywhere dense set are equal.*

This shows that the local Hölder exponent is a less versatile notion than its pointwise counterpart.

In summary, neither α nor α_l is complely satisfactory. One would thus like to find a framework which would combine the good properties of each exponent. Such a framework indeed exists : it provides the added benefit of unifying the local and pointwise exponents, and allows to define functionnal spaces that behave nicely under (pseudo) differential operators.

2.2. 2-microlocal spaces: Definition and basic properties

The original definition of the 2-microlocal spaces [13] is based on a Littlewood Paley analysis (L.P.A).A Littlewood Paley analysis is a spatially localized filter bank. One may understand it as an intermediate between a discrete and a continuous wavelet analysis. More precisely, let $\mathcal{S}(\mathbb{R})$ be the Schwartz space defined as:

$$\mathcal{S}(\mathbb{R}) = \{f \in C^{\infty}(\mathbb{R}) : \\ \forall (\alpha, \beta) \in \mathbb{N}^2, \sup_x |x^{\alpha} \partial^{\beta} f(x)| < \infty\}. \quad (4)$$

Let now

$$\varphi \in \mathcal{S}(\mathbb{R}) = \left\{ \begin{array}{l} \widehat{\varphi}(\xi) = 1, \| \xi \| < \frac{1}{2} \\ \widehat{\varphi}(\xi) = 0, \| \xi \| > 1. \end{array} \right.$$

and

$$\varphi_j(x) = 2^j \varphi(2^j x).$$

One has

$$\widehat{\varphi}_j(\xi) = \widehat{\varphi}(2^{-j}\xi).$$

The $\{\varphi_j\}$ set acts as low pass filter bank, which leads naturally to the associated band pass filter bank:

$$\psi_j = \varphi_{j+1} - \varphi_j.$$

Definition 2.3. *Let $u \in \mathcal{S}'(\mathbb{R})$. The Littlewood Paley Analysis of u is the set of distributions:*

$$\left\{ \begin{array}{lcl} S_0 u & = & \varphi * u \\ \Delta_j u & = & \psi_j * u \end{array} \right.$$

One has:

$$\boxed{u = S_0 u + \sum_{j=0}^{\infty} \Delta_j u.}$$

We can now define the two microlocal spaces $C_{x_0}^{s,s'}$.

Definition 2.4. *A distribution $u \in \mathcal{S}'(\mathbb{R})$ belongs to the 2-microlocal space $C_{x_0}^{s,s'}$ if:*

$$\exists \quad c > 0 \quad s.t. \quad \left\{ \begin{array}{l} |S_0 u(x)| < c(1 + |x - x_0|)^{-s'} \\ |\Delta_j u(x)| < c 2^{-js} \ (1 + 2^j |x - x_0|)^{-s'} \end{array} \right.$$

The Littlewood Paley analysis is a good starting point in a attempt to generalize the previous regularity notion. Indeed, L.P.A. proves to be very useful to characterize many regularity spaces. For example, LPA gives equivalent norms on function spaces such as L^P, global Hölder spaces and more generally on all the Besov spaces:

$$\| f \|_{B_p^{\alpha,q}} \sim \| f \|_p + \left(\sum_{j=0}^{+\infty} [2^{\alpha j} \| \Delta_j f \|_p]^q \right)^{\frac{1}{q}} \quad (0 < \alpha < 1).$$

Let us indicate why 2-microlocal spaces are stable under differentiation.
First, derive the relation between $\Delta_j(u)$ and $\Delta_j(u')$:

$$\widehat{\Delta_j(u')}(\xi) = \widehat{u'}.\widehat{\psi}_j(\xi) = i\xi \widehat{u}(\xi).\widehat{\psi}_j(\xi),$$

but

$$\text{supp } (u', \widehat{\varphi}_j) \subseteq \text{ supp}(\widehat{\varphi}_j) \subseteq \pm[2^{j-1}, 2^{j+1}].$$

We can thus define a function

$$E(x) \in \mathcal{S}(\mathbb{R}) : \widehat{E}(\xi) = \xi \text{ on } \pm[\frac{1}{2}, 1].$$

Let then $E_j(\xi) = 2^{2j} E(2^j x)$.
One gets:

$$\widehat{E}_j(\xi) = \xi \text{ on } \pm[2^{j-1}; 2^{j+1}].$$

Which leads to:

$$\begin{array}{rcl} \Delta_j(u') & = & i\xi \widehat{u}(\xi) \widehat{\psi}_j(\xi) = i \widehat{E}_j(\xi) \widehat{u}(\xi) \widehat{\psi}_j(\xi) \\ & = & i \ E_j * \underbrace{\Delta_j(u)}_{u_j}. \end{array}$$

Finally:

$$\begin{array}{rcl} |\Delta_j(u')| & = & 2^{2j} \left| \int f_j(t - x) E(2^j) dx \right| \\ & = & 2^j \left| \int f_j(t - 2^j x) E(x) dx \right| \end{array}$$

We now make use of the following lemma:

Lemma 2.2. *Peetre inequality,*

$$\forall x, \forall y : (1 + |x + y|)^{-\lambda} \leq (1 + |x|)^{|\lambda|} (1 + |y|)^{-\lambda}.$$

Since: $|\Delta_j(u)| < c 2^{-js} (1 + 2^j |x - x_0|)^{-s'}$.
We get:

$$\begin{array}{rl} |\Delta_j(u')(t)| \leq & 2^j 2^{-js} \int (1 + 2^j |t - 2^j x - x_0|)^{-s'} E(x) dx \\ \leq & 2^{-j(s-1)} (1 + 2^j |t - x_0|)^{-s'} \\ & \int (1 + |x|)^{-s'} E(x) dx. \quad (5) \end{array}$$

And one sees that:

$$u \in C_{x_0}^{s,s'} \Longrightarrow u' \in C_{x_0}^{s-1,s'}.$$

2.3. Wavelets, 2-microlocal analysis and connection to Hölder regularity.

We have already mentionned the similarity between an LPA and a wavelet transform: While a discrete wavelet transform can be seen as the discretisation of a continuous W.T., both in time and scale, an LPA discretizes only the scales axis.

It is therefore not a surprise that the 2-microlocal spaces can be characterized by simple conditions on continuous a discrete wavelet coefficients.

Notations:
If the $\psi_{j,k} = 2^{j/2}\psi(2^j x - k)$ form an orthonormal basis of $L^2(\mathbb{R})$, with ψ in the Schwartz class, define the discrete wavelet coefficients of f by

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

If Ψ is an admissible analysing wavelet, define the continuous wavelet transform of a function f by:

$$W_f(a,b) = \frac{1}{a}\int f(x)\Psi\left(\frac{x-b}{a}\right)dx$$

Theorem 2.1. *[8]*

Let $u \in \mathcal{S}(\mathbb{R}), \{\psi_{j,k}\}$ an orthonormal wavelet basis and Ψ an admissible wavelet. Both analyzing functions ψ and Ψ are supposed to be regular enough and with sufficient vanishing moments. The following conditions are equivalent:

1. $u \in C^{s,s'}_{x_0}$
2. $|c_{j,k}| < c2^{-js}(1 + |k - 2^j x_0|)^{s'}, \ \forall j,k$
3. $\forall a > 0, |b - x_0| < 1$:

$$|W_u(a,b)| \quad \leq \quad ca^s\left(1 + \frac{|b-x_0|}{a}\right)^{-s'}$$

Theorem 2.2. *[8]* $\forall x_0 \in \mathbb{R}, \forall s > 0$*:*

- $C^s_{x_0} \subset C^{s,-s}_{x_0}$
- $C^{s,s'}_{x_0} \subset C^s_{x_0}, \forall s + s' > 0$

2.4. Discussion.

As was mentionned in section 2.1, the Hölder exponent has been extremely useful in many applications in signal processing. However there are situations when both α and α_l are insensitive to what one would like to consider as irregularities. This can be misleading in applications such as classification or segmentation. Perhaps the simplest example where this occurs is that of the chirp $|x|^\gamma sin(\frac{1}{|x|^\delta})$ for which $\alpha = \gamma$ and $\alpha_l = \frac{\gamma}{1+\delta}$. Thus α is totally blind to the rapid oscillation in the signal around 0, and gives the same result as for the signal $|x|^\gamma$ (see figures 1 and 2). α_l on the other hand mixes γ and δ in such a way that two very different chirps can lead to the same exponent. Obviously, since there are two free parameters in a chirp, a richer characterization is needed, such as the one provided by the 2-microlocal analysis. Indeed, it is easy to show that a (γ, δ) chirp belongs to all 2-microlocal spaces $C^{s,s'}_0$ such that $s' < \frac{\alpha - s}{1+\beta}$.

While chirps are expected to be present in gravitational waves, they form a rather restricted class of oscillating singularities. Moreover, it would seem that detecting a chirp, as long as one has a reasonable PSNR, should be much easier than evaluating the $C^{s,s'}_0$ to which the signal belongs. The situation changes if one considers, instead of an isolated singularity, a finite sum of chirps (figure 3), or even a signal containing almost everywhere oscillating singularities (figure 4). This last signal is a lacunary wavelet series as described in [17]. The most noticeable fact is that it is not at all obvious from visual inspection that oscillating singularities are present. In other words, if a signal contains a large number of such singularities, they are not easily detected. In particular, a real world signal may well contain lots of "chirps", which are not detected by classical techniques. The signal on figure 6, for instance, is a lumping of two lacunary wavelet series with the same pointwise Hölder exponent but different oscillating behaviour. The segmentation of such a signal requires a more refined analysis such as 2-microlocal analysis.

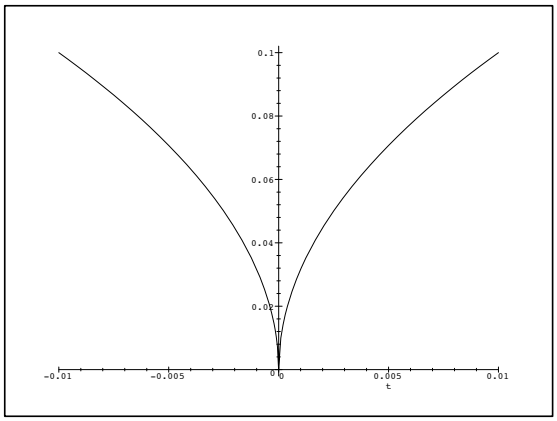

Figure 1: Singular signal with exponent 0.5

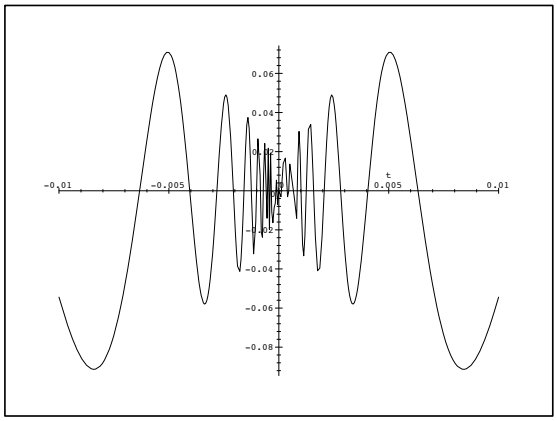

Figure 2: Singular signal with exponent 0.5 with oscillatory behaviour

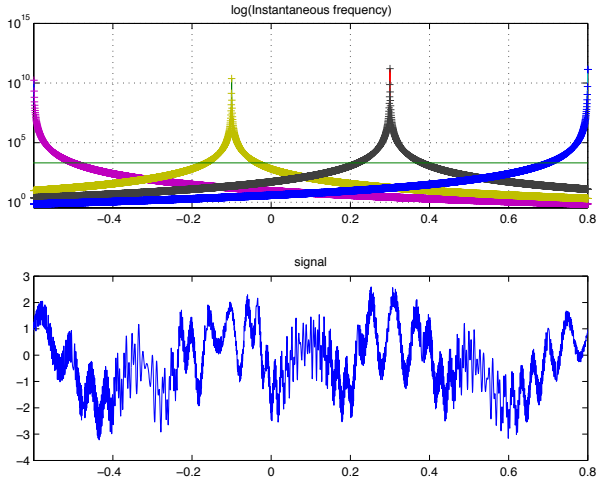

Figure 3: Finite sum of chirps and corresponding instantaneous frequencies

3. 2-MICROLOCAL FRONTIER ANALYSIS

3.1. Definition

Following the notation in [12], we will call $\Sigma(f, x_0)$ the 2-microlocal domain of a function f at x_0: $\Sigma(f, x_0)$ is the set of all couples (s, s') for which f belongs to $C^{s,s'}_{x_0}$. This set is convex and the following relation hold:

$$f \in C^{s,s'}_{x_0} \Longrightarrow C^{s-\epsilon,s'+\epsilon}_{x_0}, \forall \epsilon > 0.$$

It is thus completely defined by its frontier $\Gamma(f, x_0) = \partial\Sigma(f, x_0)$:

Definition 3.1. *2-microlocal frontier parametrization*
Let $f : \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned} S(\sigma, x) &= \sup\left\{s : f \in C^{s,\sigma-s}_x\right\} \\ S'(\sigma, x) &= \inf\left\{s' : f \in C^{\sigma-s',s'}_x\right\} \end{aligned}$$

The 2-microlocal frontier is the set of points

$$\begin{aligned} \Gamma(f, x_0) &= \{(S(\sigma), \sigma - S(\sigma))\} \\ &= \{(\sigma - S'(\sigma), S'(\sigma))\} \end{aligned}$$

We will often call $S(\sigma)$ or $S'(\sigma)$ the 2-microlocal fontier of f, because both completely define $\Gamma(f, x_0)$, and because this induces no ambiguity.

The following properties of S and S' are easy to check:
The function $S(., x_0)$ is decreasing and concave.
The function $S'(., x_0)$ is increasing and convex. Moreover, one has, for all positive τ:

$$S'_{x_0}(\sigma + \tau) \geq S'_{x_0}(\sigma) + \tau$$

See figure 7 for an example of a 2-microlocal frontier.

3.2. Connection with various regularity exponents

I
The following properties hold:

- The local Hölder exponent is given by the relation:
$$\alpha_l(x_0) = \sigma_0$$
where σ_0 is the unique value for which
$$S(\sigma_0, x_0) = \sigma_0$$
- Let $a(x_0) = \sup S(\epsilon), \epsilon > 0$
If $a(x_0) > 0$ then :
$$\alpha(x_0) = a(x_0)$$

As was announced in section 2.1, the 2-microlocal analysis thus allows to recover both the pointwise and local Hölder exponents. Of course, the 2-microlocal frontier contains a much richer information. For instance, the value the chirp exponents can be read from $S(\sigma)$: Recall that the chirp exponents of a function at one point are defined as: [8]

Definition 3.2. *Let* $F^{(-n)}$ *a n-th order primitive of* F. F *is called a* (h, β)*-type chirp at* x_0 *if*

$$\forall n \in \mathbb{N} : F^{(-n)} \in C^{h+n(1+\beta)}_{x_0}.$$

One can show that the chirp exponents of f at x_0 are

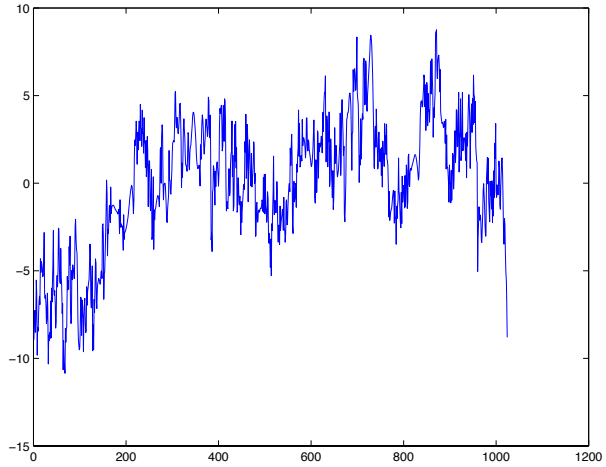

Figure 4: Singular signal with almost everywhere oscillatory behaviour

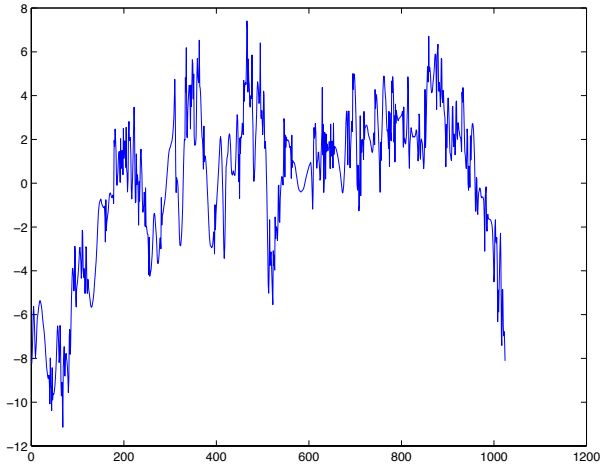

Figure 5: Signal with local regularity 0.1 and almost everywhere pointwise regularity 0.12

(h, β) if and only if the boundary of the 2-microlocal domain of f at x_0 is above the half-line starting at $(s, s') = (h, -h)$ of slope $-(\beta+1)/\beta$, and if this property no more holds for larger values of h or β (see figure 8).

3.3. Frontier prescription at one point

One may wonder what is the most general form that the 2-microlocal frontier of a signal can assume. The answer is given in [12] and is the following:

Theorem 3.1. *Any decreasing concave function $S(\sigma)$ can be the 2-micolocal frontier of a signal at one point.*

One may in fact construct such a signal via its wavelet coefficients. Suppose $S(\sigma)$ is not a line. Let

$$c_{j,k} = 2^{-j/2} \inf_{\sigma} 2^{-j\sigma} (2^{-j} + |2^{-j}k - x_0|)^{S(\sigma)-\sigma}$$

and

$$f(x) = \sum_{j,k} c_{j,k} \psi_{j,k}(x)$$

then, the 2-microlocal frontier of the function f at 0 is $S(\sigma)$.

The proof of this result uses the following observation:
Using a continuous wavelet transform instead of the orthonormal wavelet coefficients, equation (3.3) becomes, with $x_o = 0$:

$$W(a,b) = \inf_{\sigma} a^{\sigma} (a+b)^{S(\sigma)-\sigma}.$$

Applying the following coordinate change:

$$\begin{cases} a = e^{u(1+v)} \\ b = e^{u}(1 - e^{uv}) \end{cases}$$

$v > 0,\ u < 0,$
(3.3) becomes

$$\begin{aligned} T(u,v) &= e^{u \ \inf_{\sigma}[\sigma v + S(\sigma)]} \\ &= e^{-uS^{\star}(-v)} \end{aligned}$$

where $S^{\star}(.)$ is the Legendre-Fenchel transform of S:

$$S^{\star}(\sigma^{\star}) = \sup_{\sigma} [\sigma^{\star}\sigma - S(\sigma)]$$

The end of the proof mainly uses the inversion formula of the Legendre transform of concave functions.

3.4. A new characterization of the 2-microlocal frontier using the Legendre Transform

As was already suggested in [12], the Legendre transform of S seems to be an important tool in the study of the 2-microlocal domain. In this section we give a result in that further support this view. In fact, we will rather use $S'()$ as a parametrization of the frontier.

Here is the main result of this section. Its interpretation is that, rouhly speaking, the knowledge of the power law behaviour of $T(a,b)$ along curves defined by $b = x_o + a^{\alpha}$ is sufficient to obtain the whole 2-microlocal domain. More precisely:

Theorem 3.2. *Let :*

$$\tau(\alpha) = \sup\{\gamma : |T(a, \pm a^{\alpha} + x_0)| < c a^{\gamma} \quad \forall\ a < 1, \alpha \in]0,1[\}$$

$$\tau(1) = \sup\{\gamma : |T(a, b + x_0)| < c a^{\gamma} \quad \forall\ 0 \le |b| \le a < 1\}$$

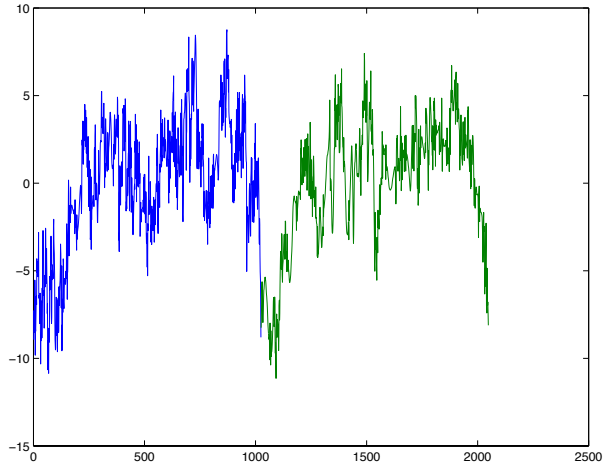

Figure 6: Lumping of lacunary wavelet series

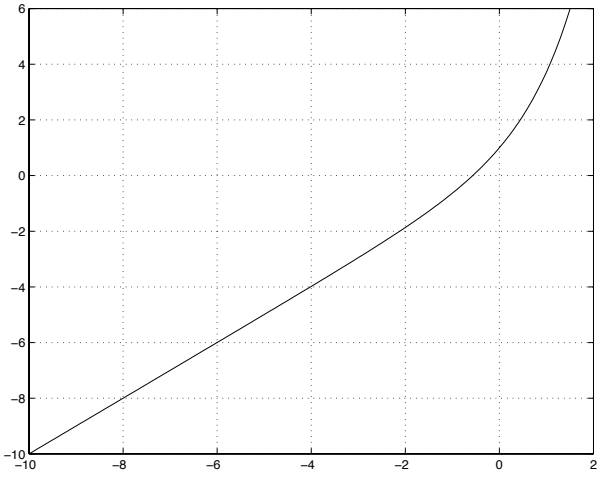

Figure 7: An admissible $S'(.)$

$$u(\beta) = \beta\tau\left(\frac{1}{\beta}\right)$$

Then, the 2-microlocal frontier of f at x_0 is given by:

$$\boxed{S'(\sigma) = u^*(\sigma)}$$

Proof. For sake of simplicity, we will take $x_0 = 0$. We will mainly use the following lemma:

Lemma 3.1. *Let σ et s' be two real numbers. The relations (6) et (7) are equivalent*

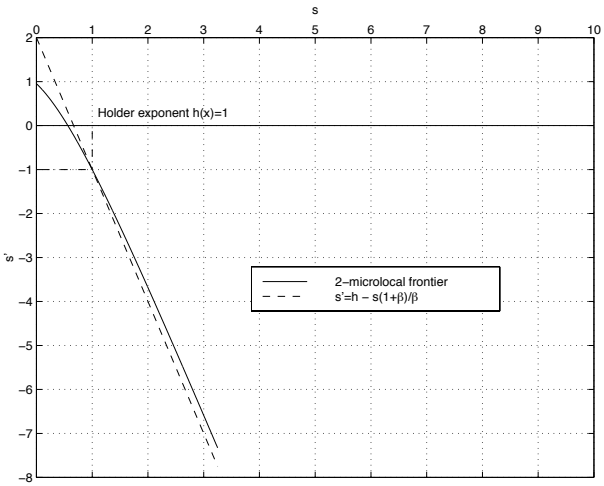

Figure 8: Relation between 2-microlocal domain and chirps exponents

$$\exists c_1(\sigma) > 0 \quad : \quad |T(a,b)| \leq c_1(\sigma)\, a^{\sigma}\, (a+|b|)^{-s'} \quad (6)$$
$$\forall\ (a,b) : (a+|b|) < 1, a > 0$$

$$\exists\ c_2(\sigma) > 0 \quad : \quad (7)$$
$$|T(a,a^{\alpha})| \leq c_2(\sigma)\, a^{(\sigma-\alpha s')}$$
$$\forall \alpha \in [0,1], \forall\ |a+a^{\alpha}| < 1$$

$$|T(a,b)| \leq c_2(\sigma)\, a^{(\sigma-s')}$$
$$\forall\ (a+|b|) < 1 : a > b$$

Proof. First, one has:

$$x^{\alpha} \leq x + x^{\alpha} \leq 2x^{\alpha} \quad (8)$$
$$\forall \alpha \in [0,1],\ \forall x \in [0,1]$$

applying this relation respectively to $(a+a^{\alpha})$ and to $(|b|+|b|^{1/\alpha})$ when $\alpha \in [0,1]$ and $\alpha \in [1,+\infty]$ gives the result.

□

We have:

$$|T(a,b)| < c a^{\sigma} (a+|b|)^{-S'(\sigma)-\varepsilon}$$
$$\forall \varepsilon > 0, \forall \sigma, \forall (a+|b|) < 1. \quad (9)$$

this implies that: $\forall \alpha \in]0,1]$

$$\begin{aligned} |T(a,a^\alpha)| &< ca^{\sigma-\alpha S'(\sigma)-\varepsilon} \\ &< ca^{\alpha\left(\frac{\sigma}{\alpha}-S'(\sigma)-\frac{\varepsilon}{\alpha}\right)} \end{aligned}$$

As $S'(.)$ is convex,

$$S'^{**} = S'$$

and

$$S'^*(\sigma^*) = \sup_\sigma [\sigma\sigma^* - S'(\sigma)]$$

Thus, $\forall\sigma^*, \forall\eta > 0, \exists\sigma_\eta$:

$$\begin{array}{rlll} & S'^*(\sigma^*) & < & \sigma_\eta\sigma^* - S'(\sigma_\eta) + \eta \\ \Longrightarrow & |T(a,a^\alpha)| & < & ca^{\alpha\left(S'^*\left(\frac{1}{\sigma}\right)-\frac{\varepsilon}{\alpha}\ -\eta\right)} \\ \Longrightarrow & \tau(\alpha) & \geq & \alpha S'^*\left(\frac{1}{\alpha}\right) \qquad \forall\alpha \in]0,1[\\ \Longrightarrow & S'^*(\beta) & \leq & u(\beta) \qquad \forall\beta > 1 \end{array}$$

Using the Legendre Transform inversion formula, this in turn implies that: forall β:

$$\underline{S'(\sigma) \geq u^*(\sigma)}$$

Conversely, we have: $\forall\varepsilon > 0,\ \forall\alpha$:

$$\begin{array}{rlll} & |T(a,a^\alpha)| & < & ca^{\tau(\alpha)-\varepsilon} \\ \Longrightarrow & |T(a,a^\alpha)| & < & ca^{\alpha\left(\frac{1}{\alpha}\tau(\alpha)-\frac{\varepsilon}{\alpha}\right)} \\ & & < & ca^{\alpha\left(u\left(\frac{1}{\alpha}\right)-\frac{\varepsilon}{\alpha}\right)} \end{array}$$

But: $u\left(\frac{1}{\alpha}\right) \geq u^{**}\left(\frac{1}{\alpha}\right),\ \forall\alpha$.
Thus:

$$|T(a,a^\alpha)| < ca^{\alpha\left(u^{**}\left(\frac{1}{\alpha}\right)-\frac{\varepsilon}{\alpha}\right)} \qquad \forall\alpha, \forall\varepsilon$$

Now:

$$u^{**}\left(\frac{1}{\alpha}\right) = \sup_\sigma \left(\frac{\sigma}{\alpha} - u^*(\sigma)\right)$$

This implies that:

$$\begin{array}{rlll} & |T(a,a^\alpha)| & < & ca^{\alpha\left(\frac{\sigma}{\alpha}-u^*(\sigma)-\frac{\varepsilon}{\alpha}\right)} \qquad \forall\sigma, \forall\alpha, \forall\varepsilon \\ \Longrightarrow & |T(a,a^\alpha)| & < & ca^{\sigma-\alpha u^*(\sigma)-\varepsilon} \qquad \forall\sigma, \forall\alpha, \forall\varepsilon \\ \Longrightarrow & |T(a,b)| & < & ca^\sigma (a+|b|)^{-u^*(\sigma)-\varepsilon} \end{array}$$

And this finally leads to: $\forall\sigma$

$$\underline{S'(\sigma) \leq u^*(\sigma)}$$

□

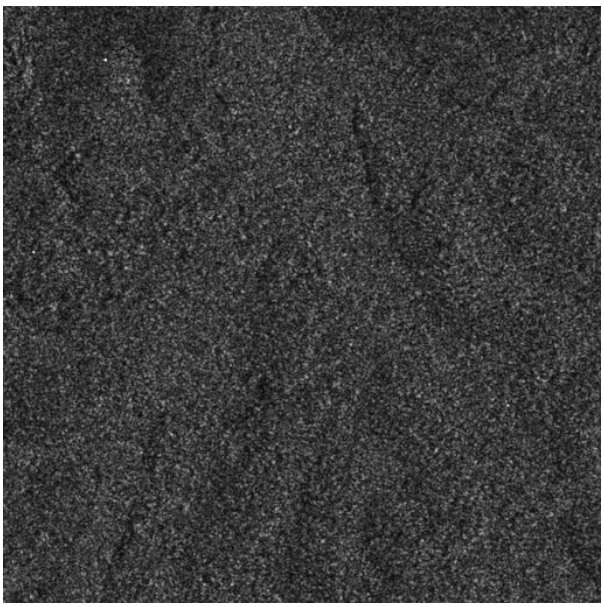

Figure 9: Original SAR image

4. APPLICATION TO IMAGE ANALYSIS

4.1. Basics of Multifractal analysis

Multifractal analysis (MA) deals with the description of the singularity structure of signals (which can be measures [18, 19, 20, 21, 22], functions [23] or capacities [24]), both in a local and a global way. The local information is given by the Hölder exponent at each point, while the global information is captured through a characterization of the geometrical or statistical distribution of the occurring Hölder exponents, called *multifractal spectrum*. Such an analysis is useful when one deals with very irregular signals (such as for instance computer data traffic traces or radar images), and when the singularity structure has practical consequences [25].

Depending on the notion of Hölder exponent in use, there are several MA of a signal. In addition, two spectra are usually defined, as explained hereafter.

4.1.1. Hausdorff spectrum

The geometrical characterization of the singularities distribution is obtained through the following function :

$$f_h(\alpha) = dim_h(E_\alpha)$$

where E_α is the set of points having Hölder exponent α, and dim_h denotes Hausdorff dimension. The Hausdorff spectrum $f_h(\alpha)$ thus gives an indication of

Figure 10: Multifractal image denoising, with shift parameter 0.3

Figure 11: Multifractal image denoising, with shift parameter 0.5

how much the singularity α is spread out over the signal support.

4.1.2. Large deviation spectrum

For simplicity we consider a signal $X(t)$ defined on [0,1] which is nowhere differentiable[1] and define the analysis with respect to the dyadic intervals:

$$I_n^k = [k2^{-n}, (k+1)2^{-n}[,\ k = 0, \dots 2^n - 1, n \in \mathbb{N}$$

Define the *coarse Hölder exponents* through

$$\alpha_n^k = -\frac{1}{n} \log \left| X((k+1)2^{-n}) - X(k2^{-n}) \right|$$

where all logarithms are taken to the base 2 and where $\log 0 := -\infty$.

The large deviation spectrum f_g measures, loosely speaking, how "fast" the probability of observing a coarse Hölder exponent different from the expected value tends to zero as the resolution tends to ∞. More precisely, f_g is related to the rate function appearing in the large deviation analysis of such quantities. A heuristic explanation is the following: assume we want to assess how much the signal T_n may vary in a (small) time interval I_n^k of duration 2^{-n}, with respect to n. The strength of the variation is measured by α_n^k, defined as: $T_n \propto |I_n^k|^{\alpha_n^k} = 2^{-n\alpha_n^k}$. If $\alpha_n^k = 1$, the signal varies smoothly with respect to the scale of measurement ε.

[1]This way we will not have to deal with polynomial trends.

It is thus of interest to evaluate the distribution of the α_n^k's when k is picked randomly from $0, \dots, 2^{n-1}$ and n is large. This will allow to characterize the signal in terms of:

- the highest degree of burstiness (resp. sparseness) occurring in the signal,
- the probability of hitting a given burstiness when measuring the signal in a "small" time interval picked randomly.

In this view, we set:

$$f_g(\alpha) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log N_n^\varepsilon(\alpha)}{n}, \tag{10}$$

where

$$N_n^\varepsilon(\alpha) = \#\{\alpha_n^k / |\alpha_n^k - \alpha| < \varepsilon\}.$$

The large deviation spectrum f_g describes the distribution of the local singularities, since the number of dyadic intervals of size 2^{-n} with coarse Hölder exponent $\simeq \alpha$ varies roughly as $2^{nf_g(\alpha)}$ for large n. Equivalently:

$$P_n(\alpha_n^k \approx \alpha) \propto 2^{-n(1-f_g(\alpha))}$$

where the probability is related to a random choice of k uniformly in $\{0, \dots, 2^n - 1\}$, i.e. P_n is the uniform distribution on the set of all dyadic intervals I_n^k of size 2^{-n}.

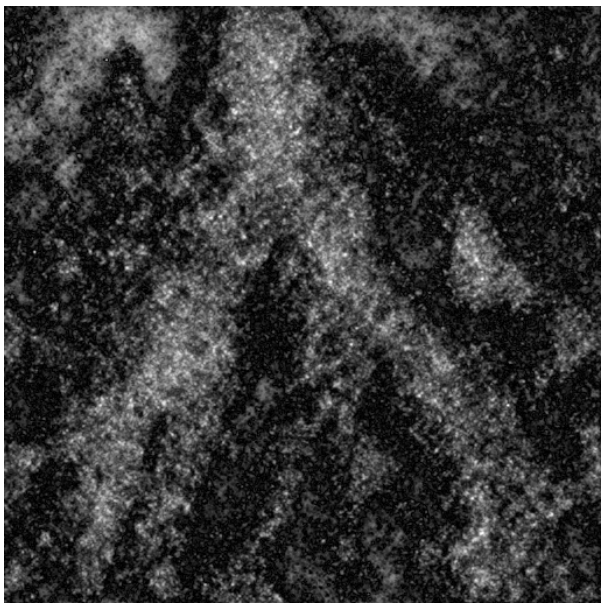

Figure 12: Multifractal image denoising, with shift parameter 0.7

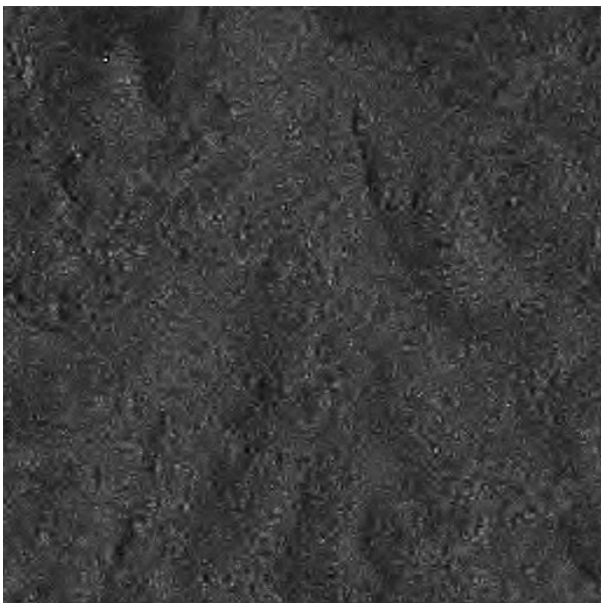

Figure 13: Denoising by wavelet coefficient thresholding

Having defined f_h and f_g, a natural question is to inquire about how they are related. This is a topic adressed by the multifractal formalism, which seeks conditions under which $f_h = f_g$. The interest of this is that f_h is typically very involved to compute theoretically, and almost impossible to estimate numerically. The evaluation of f_g is in general significantly simpler, in particular when the machinery of the large deviation theory can be used, for instance through Ellis theorem.

4.2. Recalls on the multifractal analysis of images

The multifractal approach to image analysis makes the following fundamental assumption:

$$f_h = f_g$$

In addition to providing a way to evaluate f_h, this assumption also allows to relate the geometrical and probabilistic interpretation of e.g. edge points, as explained below.

The multifractal analysis of an image consists in computing its multifractal spectrum, and then classifying/processing each point x according to its $(\alpha(x), f(\alpha(x)))$ value, both in a geometrical and a probabilistic fashion.

The value of α gives a local information pertaining to the regularity of the considered point : edge points will have exponents different from those of points in smooth regions or belonging to textured zones. However, the sole knowledge of α is not sufficient to decide, for instance, that a point belongs to a contour, for at least two reasons : first, the exponents are not invariant under non linear transformations of the image. This means that no absolute value of α can be attributed once and for all to edge points. The second reason is more subtle, and is in fact the main motivation for the use of multifractal analysis in image processing : it is based on the remark that a local information can not suffice to characterize an edge point. Indeed, if the set of candidate edge points is "too complex", or if the number of such points is "too large", the visual system will have a tendancy to see textures rather than edges. This is illustrated on figure 4.2 : the points on the black lines have the same exponent in the left and the right image. However, while one clearly sees three "contours" on the left image, most people would probably think of the right image as a "textured" zone. This is why a global information is needed in order to decide the nature of a point. More precisely, in our setting, an edge point is characterized by an exponent α (i.e. a local information) such that :

- $f_h(\alpha) = 1$, because a smooth contour fills the space like a line.
- $f_g(\alpha) = 1$, because, when choosing at random a pixel of size 2^{-n} in a square image of side 1, the

probability of hitting a smooth contour decreases as $2^{-n} = 2^{-f_g(\alpha)n}$.

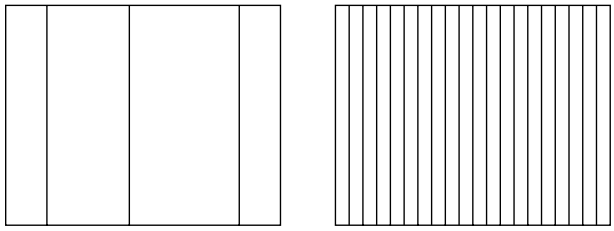

Figure 14: Three edges, a texture.

In fact, we may define the *type* of a point through its associated $f(\alpha)$ value: For $t \in [0,2]$, we say that x is a point of type t if $f_h(\alpha(x)) = t$.

More on this topic can be found in [15]

4.3. Multifractal Denoising

4.3.1. Principle

The basic idea of multifractal denoising is best explained on a simple example. The aim is to get rid of "insignificant" irregularities while keeping "meaningful" singularities. Moreover, after denoising, "most" points should lie in smooth regions. On figure 15, all points have the same regularity $\alpha = 0.5$ except 0.5 where a step occurs ($\alpha = 0$). In term of multifractal spectrum, $f_h(0.5) = 1$, $f_h(0) = 0$ and $f_h(\alpha) = -\infty$ for $\alpha \notin \{0, 0.5\}$. We want to find a simple method to obtain a shifted spectrum $\tilde{f}_h$ so that the maximum of $\tilde{f}_h$ is located at $\alpha = 1 + \epsilon$, $\epsilon > 0$. This will imply that the transformed signal is almost everywhere (*a.e.*) smooth while preserving the "rare" event "step at 0" (figure 16). More formally, our assumption regarding the presence of generic non significant singularities is modeled as:

$$\exists!\, \alpha_0 < 2 \,:\, f_h(\alpha_0) = 2 \tag{11}$$

We then apply a transformation O so that the spectrum $\tilde{f}_h$ of the modified image is

$$\tilde{f}_h(\alpha) = f_h(\alpha + \alpha_0 - \epsilon - 2) \tag{12}$$

The resulting image will be such that "most" points will have Hölder exponents a little above 2 and will consequently be *a.e.* smooth. On the other hand, since the shape of the spectrum is preserved by O, the relative strength of the singularities is not modified, i.e. salient visual features remain unchanged. In particular, such a procedure should not induce, in principle, any loss of information. We describe in the following section a practical method for deriving O.

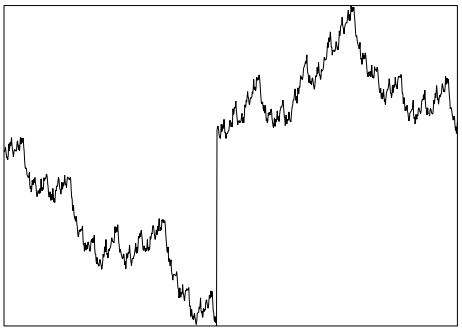

Figure 15: A signal which Hölder exponent is 0.5 every where except in 0.5 where it is 0.

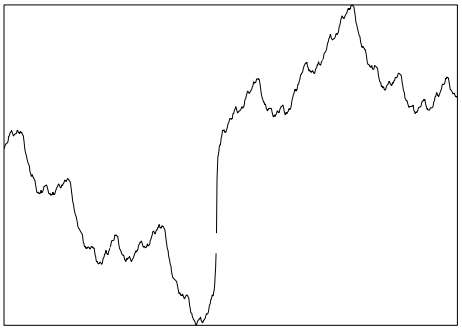

Figure 16: The signal in figure 15 which multifractal spectrum has been shifted by 0.5

4.3.2. Operator design

Recall that, loosely speaking, if f has Hölder exponent α at x_0, the $(c_{j,k})$ in $D_{x_0} = \{(j,k) : x_0 \in supp(\psi_{j,k})\}$ decrease as $O(2^{-j\alpha})$ when j tends to infinity [8]. In order to modify the Hölder regularity, a natural idea is to act upon the wavelet coefficients. The simplest way to increase the regularity from α to α' is to multiply c_k^j by $2^{-j(\Delta\alpha)}$ (where $\Delta\alpha = \alpha' - \alpha$). However, such a naive approach does not insure that transformed signal will have Hölder exponent α', because of possible oscillating behaviour. To obtain this result, it is necessary to perform the 2-microlocal analysis of the signal, and in particular to determine the function $S(\sigma)$ at each point.

Let $(\nabla_{x_0}^{\theta})^{-1}$ denote the operator which consists in

multiplying each $c_{j,k}$ in D_{x_0} by $2^{-j\theta}$. Let:

$$f^\theta = (\nabla^\theta_{x_0})^{-1}(f)$$

By (2.1), one has:

$$\begin{aligned} & S'_{f^\theta}(\sigma, x_0) = S'_f(\sigma + \theta, x_0) \\ \Rightarrow \quad & S_{f^\theta}(\sigma, x_0) = S_f(\sigma - \theta, x_0) + \theta \end{aligned}$$

and the Hölder exponent of f^θ at x_0 is:

$$\alpha(f^\theta, x_0) = S_{f^\theta}(0, x_0) = S_f(-\theta, x_0) + \theta$$

As $S_f(., x_0)$ is continuous, one can derive the following result:

Proposition 4.1.

$$\forall \alpha_\star > \alpha : \exists ! \theta_\star : f^{\theta_\star} \in C^{\alpha_\star}_{x_0}$$

This result means that, for any $\Delta\alpha > 0$, there is a unique θ such that the operator $(\nabla^\theta_{x_0})^{-1}$ increases the pointwise regularity at x_0 by $\Delta\alpha$.

We show on figures 9 to 13 an application of this technique to the denoising of a SAR image. The original image is on figure 9. Figures 10 to 12 show multifractal spectrum shiftings with increasing parameter and figure 13 shows for comparison a denoising using the classical wavelet schrinkage method. One sees in particular how the river appears more and more clearly in figures 10 to 12.

5. REFERENCES

[1] K. Daoudi and J. Lévy Véhel, "Speech modeling based on local regularity analysis," in *Proceedings of the IASTED/IEEE Interantional Conference on Signal and Image Processing.*

[2] A. Arneodo, F. Argoul, E. Bacry, and et al., "Wavelet transform of fractals. I. From the transition to chaos to fully developed turbulence. II. Optical wavelet transform of fractal growth phenomena," in *Wavelets and applications (Marseille, 1989)*, vol. 20 of *RMA Res. Notes Appl. Math.*, pp. 286–352, Paris: Masson, 1992.

[3] J. Lévy-Véhel, "Fractal approaches in signal processing," *Fractals*, vol. 3, no. 4, pp. 755–775, 1995. Symposium in Honor of Benoit Mandelbrot (Curaçao, 1995).

[4] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, part 2, pp. 617–643, 1992.

[5] W. L. Hwang and S. Mallat, "Characterization of self-similar multifractals with wavelet maxima," *Appl. Comput. Harmon. Anal.*, vol. 1, no. 4, pp. 316–328, 1994.

[6] P. Mannersalo and I. Norros, "Multifractal analysis of real atm traffic : a first look," in *COST257TD.*

[7] K. Daoudi, J. Lévy Véhel, and Y. Meyer, "Construction of continuous functions with prescribed local regularity," *To appear in Journal of Constructive Approximation*, 1998.

[8] S. Jaffard and Y. Meyer, "Wavelet methods for pointwise regularity and local oscillations of functions," *Mem. Amer. Math. Soc.*, vol. 123, no. 587, pp. x+110, 1996.

[9] P. Gonçalvès and P. Flandrin, "Scaling exponents estimation from time-scale energy distributions," in *IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-92*, (San Francisco (CA)), pp. V.157–V.160, 1992.

[10] P. Abry, P. Gonçalvès, and P. Flandrin, "Wavelets, spectrum analysis and $1/f$ processes," in *Wavelets and Statistics (A. Antoniadis, Ed.)* (A. Antoniadis and G. Oppenheim, eds.), Lectures Notes in Statistics, 1995.

[11] A. Arneodo, E. Bacry, S. Jaffard, and J. F. Muzy, "Oscillating singularities on Cantor sets: a grand-canonical multifractal formalism," *J. Statist. Phys.*, vol. 87, no. 1-2, pp. 179–209, 1997.

[12] B. Guiheneuf, S. Jaffard, and J. Lévy Véhel, "Two results concerning chirps and 2-microlocal exponents prescription," *To appear in Applied and Computational Harmonic Analysis*, 1998.

[13] J.-M. Bony, "Second microlocalization and propagation of singularities for semilinear hyperbolic equations," in *Hyperbolic equations and related topics (Katata/Kyoto, 1984)*, pp. 11–49, Boston, MA: Academic Press, 1986.

[14] K. Falconer, *Fractal geometry.* Chichester: John Wiley & Sons Ltd., 1990. Mathematical foundations and applications.

[15] J. Lévy Véhel, *Fractal Images Encoding and Analysis.* Sprinver Verlag, 1998.

[16] E. W. Hobson, *The theory of functions of a real variable and the theory of Fourier's series. Vol. I.* New York, N.Y.: Dover Publications Inc., 1958.

[17] S. Jaffard, "On lacunary wavelets series," *Preprint*, 1997.

[18] M. Arbeiter and N. Patzschke, "Random self-similar multifractals," *Math. Nachr.*, vol. 181, pp. 5–42, 1996.

[19] G. Brown, G. Michon, and J. Peyrière, "On the multifractal analysis of measures," *J. Statist. Phys.*, vol. 66, no. 3-4, pp. 775–790, 1992.

[20] B. Mandelbrot, "Intermittent turbulence in self similar cascades : divergence of high moments and dimension of the carrier," *J. Fluid. Mech.*, vol. 62, no. 331, 1974.

[21] B. B. Mandelbrot and R. H. Riedi, "Inverse measures, the inversion formula, and discontinuous multifractals," *Adv. in Appl. Math.*, vol. 18, no. 1, pp. 50–58, 1997.

[22] R. Riedi, "An improved multifractal formalism and self-similar measures," *J. Math. Anal. Appl.*, vol. 189, no. 2, pp. 462–490, 1995.

[23] S. Jaffard, "Formalisme multifractal pour les fonctions," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 317, no. 8, pp. 745–750, 1993.

[24] J. Lévy Véhel and R. Vojak, "Multifractal analysis of Choquet capacities," *Adv. in Appl. Math.*, vol. 20, no. 1, pp. 1–43, 1998.

[25] J. Lévy-Véhel, "Fractal approaches in signal processing," *Fractals*, vol. 3, no. 4, pp. 755–775, 1995. Symposium in Honor of Benoit Mandelbrot (Curaçao, 1995).