Multifractal measures: definition, description, synthesis and analysis. A detailed study

Antonio Turiel* and Conrad Pérez-Vicente†
Departament de Física Fonamental
Universitat de Barcelona
Diagonal, 647. 08028 Barcelona
Spain

Abstract

Many complex systems observed in different scientific research areas (ranging from Physics to Biology and Econometry) are found in scale-invariant, self-organized states. For that reason, during the past decades much effort has been done in order to characterize properties as scale-invariance and self-similar behaviours. One of the most commonly observed scale-invariant phenomena is multifractality. Multifractal systems are not only scale invariant, but composed of self-similar, fractal components. In this paper we review some basic concepts about processing multifractal systems of any kind. We will discuss on the advantages of using measures instead of multiaffine functions in order to characterize the properties of such systems. Finally, with the aid of recent techniques for reconstructing multifractals with the only knowledge of the most relevant fractal component, we will introduce new ideas and tools for the synthesis and analysis of multifractal signals.

1 Introduction

If someone ever wanted to coin a single word which summarizes the central paradigm of early age Physics, that word would probably be *linearity*. Indeed, the milestone laws founding Physics rely on linear properties of systems and interactions under study; at this age, Physics talked about linear superposition of forces, about linear combination of velocities, about addition of energies, entropies and other thermodynamical quantities, etc. Linearity is a simplifying requirement which, on one hand, will always provide a good first order approximation of smooth phenomena and, on the other hand, made the problems initially studied by Physics to remain simple enough to enable ready predictions. At the beginning, Physics succeeded to explain nature on the basis of a well approximating, easy to deal with, principle. In this age, quantities were always *additive*. In addition, the systems in study possessed a reduced number of degrees of freedom, and their parts were typically studied as if they were independent, that is, isolated. The independence hypothesis avoided the rise of complicated systems; besides, additivity and independence imply as a consequence that statistics was dominated by *gaussianity*, as the Central Limit Theorem constituted a pitfall difficult to avoid.

*e-mail: turiel@ffn.ub.es
†e-mail: conrad@ffn.ub.es

The progress of Classical Physics to non-linear Physics and to the study of complex systems comes from the deepening of our knowledge on natural systems, allowed by the tools that our first success granted. The description of systems now includes, on one hand, non-linear phenomena (*i.e.*, non-linear interactions, non-linear superpositions, etc) and, on the other hand, large (even infinite) number of degrees of freedom. Systems became non-linear, then complicated, finally complex. The Physics of Complex Systems had been born. New arising quantities provided a better description of natural phenomena, but they were no longer additive, nor were gaussianly distributed. New functional dependencies arose.

One early discovered non-linear behaviour of physical quantities and statistics is power-law behaviour. Power-law quantities depend on a scale parameter (characterizing the size of an observation window or the physical extent of an object under study) as a power, characterized by the so-called *scaling exponent*. A power-law quantity is *scale-invariant*: up to normalization constants it behaves the same, no matter the measure unity we are using to describe it. Scale invariant phenomena do not have any characteristic time or spatial length: they somehow repeat the same at every scale, making impossible to declare that the process started or finished in a particular stage. Physical systems become self-similar, what means fractal, what means complex. At approximately the same time, physicists realized that some classical, deterministic equations, when attaining a scale invariant limit, generate chaos, which is the closest that determinism can be from probability. The detailed description of every particle in such systems became almost impossible and unpractical, as a small variation of the initial conditions will drive the system to a completely different final state. Physicists realized that, in order to characterize chaos, in order to describe nature, they needed new variables, new quantities which could take account of the self-similar structure of systems and furnished new, independent information. Suddenly, the particular values of scaling exponents acquired a great relevance: in fact, they are scale-invariant measures (the value of exponents does not depend on the initial scale of observation) and, with the progress of Fractal Theory, they could be related with topological and functional properties, with the Physics of systems.

The referred evolution from Classical to Non-linear Physics also implied a larger scope for the new methodologies, arriving further than the conventional limits of Physics. Non-linear methods are now commonly applied in Image Processing, Biology and Econometry, among other scientific fields. Non-linear methods, and in particular scale-invariant methodologies, come in hand to describe different manifestations of Nature, no matter the conventional knowledge areas men use to analyze reality. Nowadays, scale-invariant formalisms are becoming an essential tool in the study of complex systems.

The intent of the present paper is twofold, and for that reason it consists of two parts.

In the first part, we will present a unified theoretical review, illustrated with some practical examples, on one of the most powerful scale-invariant methodologies, namely the multifractal formalism. The reader is warned: this first "reviewing" section is biased to the perspective of a physicist and more particularly to the views of the physics of turbulent flows; in that sense, the review is not complete. Besides, as our intent is rather presenting together the key facts in order to construct a coherent dissertation, sometimes we have tried to clarify some theoretical aspects, using new arguments and proofs when required; hence the first section contradicts again the "reviewing" spirit as we introduce some new features,

In the second part, and based on the line of reasoning already introduced, we will present new techniques for the synthesis and analysis of multifractal measures. We will dedicate much space to the construction of the sources field, which is called to occupy a central role in the description and analysis

of multifractal systems. The obtention of the sources field serves to show the power of multifractal formalism and, at the same time, opens new ways for the analysis of complex systems.

PART I:

Definition and description of multifractals

2 Multiscaling. Multiaffine functions

The use of statistical concepts in the mathematical treatment of turbulent flows and other chaotic systems began with the early works of Kolmogorov in 1941 [1]. Kolmogorov's "1941 theory" provides a formalism which partially describes the mechanisms underlying Fully Developed Turbulence (FDT). Although incomplete, this seminal work settled the roots for later descriptions, and some of the contributions of the original theory are still valid.

Many chaotic systems are deterministic but governed by non-linear equations which generate chaos when some control parameters are taken in an appropriate range of values. This is the case, for instance, of turbulent flows. Turbulent flows are described by means of Navier-Stokes equation, which is non-linear and possesses an infinite number of degrees of freedom. Navier-Stokes equation has a control parameter, known as Reynolds number [1]. When Reynolds number is large, the fluid behaves in a chaotic way, the turbulence gets completely developed (that is, it becomes FDT) and some properties of the fluid become universal. Kolmogorov's work showed that, far from cut-off effects (inertial range), the behaviour of FDT fluids is independent from Reynolds number. But later works have shown that many properties observed for FDT flows are also shared by very different systems. As an example, some symmetries (as translational, time and scale invariances and isotropy), which are affine and deterministic for low-Reynolds fluids and other non-chaotic systems, are recovered in the chaotic regime but in a statistical sense [2]. The key point is that no matter how this particular type of chaos has been generated, the final dynamical equilibrium state can be characterized by simple statistical descriptors, which have few to do with the particularities of the original chaos-spawning equation and are rather determined by general statistical considerations.

There are two key hypothesis, of universal character and validated through experience, which were introduced in the treatment of FDT and are also valid for other systems. The first of these hypothesis is scale invariance. Systems organize in a hierarchical way, from larger to smaller structures. The relation between two given parts only depends on the ratio of scales, and it happens the same disregarding the starting scale: a representation of linking relations resembles a pyramid, a cascade (it is common to speak of "cascade relations"). The second important hypothesis is the pass from deterministic symmetries to statistical symmetries. Typical symmetries of systems include translational invariance, isotropy and, naturally, scale invariance.

[1] Reynolds number is a parameter comparing the relative importance of flow velocities with respect to friction, *i.e.*, viscosity

When considering the first hypothesis, it is natural to think about fractals. Fractal sets (more precisely, self-similar fractal sets) are paradigms of scale invariant systems: fractal sets are usually constructed in a self-affine way, as in Figure 1, left: to pass from one scale to the following an affine transformation is performed. Self-affine fractals, however, are a very restrictive class of scale invariant objects; in order to describe natural objects, a statistical view (according to the second hypothesis) must be introduced. This gives rise to statistical fractals, as for instance random self-affine fractals (as the one shown in Figure 1, right): they are also generated by affine transformations, but at every change in scale the transformation to be applied is chosen at random in an appropriate class. Natural systems mostly behave as statistical self-affine objects: parts are not identical to the whole, but behave the same way as the whole, in the sense that statistical quantities computed from the same descriptor but at two different scales can be related through an appropriate affinity transformation.

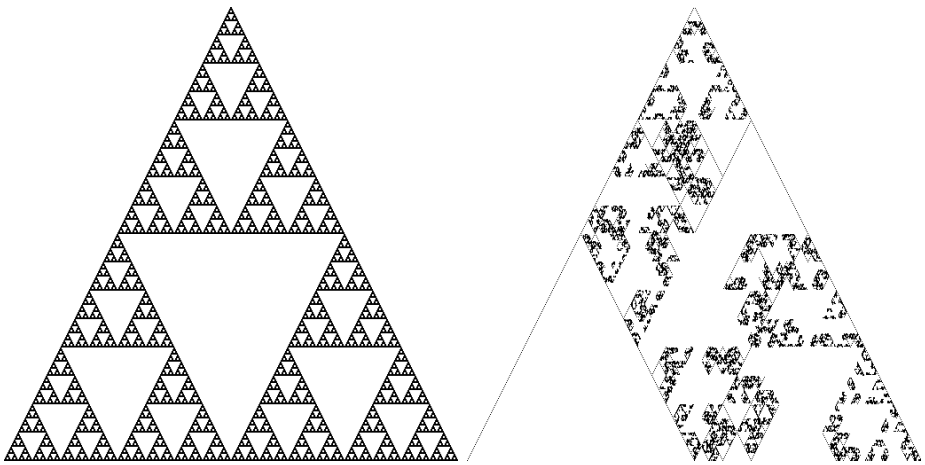

Figure 1: **Left:** Affine Sierpinski's gasket. The figure is generated by recursion, dividing an equilateral triangle in four half-height triangles and removing the central one. **Right:** Probabilistic Sierpinski's gasket. It is generated similarly to the affine one, but at each step the triangle to be removed is decided at random (some edge lines are kept to clarify the structure)

There are two main arguments in support of the validity of a statistical approach to describe natural systems, apart from the good experimental performance. The first explanation states that as chaos-spawning deterministic equations are very sensitive to initial conditions, and those will utterly be affected by quantum effects (which are of random nature), randomness is hence installed across all the scales by simple propagation. The second possible reason concerns the fact that, as evolution is chaotic, we cannot fully characterize a particular state of the fluid because we cannot know it with accuracy: we should take an almost infinite number of measures (as many as degrees of freedom) with infinite accuracy: any deviation, no matter how small, will drive the state to a different evolution. Besides, trying to proceed with such a schedule of measurements is somehow pointless, as we could never reproduce the same state. We will be more interested in providing an uncomplete characterization of the fluid for which parameters change with continuity and are hence useful from the point of view of characterization and stability.

According to the discussion above, we are not interested in obtaining the precise value of the veloc-

ities over a fluid or any other variable in a chaotically-spawn system, but some significant quantities which will be stable under the set of compatible realizations. In this context, Kolmogorov and other defined the Linear Increment (LI) of velocities at a point $\vec{x}$ in the fluid, $\delta_{\vec{r}}v(\vec{x})$, as:

$$\delta_{\vec{r}}v(\vec{x}) \equiv |v(\vec{x}) - v(\vec{x}+\vec{r})| \tag{1}$$

where $v(\vec{x})$ stands for the field of velocities (either moduli or tangential velocities) and $\vec{r}$ represents a displacement vector. Due to scale invariance, the distribution of the LI's can only depend on the size of the displacement $\vec{r}$ as a power-law, because power-laws are the only scale invariant functions[2]. Due to translational invariance, the distribution of LI's at any two points $\vec{x}$ and $\vec{x}'$ coincide[3]. Finally, due to isotropy we can (almost always) disregard the particular direction pointed by the displacement vector $\vec{r}$ and just consider its modulus. A first characterization of the statistics of LI's is given by the second-order moment, which is immediately connected to the two-point correlation of the velocity (in turn, under the assumption of translational invariance, the two-point correlation is connected to the power spectrum, which is known to show a power-law behaviour [2]). Just the second order statistics does not serve to fully characterize the distribution of LI's, but as the LI's are bounded, the probability distributions can be characterized knowing the so-called structure functions,

$$S_p(r) \equiv \langle \delta_r v^p \rangle \tag{2}$$

where the angular brackets $\langle\cdot\rangle$ mean averaging over the ensemble of realizations. Notice that we have removed the basis point $\vec{x}$, making use of the translational invariance; for the same reason (and specially if we assume ergodicity) we can extend the average procedure to different points in the same realization.

Theory predicts that the current lines of a flow in FDT are chaotically arranged so that the LI vary in a sharp, violent way with the scale r, but at the same time the statistics must show a smooth scale-invariant behaviour:

$$S_p(r) \approx \alpha_p \, r^{\tau_p} \tag{3}$$

Such a behaviour as that of eq. (3) is usually referred to as *multiscaling*, *statistical self-similarity* or even *multifractality* (although we reserve that name for the geometrical description to be introduced below). Functions (as the velocity v) verifying eq. (3) are called *multiaffine*. The exponents τ_p are scale-invariant measures (as a change in scale does not change the value of the τ_p's) which inform about the degree of variability and similarity of the ensemble of current lines. In fact, the analysis of τ_p as a function of p reveals the structure of the flow and provides a geometrical interpretation of its scaling properties. It is easy to prove that the simplest functional behaviour, namely a linear one:

$$\tau_p^{\text{linear}} = -(d-D)p + \tau_0 \tag{4}$$

corresponds to the existence of a single dissipative structure. This means that all the energy injections in the d-dimensional flow are based on a single fractal structure of dimension D. The experimental

[2]Power-law behaviour is another characteristic feature of fractal sets.

[3]Under a stronger assumption, ergodicity, averaging over the set of points $\vec{x}$ is equivalent to averaging over different realizations. However, in multifractal systems (as FDT fluids) ergodicity is sometimes locally broken due to intermittency, and should be dealt with care.

evidence on FDT flows and other chaotic system shows that linear τ_p's are rare; the common situation is to have a τ_p which is a non-trivial concave function of p, and in that situation we speak about *anomalous scaling*[4]. When τ_p has a non-trivial curvature, it is interpreted as if in the flow there are many dissipative structures, each one dominating at a different regime captured by the order p [3, 4]. The slope of the tangent to the curve τ_p at a given p, according to eq. (4), is interpreted as the co-dimension of the dissipative structure which dominates at that value of the order of moments.

We are claiming that there exist different, hierarchized structures in the flow which can be recognized by its statistical fingerprint in the structure functions. We can essay to go further and try to explicitly realize this geometrical picture. The idea is that not only the p-moments of the LI's (the structure functions) exhibit power laws in the scale parameter r, buy also the LI's at every particular point, in the way:

$$\delta_r v(\vec{x}) \;\approx\; \omega(\vec{x})\, r^{h_v(\vec{x})} \tag{5}$$

With that scheme in mind, we explain multiscaling, eq. (3), which is statistical scale invariance, with the existence of microscopic scale invariance, in the sense of eq. (5). Besides, recalling the definition of LI, eq (1), the exponent $h(\vec{x})$ can be interpreted as a Hölder exponent, that is, a measure of the regularity of the function $h(\vec{x})$ at the point $\vec{x}$. The microscopic picture provides geometrical and functional interpretations of the structure of the system.

But even if a model for microscopic scale invariance is very appealing, it is not always correct in practice. The first difficulty comes from the fact that LI's are very coarse measures of the variations of velocity in the flow (or another scalar quantity in other multiscaling contexts) and they are easily affected by artifacts and long-range correlations. The later problem is easy to understand: let us suppose that the exponent $h(\vec{x})$ associated to the microscopic structure is rather large, meaning that the function is very regular at that particular point. If due to the presence of some external contribution the field has a long-range smooth correlation, this contribution will be added to the microscopic behaviour, making difficult or impossible to decide the local structure, the local $h(\vec{x})$. Arneodo and collaborators [5] designed a method for filtering these long-range correlations which are commonly found. In order to characterize the local Hölder exponent $h(\vec{x})$ by a simple substraction, they filtered the velocity with an analysis wavelet $\Psi(s)$. The LI's are then generalized to Linear Wavelet Projections (LWP), in the way:

$$\delta_\Psi v(\vec{x}_0, r) \;\equiv\; \int ds\; v(\vec{x} - s\vec{u}_r)\frac{1}{r}\Psi(\frac{s}{r}) \tag{6}$$

where $\vec{u}_r$ stands for a unitary vector pointing in a direction of choice. Notice that the LI is a particular case of LWP; if $\Psi(s) = \delta(s) - \delta(s-1)$ (the so-called "poor man's wavelet") the LWP becomes a LI. The idea of Arneodo et al. was to use wavelets capable to filter a large enough number of integer moments, intending to remove smooth long-range correlations. As it has been reported [5, 6], this method is very efficient not only to extract local exponents, but also to eliminate artifacts in the determination of τ_p (redefining the structure functions in terms of LWP's instead of LI's). Besides, the wavelet projections can be generalized in order to avoid choosing a privileged direction $\vec{u}_r$ [7].

[4]The name "anomalous scaling" has been kept for historical reasons, as Kolmogorov's model was commonly accepted for so long and it predicted linear scaling of τ_p; the first experimental evidences of non-linear dependence of τ_p on p were considered anomalies.

The method of wavelet analysis of multiaffine signals has proven to be an efficient way to determine statistical properties of complex systems, and provides a good basis to analyze their geometry. In fact, the analysis of wavelet projections lies in the roots of the so-called Wavelet Transform Moduli Maxima method, which was suggested by Mallat et al [8] and developed by Arneodo and collaborators in the context of multiaffine/multifractal signals [5, 7]. This method is very powerful in removing some artifacts, as oscillating singularities, but at the cost of loosing resolution. It has other limitations, as for instance the requirement of isolation for the singularities to be detected [9]. Even though the main singularities can be used to describe the statistics, all those limitations do not allow for a precise geometrical determination. We will see that we can overcome those difficulties by the introduction of multifractal measures.

3 Why multifractal measures?

3.1 Definition and geometrical interpretation

We will denote a generic signal by $s(\vec{x})$, where $\vec{x} \in I\!R^d$. The function $s(\vec{x})$ is considered to be of finite variation, a.e. differentiable; this is a mathematical formalization of the fact that physical signals do not explode to infinity (they are locally bounded, hence of finite variation) and they vary smoothly during long periods. Even more, we will assume that $|\nabla s|(\vec{x})$ is locally integrable. We can hence define the (positive) measure μ as:

$$d\mu(\vec{x}) \;\equiv\; d\vec{x}\,|\nabla s|(\vec{x}) \tag{7}$$

that is, for each set A, the measure of this part $\mu(A)$ is given by:

$$\mu(A) \;=\; \int_A d\vec{x}\,|\nabla s|(\vec{x}) \tag{8}$$

According to the definition, this measure is σ-finite although not necessarily finite. The measure μ gives an idea of the local distribution of the gradient of s. It is possible to characterize the behaviour at any particular point $\vec{x}_0$ by evaluating μ over balls centered on $\vec{x}_0$ with different radii r; we will denote these balls by $B_r(\vec{x})$. For a general measure μ, the collection $\{\mu(B_r(\vec{x}_0))\}_{r>0}$ will have a complicated dependence on r. Some scale-invariant systems exhibit *multifractal behaviour*, that is,

$$\mu(B_r(\vec{x}_0)) \;=\; \alpha(\vec{x}_0)\, r^{d+h(\vec{x}_0)} \;+\; o(r^{d+h(\vec{x}_0)}) \tag{9}$$

(see, for instance, [10, 11, 12, 13, 14]) where d is the dimension of the space, introduced for convenience: so that the exponent h is the same for functions of the same degree of smoothness, disregarding the dimensionality of the space. We call μ "multifractal" because eq. (9)[5] allows to split the space into multiple fractal sets. First, at each point $\vec{x}$ the singularity exponent[6] $h(\vec{x})$ is computed. Then, all the points in the space are arranged according to the value of their singularity exponent. The fractal component F_{h_0} is defined as:

[5] In theory, it is eq. (32) which should be used; see Appendix A.

[6] We will call $h(\vec{x})$ "singularity exponent", although sometimes this quantity is positive. It actually represents the degree of singularity (if negative) or regularity (if positive) of ∇s [11].

$$F_{h_0} \equiv \{\vec{x} : h(\vec{x}) = h_0\} \tag{10}$$

The arrangement of points around fractal components is very revealing about the geometrical structure of the image; for instance, the component associated to the smallest singularity exponent (called the *Most Singular Manifold*, MSM) has usually a fractal co-dimension of 1.0 and resembles the edges or contours of the objects present in the case of natural images [11], or relevant dynamical points in other systems [15, 14]. Another reason for the relevance of the multifractal decomposition is that the multifractal hierarchy determines the statistical properties. More precisely, knowing the *singularity spectrum* (the function $D(h) = \dim F_h$, which gives the fractal dimension of the fractal component as a function of the exponent) it is possible to obtain the multiscaling exponents τ_p in eq. (3), in application of Parisi and Frisch's formula [16][7]

3.2 Numerical reasons

When dealing with digitized signals, where distances are discretized, the direct application of eq. (9) to compute the singularity exponents $h(\vec{x})$ gives a very poor performance: when the radius r is a non-integer number of units, some interpolation must be applied. In such cases, it is convenient to consider the projections of the measure μ over an appropriate wavelet Ψ around the point $\vec{x}$ to be studied and at different sizes r. This is completely analogous to what happened in the extension of the definition of LI's to LWP's, although here the reasons are different: we do not want to filter correlations but to interpolate discretized samples. For the measure μ we define the wavelet projections $T_\Psi\mu(\vec{x}_0, r)$ as:

$$T_\Psi\mu(\vec{x}_0, r) \equiv \int d\mu(\vec{x}) \, \frac{1}{r^d} \Psi(\frac{\vec{x}_0 - \vec{x}}{r}) \tag{11}$$

A smooth wavelet has wavelet projections which vary smoothly with the size r; a wavelet supported over a compact set K will give a rise to wavelet projections close to the measures of the r-dilations of K. For a large class of functions Ψ, the wavelet projections provide the same singularity exponents[8] as the measure μ, that is:

$$T_\Psi\mu(\vec{x}_0, r) = \beta(\vec{x}_0) \, r^{h(\vec{x}_0)} + o(r^{h(\vec{x}_0)}) \tag{12}$$

In contrast to what happens with wavelet analysis on multiaffine functions, there is little scientific literature about analysis of multifractal measures. For that reason, we have found convenient to present here some simple theorems about the requirements on a function Ψ in order to have wavelet projections $T_\Psi\mu(\vec{x}_0, r)$ verifying eq. (12); see the discussion in Appendix A. The conclusion of our proofs is that the range of resolved singularities depends on the behaviour of the tail of the analysis wavelet Ψ: the faster Ψ decreases at infinity, the larger singularities can be detected. There is an experimental tradeoff, however: wavelets able to detect a wider range of singularities have poorer numerical and

[7] The reader is warned however that from the theoretical point of view up to three different singularity spectra can be defined, according to some topological or functional features of the system, and they do not necessarily coincide; the reader is referred to [17] for a wider explanation.

[8] The wavelet projections are normalized by a factor r^{-d} and for that reason the exponents appearing in the power law are shifted by $-d$ with respect to those of μ. That normalization of the wavelet is convenient to make the limit $r \to 0$ stable.

spatial accuracy, as can be seen in Figure 2. For the experiences in Figure 2, the wavelets that we used were $\Psi_1(\vec{x}) = (1 + |\vec{x}|^2)^{-1}$, $\Psi_2(\vec{x}) = (1 + |\vec{x}|^2)^{-2}$ and $\Psi_3(\vec{x}) = e^{-|\vec{x}|^2/2}$. According to the theory, Ψ_1 truncates singularities above $h = 0$, Ψ_2 above $h = 1$ and Ψ_3 resolves any singularity. However, the accuracy for the singularities in the corresponding range is better when less singularities are resolved.

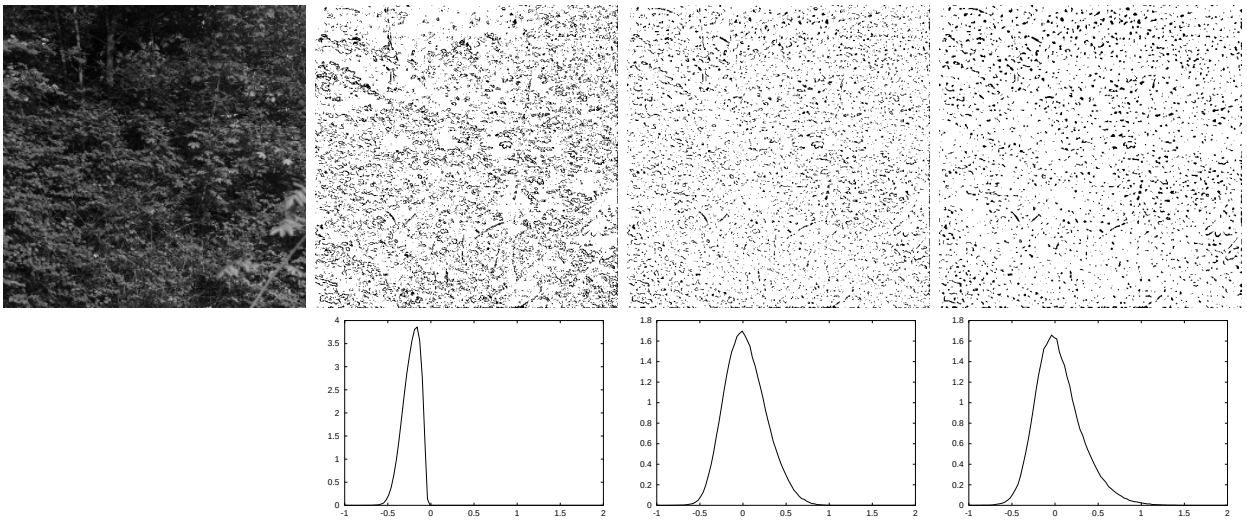

Figure 2: **Top:** (from left to right) Central 512×512 patch in imk01375.imc from van Hateren's database [18] and MSMs for Ψ_1 (estimated $h_\infty = -0.45 \pm 0.1$), Ψ_2 ($h_\infty = -0.40 \pm 0.1$) and Ψ_3; ($h_\infty = -0.40 \pm 0.1$)
Bottom: (below the corresponding MSM) Experimental singularity distributions for Ψ_1, Ψ_2 and Ψ_3.

The theory summarized in Appendix A and exemplified in Figure 2 supports the choice of measures in front of LI's or their generalization, LWP's. Even if in practical grounds we need to perform wavelet projections of the measure which are analogous to LWP's, the requirements on the analysis wavelet Ψ are less constraining when working with measures, specially in what concerns to minimum resolution. For the wavelet analysis of measures we can use positive wavelets, which are well behaved when resolution is reduced. In opposition, wavelets for the analysis of LI's require two or more zero-crossings, introducing a minimum discretized scale of about four points at least (see Figure 3). That is, in order to attain a good accuracy and spatial localization, if measures can be used they should be used in preference.

3.3 Theoretical reasons

From the theoretical point of view, the interest in working with measures rather than with functions is because measures are constructed by accumulating intermittent variables (usually derivatives). A variable is said intermittent if any deviation from the mean value lasts only during a short period or spatial distance. Intermittence is connected with the dynamics of turbulence (the sudden, rare bursts in the value of the variable are associated with microscopic structure, that is, with crossings of the different fractal components) and for that reason the study of intermittent variables is more

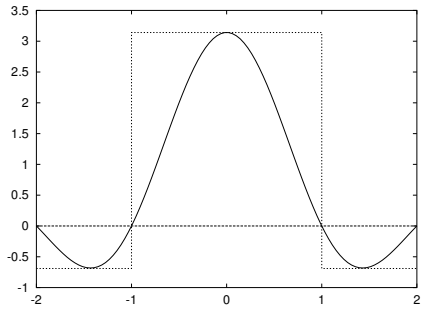

Figure 3: The theoretical continuous curve can only be approximated over a discretized grid (dashed steps) up to a resolution which allows separating positive and negative parts; the zero-crossings correspond to pixel boundaries. For that reason, the number of zero-crossings determines the minimum attainable resolution.

likely to reveal about the structure of flow. There is yet another reason to prefer measures: they are more stable. As the main contributions to measures come from bursts, that is, from microscopic structure, they are less affected by long-range correlations and henceforth numerically they are more stable. Recently the measure-based approach have been used for perform actual geometrical analysis of systems, and not only statistical analysis. It has been shown that, using multifractal measures, it is easy to obtain consistent multifractal decompositions for many different systems, as natural images [11], geophysical atmospheric variables [14] and even econometrical series [15].

4 Bounded multifractals: reconstruction

So far we are interested in physical multifractal systems, that is, we consider the analysis of real signals, not theoretical ones. Physical signals are more constrained than mathematical, theoretical, ideal ones. One of the most important constraints on physical signals is boundness: a physical signal cannot diverge to infinity, as this would imply an infinite energy or mass expense or supply. In a more precise way, a physical signal is of bounded variation: over a given compact set the maximum difference between two values that the function can take is finite (although this difference may be different depending on the particular compact set chosen). As a direct consequence of the definition, finite variation signals have a minimum Hölder exponent of zero (differences are bounded at least by constants). Remark that in the case of a multifractal measure as μ, the singularity exponent coincides with the Hölder exponent of the derivative (when it exists). Then, we can conclude that the minimum possible value for the singularity exponent for measures μ constructed from finite variation signals is -1. Experimental verifications confirm that property [11, 14, 15].

The boundness by below of the exponents $h(\vec{x})$ implies the existence of a fractal component associated to that particular value, to the most singular exponent h_∞. That fractal component receives the name of Most Singular Manifold (MSM; hereafter denoted by F_∞). The MSM plays a fundamental role in the characterization of physical multifractal systems. First, it forces exclusion of some popular models used to represent multifractal/multiaffine functions. Many generative models are constructed

on the basis of an Infinitely Divisible Multiplicative Process (IDMP), as statistical self-similarity implies the existence of an underlying IDMP [19]. From the IDMP it is possible to obtain the multiaffine exponents τ_p [4, 20]; the simplest IDMP's are log-Normal processes. However, log-Normal processes have unbounded singularity exponents and thus cannot represent physical signals except as an approximation for lower order moments. For the same reason, many log-Lévi processes are excluded as well. Then, the simplest IDMP's verifying lower boundness are log-Poisson [21, 3]. One of the appealing characteristics of log-Poisson processes is that the system can be interpreted as a cascade of energy injection starting from the MSM through the less and less singular components in the multifractal: the systems is then hierarchically organized in a pyramidal fashion. In fact, the parameters defining the MSM (namely, its associated singularity exponent h_∞ and its fractal dimension D_∞) characterize completely log-Poisson IDMP's. Can this strong dependency be further exploited?

In [22] it was proposed for the first time that the whole measure μ, and even more, the whole signal s, could be reconstructed just knowing the values of the measure over the MSM only. The reconstruction paradigm was designed under five hypothesis (compatible with known properties for scale-invariant systems): determinism, linearity, translational invariance, scale invariance and compatibility with the known power spectrum. The data which are needed to reconstruct the signal are the essential gradient ∇s_∞, that is, the gradient of the signal over the MSM F_∞,

$$\nabla s_\infty(\vec{x}) \;\equiv\; \nabla s(\vec{x})\, \delta_{F_\infty}(\vec{x}) \tag{13}$$

where $\delta_{F_\infty}(\vec{x})$ stands for the delta-like (distributional) density measure for the MSM F_∞. The reconstruction formula states that the whole signal can be retrieved from the essential gradient using a universal convolution kernel $\vec{g}(\vec{x})$:

$$s(\vec{x}) \;=\; \vec{g} \otimes \nabla s_\infty(\vec{x}) \tag{14}$$

where the symbol $\otimes$ must be understood as a convolution dot-product of vectors and the universal kernel $\vec{g}(\vec{x})$ is given by its Fourier transform $\hat{\vec{g}}(\vec{f})$,

$$\hat{\vec{g}}(\vec{f}) \;=\; i\frac{\vec{f}}{f^2} \tag{15}$$

i stands for the imaginary unit, $\vec{f}$ is the spatial frequency variable which parametrizes Fourier space and $f \equiv |\vec{f}|$. The experimental performance of the reconstruction formula, eq. (14), over real data is very good: see Figure 4 for an example over Lena's image. The quality of the reconstruction depends strongly on the quality in the detection of the MSM. If some important structures which should belong to the MSM are missing, also the reconstruction will lack of them: the reconstruction kernel $\vec{g}$ just diffuses the information out of the MSM, but without creating new geometrical features; in that sense, it is just a functional (in opposition to geometrical) propagator. All the information in the geometry of the scene is contained in the MSM, and the relevant functional information, in the values of the gradient over the MSM, ∇s_∞

Although the reconstruction formula, eq. (14), was first designed for a particular system (namely natural images, see [22]), its validity seems broader, as it has been successfully applied to 1D econometrical series [15] and meteorological 2D signals [14]; thus, it seems that the capability of reconstructing a signal from the MSM is in correspondence with more general principles. The reconstruction kernel

Figure 4: Lena's original images (left), MSM with a coarse quantization (it contains almost 43 % of the points) (middle) and reconstruction (PSNR: 27.54 dB) (right). Observe that errors are local and smoothly propagated, providing good quality under visual inspection.

represents a smooth spatial version of the cascade: energy is propagated across the hierarchy of fractal sets. The relevance of the "reconstructibility" property is twofold. On one hand, it authorizes concentrating the analysis efforts on the MSM, as it truly contains all the information required to describe the system. On the other hand, it settles the roots for a coding scheme which is very compact[9] and which emulates physical properties of the system (information is propagated the same way as energy: from the vertex of the hierarchy). But the reconstruction formula also allows to determine deeper properties on the dynamics of the system, using the techniques we will present now.

PART II:

Multifractal synthesis and analysis

In the following we will explore two different, complementary problems. Although we will not provide a complete answer to neither, we will obtain some valuable information for the understanding of multifractal signals. All the results presented in this section are novel and the roots of forthcoming works. In all instances we will assume that multifractals are reconstructible, which in particular implies that there exists a MSM, that is, that there exists a minimum possible singularity exponent. This restriction excludes many mathematical, theoretical multifractals, although fits well for physical, multiscaling signals, which at the end are the class of objects we are interested in describing.

The first problem is how to construct new multifractal signals having the same multifractal structure as a given one (*multifractal synthesis*). The second problem is how to split a given multifractal into two parts: a canonical representative containing all the multifractal structure and a non-multiscaling

[9]Besides, the MSM can be regarded as the most informative set in the image [14], as was already conjectured in [11].

function which comprises all the functional variability of that particular realization (*multifractal analysis*). Multifractal synthesis is important as we gain knowledge about the possible functional differences giving rise to the same multifractal geometry, thus allowing to define equivalence classes among different signals according to this geometry; we can then discriminate signals with different scaling mechanisms as they lie in different classes. Multifractal analysis, on the contrary, extracts a canonical representative of the equivalence class and explicitly expresses any member of the equivalence class as a simple product of the representative by a non-multiscaling (although scale invariant) function that we will call "sources field". This sources field can be interpreted as the part of the dynamics of the system which cannot be explained looking only at its geometry.

5 Multifractal synthesis

Let us suppose that we have a multifractal signal s with associated multifractal hierarchy $\{F_h\}_h$, where h is bounded by below, $h \geq h_\infty$. We would like to identify the class of functional transformations T such that $T[s]$ is a multifractal signal with the same associated hierarchy. In essence, we look for transformations preserving the fractal structure of each one of the fractal components F_h. The best known type of fractal-preserving functional is the one associated to bi-Lipschitz functions. A *bi-Lipschitz function* $\mathcal{L} : I\!R \to I\!R$ is such that there exists two positive constants $0 < A < B$ such that for any two real numbers a, b the following inequalities are verified:

$$A|a-b| \;\leq\; |\mathcal{L}(a) - \mathcal{L}(b)| \;\leq\; B|a-b| \tag{16}$$

One of the particularities of bi-Lipschitz functions is that they preserve fractal dimensions: the image of a fractal set $F \subset I\!R$ by a bi-Lipschitz function is a fractal of the same dimension [23]. When applied to multiscaling functions (in the sense of eq. (5)) we obtain a stronger statement: it follows trivially combining eq. (16) and eq. (5) that if $s(\vec{x})$ is multiscaling, the function $\mathcal{L}(s(\vec{x}))$ is also multiscaling with exactly the same singularity exponents. So that, bi-Lipschitz functions are multifractal-preserving for multiaffine functions. Reasonably, they should also be multifractal-preserving when working with multifractal measures. We have not however a proof of such a statement; instead, we can prove that multifractal measures are preserved by a more restricted class of bi-Lipschitz functions, namely monotonic smooth functions.

Let μ be a multifractal measure with density given by a gradient, $d\mu(\vec{x}) = d\vec{x}|\nabla s|(\vec{x})$, and $\mathcal{L} : I\!R \to I\!R$ a monotonic function with continuous derivative, then the signal $\mathcal{L}(s)$ is multifractal also, and with the same singularity exponents.

Let us calculate the measure associated to $\mathcal{L}(s)$, that we will denote it by $\mathcal{L}\mu$. It follows that the measure $\mathcal{L}\mu$ of any set $\mathcal{A}$ is:

$$\mathcal{L}\mu(\mathcal{A}) \;=\; \int_{\mathcal{A}} d\vec{x}\, |\nabla \mathcal{L}(s(\vec{x}))| \;=\; \int_{\mathcal{A}} d\vec{x}\, |\mathcal{L}'(s(\vec{x}))|\, |\nabla s|(\vec{x}) \tag{17}$$

If we take an arbitrary point $\vec{x}$ and we consider balls centered on $\vec{x}$ and of radii r small enough, we can find two bounding constants $0 < A < B$ such that $A < |\mathcal{L}'(s(\vec{y}))| < B \;\forall \vec{y} \in B_r(\vec{x})$ because s is of bounded variation and $\mathcal{L}'$ is continuous; we then conclude:

$$A\,\mu(B_r(\vec{x})) \;\leq\; \mathcal{L}\mu(B_r(\vec{x})) \;\leq B\,\mu(B_r(\vec{x})) \tag{18}$$

from which it follows that if $\mathcal{L}\mu$ is multifractal it should possess the same singularity exponents as μ. To finish the proof, it should be realized that the function $s(\vec{x})$ is scale invariant, so it is $\mathcal{L}'(s(\vec{x}))$, hence eq. (18) is valid for an arbitrary radius r, not necessarily small enough. It follows that $\mathcal{L}\mu$ has the same singularities as μ. In Figure 5 we can see an example of the experimental correspondence between the singularities of Lena's image and its transform under a particular type of monotonic transformation commonly used in image processing, namely histogram equalization.

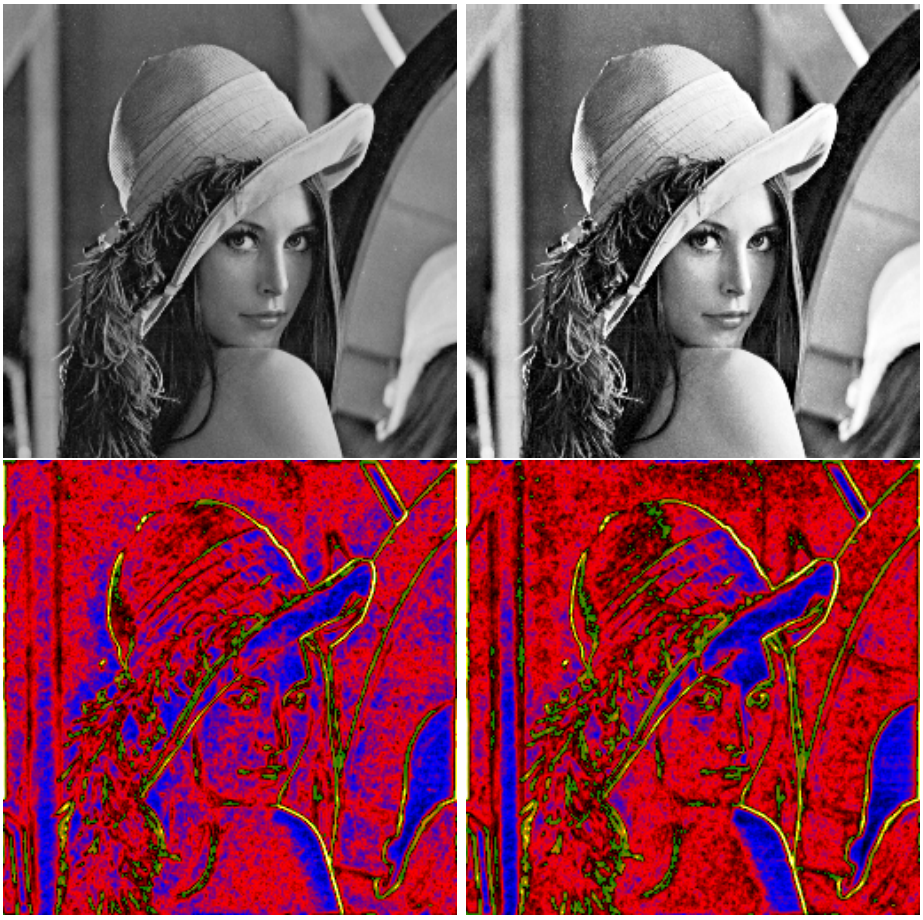

Figure 5: **Top:** Lena's images: Original (left) and after histogram equalization (right). **Bottom:** Singularity exponents associated to both images, from most singular (lighter colors) to less singular (darker colors)

The proof above gives a hint of how multifractal measures should be manipulated in order to construct new multifractal measures associated to the same multiscaling hierarchy. The intuitive idea is to perform a simple integration of a positive function $f(\vec{x})$ with respect to the measure μ; we produce then the measure μ_f, defined by its actuation on an arbitrary set $\mathcal{A}$ as:

$$\mu_f(\mathcal{A}) \;=\; \int_{\mathcal{A}} d\mu(\vec{x})\; f(\vec{x}) \tag{19}$$

If the function $f(\vec{x})$ can be locally bounded by below and above, as in the previous proof, we can conclude that for the measures of balls of radii small enough we will observe a multifractal behaviour with the same local singularity exponents as μ. However, we cannot state that for larger radii the local power-law scaling is preserved; in fact it is easy to design functions f with compact support thus truncating the scaling effects. In analogy to the previous proof, it would be convenient to require the function f to be scale invariant. If we further require that the measure μ_f should be derived from the gradient of an appropriate function s_f, it follows that

$$d\mu_f(\vec{x}) \;\equiv\; d\vec{x}\,|\nabla s|(\vec{x})\; f(\vec{x}) \;=\; d\vec{x}|\nabla s_f|(\vec{x}) \tag{20}$$

The easiest way to verify such a relation is to assume that ∇s and ∇s_f are parallel, that is:

$$\nabla s_f(\vec{x}) \;=\; f(\vec{x})\,\nabla s(\vec{x}) \tag{21}$$

but then locally $f(\vec{x})$ must be a function of $s(\vec{x})$, and then by integration $s_f = \mathcal{L}(s)$ for an appropriate real function $\mathcal{L}$. We conclude that the simplest way to obtain multifractal measures of the same type is by means of appropriate functional transformations. This class of transform is yet a bit restrictive. A more general transform is possible; we do not need to assume that the gradients are parallel, but that they are related by a matrix field $\mathbb{M}_f$, in the way

$$\nabla s_f(\vec{x}) \;=\; \mathbb{M}_f(\vec{x})\,\nabla s(\vec{x}) \tag{22}$$

and where $\parallel \mathbb{M}_f(\vec{x}) \parallel = f(\vec{x})$. Matricial transforms of this kind give rise to new functions s_f which are not necessarily of the form $\mathcal{L}(s)$, and the matrix $\mathbb{M}_f$ informs about the degree of distortion in the mutual functional dependence. We will see in the following that given two equivalent multifractals the matrix relating them can be extracted, in order to analyze the geometry independently of the functional aspects.

6 Multifractal analysis: sources

Given a multifractal measure μ associated to a reconstructible signal s, there is a simple canonical way to construct a reduced signal $s_{_R}$ with the same associated multifractal hierarchy, provided that the MSM is a set of co-dimension 1 and, even more, hypersurface-like[10]. The construction goes as follows: first, we extract the MSM F_∞ for the signal s. As this signal is reconstructible, it suffices to know ∇s_∞ to retrieve s. In order to construct the reduced counterpart $s_{_R}$ we will substitute the actual values of the gradients of s over the MSM (that is, ∇s_∞) by unitary vector gradients, directed in the perpendicular direction to the MSM at each point (that we can determine as this set is hypersurface-like). There is an ambiguity due to the existence of two possible orientations for a hypersurface; we choose to keep the same orientation as the actual gradient ∇s_∞. Then, knowing the

[10]The later statements - for systems of dimension 2 (images, surface signals, etc), those conditions are equivalent to require the MSM to be of dimension 1 and curve-like - could be relaxed under more general requirements to generate the reduced counterpart of s.

oriented MSM (an example is depicted in Figure 6), that is, the set of points belonging to the MSM plus the orientation to be kept, we generate a unitary field perpendicular to the MSM and use it to reconstruct with the reconstruction formula, eq. (14). The resulting signal is the reduced signal $s_{_R}$ (see an example in Figure 6).

Figure 6: **Left:** Oriented MSM, from the MSM represented in Figure 4 . Black represents positive orientation, while gray means negative. **Right:** Reduced signal, generated by the only knowledge of the oriented MSM depicted on the left.

As a matter of fact, the gradients ∇s_∞ change smoothly for displacements inside the MSM. It is so reasonable to think that passing from the original signal to the reduced signal can be verified through a transforming smooth matrix as the ones studied in the previous section. It is an experimental evidence that reduced signals have the same multifractal structure as original ones, confirming the smoothness statement above.

The reduced signal represents a simple guess on the function s; we just introduce information about the geometry (*i.e.*, the points belonging to the MSM) and a small amount of information about the function, that is, the orientation, which cannot be deduced locally as it is a global feature. We would now want to extract the matrix $\mathbb{M}_f$ relating both signals, the original one s and its reduced counterpart $s_{_R}$. The procedure of application in this case is a Radon-Nykodin derivative [24].

We will next label measures according to the function used to produce them; so, the measure associated to the signal s will be μ_s and the measure associated to the reduced signal $s_{_R}$ is $\mu_{s_{_R}}$, in the way:

$$\begin{aligned} \mu_s(\mathcal{A}) &\equiv \int_{\mathcal{A}} d\vec{x}\, |\nabla s|(\vec{x}) \\ \mu_{s_{_R}}(\mathcal{A}) &\equiv \int_{\mathcal{A}} d\vec{x}\, |\nabla s_{_R}|(\vec{x}) \end{aligned} \tag{23}$$

As both measures are multifractal with the same singularity exponents we have:

$$\begin{aligned} \mu_s(B_r(\vec{x})) &= \alpha_s(\vec{x})\, r^{d+h(\vec{x})} \\ \mu_{s_R}(B_r(\vec{x})) &= \alpha_{s_R}(\vec{x})\, r^{d+h(\vec{x})} \end{aligned} \tag{24}$$

Both measures are absolutely continuous [24] the one with respect to the other, so one can be represented in terms of the other using an appropriate density function. The density function, that we will represent by $\rho(\vec{x})$, can be obtained by means of a Radon-Nykodin derivative, namely:

$$\rho(\vec{x}) = \lim_{r\to 0} \frac{\mu_s(B_r(\vec{x}))}{\mu_{s_R}(B_r(\vec{x}))} \tag{25}$$

or, alternatively, it can be calculated as the limit of the ratio of wavelet transforms at decreasing size parameters r, that is,

$$\rho(\vec{x}) = \lim_{r\to 0} \frac{T_\Psi \mu_s(\vec{x}, r)}{T_\Psi \mu_{s_R}(\vec{x}, r)} \tag{26}$$

In this way only the scalar and not the matricial character can be obtained, however. In order to deduce the matrix relating the actual gradients some sort of generalization is required. This generalization is simple for 2D signals; we will concentrate in this particular case for the rest of the section. First, we will consider any 2D vector as a complex number: any multiplication or division of vectors will assume the field structure of the complex plane. We will overline arrows on vectorial quantities in order to remark their vectorial/complex-number character. The multifractal measures above can be generalized to vector valued measures $\vec{\mu}_s$ and $\vec{\mu}_{s_R}$ in the following way:

$$\begin{aligned} \vec{\mu}_s(\mathcal{A}) &\equiv \int_{\mathcal{A}} d\vec{x}\, \nabla s(\vec{x}) \\ \vec{\mu}_{s_R}(\mathcal{A}) &\equiv \int_{\mathcal{A}} d\vec{x}\, \nabla s_R(\vec{x}) \end{aligned} \tag{27}$$

According to this definition, the absolute variation of the measure $\vec{\mu}_s$ [24] corresponds to the scalar measure μ_s, $\| \vec{\mu}_s \| = \mu_s$, and analogously for $\vec{\mu}_{s_R}$. We can now define the Radon-Nykodin derivative, $\vec{\rho}$ using complex division, in the way:

$$\vec{\rho}(\vec{x}) \equiv \lim_{r\to 0} \frac{\vec{\mu}_s(B_r(\vec{x}))}{\vec{\mu}_{s_R}(B_r(\vec{x}))} \tag{28}$$

which can also be calculated by the limit of the ratio of wavelet transforms as before, namely:

$$\vec{\rho}(\vec{x}) = \lim_{r\to 0} \frac{T_\Psi \vec{\mu}_s(\vec{x}, r)}{T_\Psi \vec{\mu}_{s_R}(\vec{x}, r)} \tag{29}$$

where the wavelet transforms are performed over each scalar component of the vectors independently. We will call the vector field $\vec{\rho}$ the sources of the signal s. The sources represent the matrix field of transformation of the gradient ∇s_R to the gradient ∇s,

$$\nabla s(\vec{x}) = \vec{\rho}\, \nabla s_R(\vec{x}) \tag{30}$$

where the product in the rhs must be understood as the product of complex numbers. The vector $\vec{\rho}$ can thus be written as a matrix $\mathbb{M}$ in the way:

$$\mathbb{M} = \begin{pmatrix} \rho_x & -\rho_y \\ \rho_y & \rho_x \end{pmatrix} \tag{31}$$

By construction, the modulus of the vector $\vec{\rho}$ corresponds to the Radon-Nykodin derivative of the scalar measures, eq. (25); that is, $\| \vec{\rho} \| = \rho$. The modulus of the sources field is represented in Figure 7 for Lena's image; the original and the reduced images are also shown to help understanding.

Figure 7: Lena's original images (left), reduced image (middle) and sources (just the moduli are represented; the greater the modulus is, the brighter the point is represented) (right). Notice that sources deviate from average where light or darkness need to be introduced to pass from one image to the other. The sources were calculated by wavelet projection over $\Psi(\vec{x}) = \frac{1}{|\vec{x}|}$, which is scale-invariant and simplifies taking limits in eq. (29)

Obtaining the sources $\vec{\rho}$ is helpful in the analysis of the structure of a given multifractal signal. The sources informs about how the signal has deviated from a naive template (the reduced signal) which can be constructed with geometrical information only (*i.e.*, the oriented MSM). A constant source field means that the signal corresponds to the template up to a change in units. Deviations from the average imply the existence of a perturbation which cannot be explained in the framework of multifractality only. In particular, divergences and zeroes are very interesting, as they represent critical points for the sources field and tells us about places at which the functional character undergoes a transition, a sharp change. Let us also remark that any signal can be reconstructed just knowing the oriented MSM (which spawns the reduced signal) and the sources: according to eq. (30) the original gradient field can be recovered with the reduced signal and the sources, so the original signal can then be reconstructed. The decomposition of any image as reduced signal + sources is a useful tool in both classification and coding. Let us finish by remarking that the sources field must be scale invariant by construction, but at the same time cannot introduce more new singularities, as both signals contain

them all. This restricts enormously the scaling behaviour for this field, a subject which needs to be further investigated.

7 Conclusions

In this paper we have first performed a fast review of the basic concepts for processing physical multifractal systems, then we have presented new ideas and techniques for the analysis and synthesis of multifractal signals. We have given some theoretical reasons supporting why physical signals are easiest to process using multifractal measures rather than linear increments or wavelet projections of multiaffine functions. We have shown examples on the performance of using measures for multifractal analysis.

Besides, we have discussed on the bounded character of singularity exponents for physical signals: physical signals cannot diverge at any point, so there necessarily exists a minimum exponent which bounds by below the possible values to be obtained from any physical signal. We have discussed the consequences of this property in the perspective of several commonly used models.

We have reviewed the concept of reconstructible multifractals: such signals can be reconstructed using information conveyed by the Most Singular Manifold, MSM, which is the fractal set which is associated to the most singular value of the exponents (thus, reconstructible multifractals are bounded, as physical signals). We have seen that the reconstruction algorithm gives a remarkable performance in experimental situations, for very diverse systems.

With the help of the reconstruction property, we have explored new techniques for the synthesis and analysis of multifractal systems. We have seen some simple procedures to construct signals with the same associated multifractal hierarchy starting from a given one; we conclude that there exists equivalence classes of signals sharing the same multifractal structure. On the other hand, we have shown that, under the appropriate conditions, a canonical representative (called reduced signal) of the multifractal equivalence class can be deduced for any signal, and the matricial field relating the representative and the actual signal can be calculated. This field, that we call the sources field, is very informative about the dynamical particularities of the signal which cannot be explained just using multiscaling/multifractal arguments.

We have verified that the sources field is a valuable tool to characterize physical processes in the systems, as energy injection, crisis and other events. We are presently working in the development of such applications.

Acknowledgements

A. Turiel is financially supported by a research contract from *Generalitat de Catalunya* (RED 2002). We also acknowledge financial support from MCyT, contract BFM2000-0626.

References

[1] A. N. Kolmogorov, "Dissipation of energy in a locally isotropic turbulence," *Dokl. Akad. Nauk. SSSR*, vol. 32, pp. 141, 1941.

[2] U. Frisch, *Turbulence*, Cambridge Univ. Press, Cambridge MA, 1995.

[3] Z. S. She and E. Leveque, "Universal scaling laws in fully developed turbulence," *Physical Review Letters*, vol. 72, pp. 336–339, 1994.

[4] Z. S. She and E. C. Waymire, "Quantized energy cascade and log-poisson statistics in fully developed turbulence," *Physical Review Letters*, vol. 74, pp. 262–265, 1995.

[5] A. Arneodo, F. Argoul, E. Bacry, J. Elezgaray, and J. F. Muzy, *Ondelettes, multifractales et turbulence*, Diderot Editeur, Paris, France, 1995.

[6] E. Bacry, J. F. Muzy, and A. Arneodo, "Singularity spectrum of fractal signals from wavelet analysis: exact results," *J. of Stat. Phys.*, vol. 70, pp. 635–673, 1993.

[7] S. G. Roux, A. Arneodo, and N. Decoster, "A wavelet-based method for multifractal image analysis. iii. applications to high-resolution satellite images of cloud structure," *Eur. Phys. J. B*, vol. 15, pp. 765–786, 2000.

[8] S. Mallat and S. Zhong, "Wavelet transform maxima and multiscale edges," in *Wavelets and their applications*, Ruskai M. B. et al, Ed. Jones and Bartlett, Boston, 1991.

[9] Z. R. Struzik, "Determining local singularity strengths and their spectra with the wavelet transform," *Fractals*, vol. 8, no. 2, pp. 163–179, June 2000.

[10] A. Turiel, G. Mato, N. Parga, and J. P. Nadal, "The self-similarity properties of natural images resemble those of turbulent flows," *Physical Review Letters*, vol. 80, pp. 1098–1101, 1998.

[11] A. Turiel and N. Parga, "The multi-fractal structure of contrast changes in natural images: from sharp edges to textures," *Neural Computation*, vol. 12, pp. 763–793, 2000.

[12] B. B. Mandelbrot, A. Fisher, and L. Calvet, "A multifractal model of asset returns," *Cowles Foundation Discussion Paper No. 1164*, 1997.

[13] P. Ivanov, L. Amaral, A. Goldberger, S. Havlin, M. Rosenblum, Z. Struzik, and H. Stanley, "Multifractality in human heartbeat dynamics," *Nature*, vol. 399, pp. 461–465, 1999.

[14] J. Grazzini, A. Turiel, and H. Yahia, "Entropy estimation and multiscale processing in meteorological satellite images," in *Proc. of ICPR 2002*, 2002, vol. 3, pp. 764–768.

[15] A. Turiel and C. Pérez-Vicente, "Multifractal geometry in stock market time series," *Physica A*, vol. 322, pp. 629–649, May 2003.

[16] G. Parisi and U. Frisch, "On the singularity structure of fully developed turbulence," in *Turbulence and Predictability in Geophysical Fluid Dynamics. Proc. Intl. School of Physics E. Fermi*, M. Ghil, R. Benzi, and G. Parisi, Eds., Amsterdam, 1985, pp. 84–87, North Holland.

[17] J. Lévy-Véhel, "Introduction to the multifractal analysis of images," in *Fractal Image Encoding and Analysis*, Y. Fisher, Ed. Springer Verlag, 1997.

[18] J. H. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proc. R. Soc. Lond.*, vol. B265, pp. 359–366, 1998.

[19] E. A. Novikov, "Infinitely divisible distributions in turbulence," *Physical Reviev E*, vol. 50, pp. R3303, 1994.

[20] R.N. Bhattacharya and E.C. Waymire, *Stochastic processes with applications*, Wiley, 1991.

[21] B. Dubrulle, "Intermittency in fully developed turbulence: Log-poisson statistics and generalized scale covariance," *Physical Review Letters*, vol. 73, pp. 959–962, 1994.

[22] A. Turiel and A. del Pozo, "Reconstructing images from their most singular fractal manifold," *IEEE Trans. on Im. Proc.*, vol. 11, pp. 345–350, 2002.

[23] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley and sons, Chichester, 1990.

[24] W. Rudin, *Real and Complex Analysis*, Mc Graw Hill, New York, USA, 1987.

[25] I. Daubechies, *Ten lectures on wavelets*, CBMS-NSF Series in Ap. Math. Capital City Press, Montpelier, Vermont, 1992.

A Scaling requirements for analyzing wavelets over multifractal measures

First of all, it should be remarked that eq. (9) is usually relaxed to a more suitable, easier to deal with, expression, in the way:

$$A(\vec{x}_0)\, r^{d+h(\vec{x}_0)} \;\leq\; \mu(B_r(\vec{x}_0)) \;\leq\; B(\vec{x}_0)\, r^{d+h(\vec{x}_0)} \tag{32}$$

for appropriate positive constants $A(\vec{x}_0)$, $B(\vec{x}_0)$. It is very usual in Physics to present multifractality in a non-rigorous way, as the verification of eq. (9); the theoretical derivations, however, can be worked out only using expressions such as eq. (32). We sometimes make use of expressions like eq. (9), but just as an intuitive, loose notation of eq. (32).

We will consider multifractal measures μ defined over d-dimensional spaces for any value of d. The wavelet Ψ will be assumed to be positive to simplify the discussion. A positive wavelet is not an *admissible* wavelet, that is, the wavelet projections cannot be used to retrieve the original function [25]. However, it can be still used to analyze the singularities of μ, as we will see. In our proof, we will also restrict the discussion to the case of a continuous function Ψ and $\Psi(0) \neq 0$. As it will be shown, Ψ will be required to decrease fast enough.

We will assume that for any point $\vec{x}$ and any size r the wavelet projection $T_\Psi\mu(\vec{x}, r)$ is finite. If it was not the case, as μ is σ-finite it would be possible to define a sequence of finite measures μ_n defined over compact supports A_n such that $\mu|_{A_n} = \mu_n$; but Ψ is continuous so the wavelet projections of the μ_n are finite. The proof for μ would be obtained as a limit case, once the theorem is shown to be valid for the μ_n's,

Finally, we will just study the singularity at $\vec{x} = 0$, the extension for the other points being absolutely trivial. We will denote the singularity at $\vec{x} = 0$ by h_0, that is,

$$\mu(B_r(0)) \approx \alpha_0 \, r^{d+h_0} \tag{33}$$

where the symbol $\approx$ should be understood in the sense of eq. (32). The wavelet projections we are interested in are $T_\Psi \mu(0, r)$, given by:

$$T_\Psi \mu(0, r) \equiv \int d\mu(\vec{x}) \, \frac{1}{r^d} \, \Psi(\frac{\vec{x}}{r}) \tag{34}$$

The statement of the theorem is as follows:

Theorem: *If eq. (33) holds, then for any wavelet Ψ belonging to an appropriate class the following relation holds:*

$$T_\Psi \mu(0, r) \approx \beta_0 r^{h_0} \tag{35}$$

We will present the proof in three stages, for three different classes of functions: set functions, compact support functions and fast decreasing functions.

A.1 Set functions

Let $\Psi(\vec{x})$ be the characteristic or set function of the ball of radius l centered around the origin, that is, $\Psi(\vec{x}) = \chi_{B_l(0)}(\vec{x})$ where the general expression of $\chi_{B_r(\vec{x}_0)}$ is given by:

$$\chi_{B_r(\vec{x}_0)}(\vec{x}) = \begin{cases} 1 & |\vec{x} - \vec{x}_0| < r \\ 0 & |\vec{x} - \vec{x}_0| > r \end{cases} \tag{36}$$

According to eq. (34). the wavelet projections of Ψ at $\vec{x} = 0$ are given by:

$$T_\Psi \mu(0, r) = \int d\mu(\vec{x}) \, \frac{1}{r^d} \, \chi_{B_l(0)}(\frac{\vec{x}}{r}) = \frac{1}{r^d} \int d\mu(\vec{x}) \, \chi_{B_{rl}(0)}(\vec{x}) \tag{37}$$

Hence it follows:

$$T_\Psi \mu(0, r) = \frac{1}{r^d} \int_{B_{rl}(0)} d\mu(\vec{x}) = \frac{\mu(B_{rl}(0))}{r^d} \approx \alpha_0 \, l^{h_0+d} \, r^{h_0} \tag{38}$$

so for those simple functions the theorem is trivially verified.

A.2 Compact support functions

We will just try to find two constants $0 < A < B$ such that:

$$A \, r^{h_0} \leq T_\Psi \mu(0, r) \leq B \, r^{h_0} \tag{39}$$

A lower bound constant A is easily obtained by the condition of continuity and $\Psi(0) \neq 0$, as follows: there exists a finite radius l such that $\Psi(\vec{x}) \neq 0 \; \forall \vec{x} \in B_l(0)$. Let $m > 0$ be the minimum value of $\Psi(\vec{x})$ in $B_l(0)$; then, for all $\vec{x}$ the following inequality holds:

$$m \, \chi_{B_{l(0)}}(\vec{x}) \; \leq \; \Psi(\vec{x}) \tag{40}$$

Let $0 < A_0 < B_0$ be the bounding constants for μ at $\vec{x} = 0$, that is, such that $A_0 r^{d+h_0} < \mu(B_r(0)) < B_0 r^{d+h_0}$ (as μ is multifractal, they exist). It follows:

$$m \, A_0 \, l^{h_0+d} \, r^{h_0} \; \leq m \, T_{\chi_{B_{l(0)}}} \mu(0,r) \leq T_\Psi \mu(0,r) \tag{41}$$

So, A can be taken as $A = mA_0 l^{h_0+d}$.

To obtain B we will use that Ψ has bounded support; so there exists a finite radius L such that $\Psi(x) = 0 \; \forall \vec{x} : |\vec{x}| > L$. As Ψ is continuous, it possesses a finite maximum M in $B_L(0)$. Hence, the following functional inequality holds:

$$M \, \chi_{B_L(0)}(\vec{x}) \; \geq \; \Psi(\vec{x}) \tag{42}$$

and analogously to the case of the lower bound, we conclude that a possible upper bound is given by $B = MB_0 L^{h_0+d}$.

A.3 Fast decreasing functions

We will search constants $0 < A < B$ as in the previous case. The constant A can be calculated exactly like for compact support functions, so we just need to find B.

Let the radius l be defined as in the previous subsection and let us define the sets R_i as follows: $R_i = B_{2^i l}(0) - B_{2^{i-1} l}(0)$ for $i \geq 1$ and $R_0 = B_l(0)$. Let M_i be the maximum of Ψ over R_i. Hence, the following functional inequality holds:

$$\Phi(\vec{x}) \equiv \sum_{i=0}^{\infty} M_i \, \chi_{R_i}(\vec{x}) \; \geq \; \Psi(\vec{x}) \tag{43}$$

where χ_{R_i} is the characteristic function of R_i, which equals 1 over R_1 and 0 outside. We can conclude by proving that there exists an upper bound B for Φ such that $T_\Phi \mu(0,r) \leq B r^{h_0}$, which by eq. (43) is also an upper bound for Ψ. To obtain this bound, let us suppose that there exists a finite radius $L > 0$ such that if $|\vec{x}| > L$, for all $K > 1$ the function Ψ verifies:

$$\Psi(K\vec{x}) < \Omega(K)\Psi(\vec{x}) \tag{44}$$

where $\Omega(K)$ decreases to zero as K goes to infinity faster than any polynomial, that is,

$$\lim_{K \to \infty} K^N \Omega(K) \; = \; 0 \quad \forall N > 0 \tag{45}$$

This requirement on Ψ above is a small modification on the condition defining Schwartz's class. We introduce the radius L to avoid forcing the function to be strictly decreasing from the origin, which would define a very restrictive class of functions. Let i_0 be the least integer greater than $\log_2(L/l)$. Following eq. (44), for all $i' > i > i_0$:

$$M_{i'} < \Omega(2^{i'-i})\, M_i \tag{46}$$

We will first take r as $r = 2^{-J}$, for $J > 0$ integer. We can thus decompose the dilation of Φ by a factor r, $\Phi(\vec{x}/r)$, as:

$$\Phi(\vec{x}/r) = \sum_{i=0}^{i_0+J} M_i\, \chi_{R_i}(\vec{x}/r) + \sum_{i=i_0+J+1}^{\infty} M_i\, \chi_{R_i}(\vec{x}/r) \;= \Phi_J(\vec{x}) + \delta_J(\vec{x}) \tag{47}$$

For $i > i_0 + J$, $\chi_{R_i}(\vec{x}/r) = \chi_{R_i}(2^J\vec{x}) = \chi_{R_{i-J}}(\vec{x})$. The residual function $\delta_J(\vec{x})$ verifies the following bound (obtained applying eq. (46))

$$\delta_J(\vec{x}) < \Omega(1/r) \sum_{i=i_0}^{\infty} M_i\, \chi_{R_i}(\vec{x}) \;=\; \Omega(1/r)\, \delta_0(\vec{x}) \tag{48}$$

Given the dependence of $\delta_J(\vec{x})$ in r it decays very fast, and this function is negligible in comparison with any power of r, so we can ignore this part. Hence, $T_\Phi\mu(0,r) \approx T_{\Phi_J}\mu(0,r)$ for $r = 2^{-J}$. For a general value of r, we also obtain $T_\Phi\mu(0,r) \approx T_{\Phi_J}\mu(0,r)$ where J is such that $2^{-J-1} < r \leq 2^{-J}$. Taking into account that:

$$T_{\chi_{R_i}}\mu(0,r) \;\leq\; l^{h_0+d}\, 2^{(h_0+d)i}\, (B_0 - 2^{-(h_0+d)}\, A_0)\, r^{h_0} \tag{49}$$

it follows:

$$T_{\Phi_J}\mu(0,r) \;\leq\; l^{h_0+d}\, (B_0 - 2^{-(h_0+d)}\, A_0) \left[\sum_{i=0}^{i_0+J} 2^{(h_0+d)i}\, M_i\right] r^{h_0} \tag{50}$$

so the upper bound B is given by:

$$B \;=\; l^{h_0+d}\, (B_0 - 2^{-(h_0+d)}\, A_0) \sum_{i=0}^{\infty} 2^{(h_0+d)i}\, M_i \tag{51}$$

where the series $\sum_{i=0}^{\infty} 2^{(h_0+d)i}\, M_i$ is finite due to the fast decreasing of M_i.

A.4 Refined choices for the wavelet

The previous section gives a proof of the capability as singularity analyzers for wavelets Ψ chosen in a very restrictive class of functions. Some of the conditions can be relaxed from the mathematical point of view (for instance, the requirement of continuity could be relaxed to integrability and boundness), although their precise formulation does not change very much the result in practical applications. It is also clear that singularities can be detected just using positive wavelets, but non-positive wavelets could be used as well. So none of those conditions is going to change significantly the main result nor the experimental performance (except for the questions already discussed concerning the minimum distinguishable resolution). However, there is a requirement which is critical to extract singularities and whose correct tuning allows to improve significantly the performance; namely, the condition of fast decay expressed in eq. (45).

Any *a priori* knowledge about the properties of the multifractal measure μ allows enlarging the class of valid wavelets by relaxing the fast decay condition. The most relevant information is the knowledge of bounds on the range of possible singularity exponents h_0. Let us assume that there exists a maximum singularity exponent h_M; in that case, the fast decay condition can be modified so that eq. (45) is only required for $N \leq N_0 \equiv d + h_M$. The argument goes as follows: let us assume that for $|\vec{x}|$ large enough, $|\Psi(\vec{x})| \sim |\vec{x}|^{-N_0}$ (in the sense that $\lim_{|\vec{x}|\to\infty} |\vec{x}|^{N_0}|\Psi(\vec{x})|$ is not divergent). So, $\Omega(|\vec{x}|) \propto |\vec{x}|^{-N_0}$ and M_i is then given by $M_i \propto 2^{-N_0 i}$. The existence of the upper bound B given in eq. (51) is limited to the cases in which the series appearing in the definition is finite; but for the wavelet we consider now this series is proportional to

$$\sum_{i=0}^{\infty} 2^{(h_0+d-N_0)i} \tag{52}$$

which converges if $N_0 > h_0 + d$ and diverges for $N_0 \leq h_0 + d$. So, it suffices to take $N_0 > h_M + d$ to assure the convergence. This argument can be refined as follows: let $\Psi(\vec{x})$ a wavelet decaying as $|\vec{x}|^{-N_0}$ at the infinity. Then, at any point $\vec{x}$ in which $\mu(B_r(\vec{x})) \approx \alpha(\vec{x})\, r^{h(\vec{x})+d}$, the wavelet projection is given by $T_\Psi\mu(\vec{x}, r) \sim r^{h_\Psi(\vec{x})}$, where:

$$h_\Psi(\vec{x}) \;=\; \begin{cases} h(\vec{x}) & , \quad h(\vec{x}) < N_0 - d \\ \\ N_0 - d & , \quad h(\vec{x}) \geq N_0 - d \end{cases} \tag{53}$$

This result allows designing high performance wavelets at the cost of reducing the range of observed singularities. Let us finishing by remarking that this result is analogous to the one given by Arneodo et al. [5], in which the authors shown that in order to generalize LI's to LWP's the wavelet of choice must vanish a certain number of moments; notice that singularities beyond the maximal number of moments are truncated to the maximum detectable value, very similar to our result here. The main difference lies in that we can use positive wavelets to analyze measures, what allows in principle a better spatial resolution.