Multiscale representations: fractals, self-similar random processes and wavelets

Marie Farge[1], Kai Schneider[2], Olivier Pannekoucke[3]
and Romain Nguyen van yen[1]

December 15, 2010

[1] *LMD–IPSL–CNRS, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France*
[2] *M2P2–CNRS & CMI, Université de Provence, 39 rue F. Joliot–Curie, 13453 Marseille Cedex 13, France*
[3] *Centre National de Recherches Météorologiques, 42, avenue Gaspard Coriolis 31057 Toulouse Cedex 01, France*

Contents

1 Introduction

Many growth processes which shape the human environment generate structures at all scales, *e.g.*, trees, rivers, ligthning bolts. Likewise, most geophysical flows happen on a wide range of scales, *e.g.*, winds in the atmosphere, currents in the oceans, seismic waves in the mantle. In general, both kinds of phenomena are governed by nonlinear dynamical laws which give rise to chaotic behaviour, and it is thus very difficult to follow their evolution, let alone predict it. Only in the last few decades could the systems of nonlinear equations modelling environmental fluid flows be solved, thanks to the development of numerical methods and the advent of super-computers. Although the present computer performances still remain insufficient to simulate from first principles, *i.e.*, by Direct Numerical Simulation (DNS), many environmental fluid flows, especially those which are turbulent, appropriate multiscale representations may contribute to the success of that ongoing enterprise. The goal of this review is to present three of them: fractals, self-similar random processes and wavelets.

A fractal is a set of points which presents structures that looks essentially the same at all scales. When only its large scale features are considered, a certain shape is observed, which does not become simpler when zooming towards small scales, but on the contrary remains quite similar as it is at large scale. This goes on from one scale to the other, up to the point that one cannot tell what is the scale of observation. When measuring the length, surface, or volume of a fractal object, it is found that, in contrast to classical geometrical objects, *e.g.*, circle or polygones, a definite answer cannot be obtained since the measured value increases when the scale of observation decreases. Let us now consider a simple example of a drop falling into water, an experiment that can be easily done with a glass of water, a drop of oil and a drop of ink. While falling, the shape of the oil drop becomes more and more spherical, therefore more regular than it was at the instant of impact. Since oil is hydrophobic, the drop tends to minimize the interface between oil and water for a given volume. In contrast, the shape of the ink drop becomes more and more convoluted, since the drop is unstable and splits into smaller drops. In absence of surface tension the interface between ink and water would then become fractal in the limit of long times. Indeed, since ink is hydrophilic the drop tries to maximize the interface for a given volume. Both systems satisfy the same equations and only one parameter, the surface tension, differs, which implies either minimization or maximization of the interface. The solution of the former exists and is

smooth, while the maximum does not exist. John Hubbard, who suggested this example, concludes: *'The world is full of systems which are trying to reach an optimum which does not exist, and consequently they evolve towards structures which are complicated at all scales. This happens for trees, which try to maximize their exposure to light, for lungs and capillaries, which try to maximize the interface between tissue and blood. The great work of Mandelbrot has been to tell, very loudly and in a very convincing way, that the world is full of complicated phenomena, of complicated objects having structure at all scales.'* [27]

Fractals can be traced back to the discovery of continuous non differentiable functions, *e.g.*, the Weierstrass function, and non rectifiable curves, *e.g.*, the Sierpinski gasket. Measure theory, as developed in particular by Felix Hausdorff at the end of the 19th century, and integration theory, as redesigned by Henri Lebesgue and others at the beginning of the 20th century, together with the study of recursive sequences in the complex plane, by Pierre Fatou and Gaston Juia, were all precursors of fractals, although a different terminology was used at those times. Only when computer graphics became widely available in the 60's was one able to visualize fractals and wonder about their apparent complexity. Although the mathematical tools were already there, it is Benoît Mandelbrot, from IBM Research Center in Yorktown (USA), who popularized fractals and named them in the seventies. Actually, before he started talking about fractals, Mandelbrot was a specialist of Brownian motion that he had learned about during the time he was at Ecole Polytechnique in Paris where he studied under the French probabilist Paul Lévy [36]. It was Mandelbrot who gave in 1968 the name 'Fractional Brownian Motion' [40] to the self-similar stochastic processes proposed by Kolmogorov in 1940 [32], which are generalization with long-range correlated increments of the classical Brownian motion.

The mathematical foundation of wavelets is more recent, since the continuous wavelet transform has been introduced only in the eighties by Jean Morlet and Alex Grossmann. Jean Morlet was working in oil exploration for the French company Elf, while Alex Grossmann was a specialist of coherent states in quantum mechanics and a member of CPT (Centre de Physique Théorique) in Marseille (France). From their work Ingrid Daubechies, Pierre-Gilles Lemarié and Yves Meyer constructed several orthogonal wavelet bases. Soon after, Stéphane Mallat and Yves Meyer introduced the concept of multiresolution analysis (MRA) which lead to the fast wavelet transform (FWT). Without the FWT, the wavelet transform would have remained confined to text books and theoretical papers. The same was true for the Fourier transform that would not have entered our everyday's life without the combination of computers and FFT (fast Fourier transform), invented by Gauss around 1805 and rediscovered by Cooley and Tukey in 1965.

The aim of this paper is to give researchers working in environmental fluid dynamics some mathematical tools to study the multiscale behaviour of many natural flows. For the sake of clarity, we propose to divide what is presently named 'fractals' into two classes: deterministic fractals and self-similar random processes. We will keep the terminology 'fractals' to designate the former, which are constructed following some deterministic procedure iterated scale by scale. For the latter we propose to return to the 'pre-fractal' terminology of 'self-similar random processes', which are ensembles of

random realizations whose statistics exhibit some scaling behaviour. We have thus organized the multiscale methods presented here into three classes: fractals, self-similar random processes and wavelets. Note that they are all mathematical tools which do not have any explanatory power *per se*. They require the scientist who use them to have enough physical insight to interpret the results and decide if this tool is actually appropriate to his problem. If a new technique is not mastered well enough, it would provide an *a priori* interpretation, which is built-in the tool without the user being sufficiently cautious about that risk. To avoid such a drawback, we will here limit ourselves to give definitions, expose methods and illustrate their use on academic examples rather than applications. We will justify this choice in the conclusion by showing how such misinterpretation has happened in one field of application, with both fractals and wavelets.

2 Principles

2.1 Fractals

2.1.1 Definition and history

To define what 'fractal' means is quite a difficult endeavour since one finds in the literature different definitions. We propose the following definition: a fractal is a shape which is so convoluted, irregular or fragmented that it is not rectifiable, *i.e.*, one cannot measure its length. Its boundary is a set of points, either connected or disconnected, which looks the same at different scales and tends to be space-filling. If the points remain connected the boundary can be parametrized by a continuous but non-differentiable function. Otherwise, the fractal is a dust of disconnected points which can only be parametrize by a measure. A fractal shape looks complicated although it is not, since it can be generated by a simple iterative procedure. The difficulty is, given an observed complicated shape, can we infer the simple rule which has generated it? In most cases the answer is no and this is why methods developed under the trademark 'fractals' are rather descriptive than predictive.

Benoît Mandelbrot introduced the word 'fractal' in 1975, in a book first published in French [42] and two years later in English [43], but he managed to keep the definition vague and varied them throughout his books. The first definition he gave is: *'...'fractal object' and 'fractal', terms that I have formed for this book from the Latin adjective 'fractus' which means irregular or broken'* [42]. Subsequently Mandelbrot succeeded in gathering under the same name different mathematical objects which were proposed before but were considered by most mathematicians as surprising, anecdotic or weird. Poincaré recalled that *'we have seen a rabble of functions arise whose only job, it seems, is to look as little as possible like decent and useful functions. No more continuity, or perhaps continuity but no derivatives [...] Yesterday, if a new function was invented it was to serve some practical end, today there are specially invented only to show up the arguments of our fathers, and they will never have any other use'*[6]. An example of such entertaining mathematical object was the fractal curve known as the 'snow flake', see Fig. 1(a), published in 1904 by Helge von Koch in the Swedish journal 'Arkiv for Matematik' [31].

In 1918 the French Academy set for its 'Grand Prix des Sciences Mathématiques' the iteration of fractional functions and Gaston Julia won that prize. Independently Gaston Julia and Pierre Fatou were studying rational maps in the complex plane by iterating polynomials, *e.g.* quadratic maps. In 1977 Adrien Douady and John Hubbard used Newton's method to solve the quadratic map $f_c(z) = z^2 + c$, with $z \in \mathbb{C}$, $c \in \mathbb{C}$ a parameter. This quadratic map is the simplest nonlinear dynamical system one can think of in the complex plane and they studied the set K_c of z for which the n-th iterate of f, $f_c^n(z)$, converges. The frontier of K_c is now called the Julia set of f_c. Benoît Mandelbrot, who worked for IBM and had thus access to large computers, graphical facilities and good programmers, made visualizations to help understanding that problem. In a paper, published in 1982, Douady and Hubbard [15] showed that if $0 \in K_c$ the set K_c is connex, and they denoted it M to pay tribute to Mandelbrot for his visualizations. They commented as follows: *'Benoît Mandelbrot has obtained on a computer a very beautiful picture of M, exhibiting small islands which are detached from the principal component. These islands are in fact connected by filaments which escape the computer'* [15]. Without any doubt computer visualization has played an essential role in the dissemination of fractals outside mathematics.

The main contribution of Mandelbrot has been to widely popularize fractals, thanks to computer visualization. His argument is that fractals are more appropriate to describe natural phenomena than the classical objects geometers have been using for centuries, namely rectifiable curves (*e.g.*, circle and other ovals) or piece-wise regular curves (*e.g.*, triangle and other polygons). He illustrated that with many examples [42, 43] such as: the length of the coast of Britain, fluctuations of stock exchange, flood data...

2.1.2 Fractal dimension

The box-counting dimension d of a simple geometrical object A is defined by

$$N(l) \underset{l \to 0}{\sim} l^{-d}, \tag{1}$$

where $N(l)$ is the minimal number of boxes of side length l required to cover the whole set of points A. For instance, if A is a regular curve (*i.e.*, everywhere differentiable), like a segment, then $d = 1$. If A is as simple surface (respectively a simple volume), then $d = 2$ (respectively $d = 3$). In those cases, d corresponds to the topological dimension of the manifold. The definition of d given by Eq. (1) can be extended to more general sets, for which d is in general no more an integer, which brings up the concept of fractal set for which d is then called the fractal dimension. A more rigorous definition of the fractal dimension relies on the Hausdorff dimension [23]. But this is less easy to compute from data, and in most cases the box-counting dimension d and the Hausdorff dimension are equal. Hence, we consider thereafter that the Hausdorff dimension is equivalent to the fractal dimension as defined in Eq. (1).

Classical illustrations of fractal sets of points are given by the Cantor dust and the von Koch curve. The former is a set of points obtained by dividing recursively a segment into three parts, where only the first and the third sub-segment are retained, this construction is illustrated in Fig. 1(a). Since each step of the algorithm doubles the number of segments while their length is divided by three, after n iterations there are 2^n segments of length 3^{-n}. Since each segment includes all the sub-segments of the

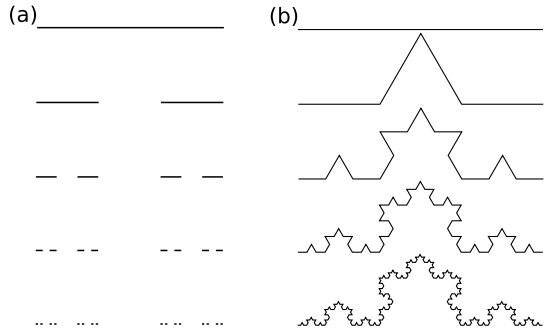


Figure 1: Illustration of the first four iteration steps leading to the Cantor dust (a) and to the von Koch snow flake (b).

following iterations, it results that one can cover this ensemble of segments with 2^n balls of radius 3^{-n}. The fractal dimension of the Cantor set as defined by the box counting method is thus:

$$d_C = \lim_{n\to\infty}\left(-\frac{\ln 2^n}{\ln 3^{-n}}\right) = \frac{\ln 2}{\ln 3}. \tag{2}$$

Therefore, the fractal dimension of the Cantor set is between 0 and 1, which implies that the set is neither an ensemble of isolated points nor a line.

The second example, the von Koch curve, is also obtained using a recursive process where in this case each segment of length l is replaced by four segments of length $l/3$ as illustrated in Fig. 1(b). Starting from the unit length segment, after n iterations there are 4^n segments of length 3^{-n}. The fractal dimension of the von Koch curve, as defined by the box counting method, is thus:

$$d_K = \lim_{n\to\infty}\left(-\frac{\ln 4^n}{\ln 3^{-n}}\right) = \frac{\ln 4}{\ln 3}. \tag{3}$$

The fractal dimension is hence contained between 1 and 2, implying that the length of the von Koch curve is infinite while its surface is zero.

2.1.3 Hölder exponent and singularity spectrum

Fractal dimension was defined above as a geometrical property that characterizes a set of points, but it can also be used to analyze the regularity of functions or distributions as detailed now. Complex signals, like those encountered in environmental data analysis, can be seen as superpositions of singularities. One way of detecting a singularity of a function f at a point x is to measure its Hölder regularity. The function f is said to be α-Hölder in x if there exists a polynomial P_n of degree n and a constant K such that for sufficiently small l

$$|f(x+l) - P_n(l)| \leq K|l|^\alpha, \tag{4}$$

where n is the integer part of α (*i.e.*, $n \leq \alpha < n+1$). The Hölder regularity of f in x is the maximum α such that f is α-Hölder in x. Note that for $\alpha = 1$ the function is called Lipschitz-continuous in x. If f is $n+1$ times differentiable in x, then

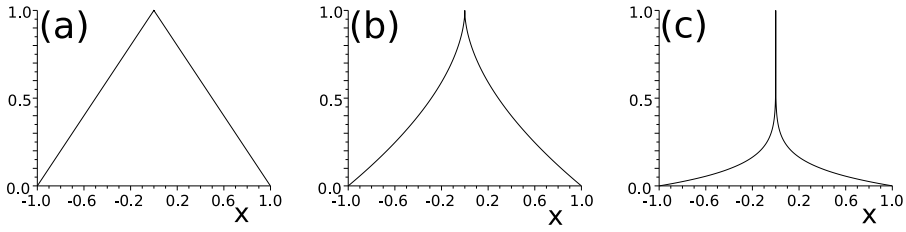


Figure 2: Illustrations of singularities at point $x = 0$ with the graph of the function $f(x) = 1 - |x|^\alpha$, with $\alpha = 1, 5/9$ and $1/9$ respectively for (a), (b) and (c).

$P_n(l) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} l^k$, the Taylor expansion of f in x. The smaller the Hölder exponent, the stronger the singularity is (Fig. 2).

Some functions, sometimes called multi-fractal functions [46], have Hölder regularity which varies from one point to the other. It is thus interesting to analyze the set of points A_α where a function has Hölder regularity α, for example by computing its fractal dimension $d(\alpha)$. The singularity spectrum is the function which associates $d(\alpha)$ to each value of the Hölder regularity α. It is not easy to compute directly, but a trick can be used to estimate it. We briefly sketch the idea without giving a rigorous demonstration.

If we consider a covering $\mathcal{B}_l$ of the support of the function f by boxes of the form $B_{x,l} = [x, x + l]$, then, by definition of the regularity, we obtain that

$$|f(x + l) - f(x)| \sim l^{\alpha_x}, \tag{5}$$

where $\sim$ stands for the magnitude order. Hereafter l is assumed to be small ($l \ll 1$). By definition of the fractal dimension, the minimal number of balls needed to recover the support of A_α is

$$N_{A_\alpha}(l) \sim l^{-d(\alpha)}. \tag{6}$$

The moment function $Z_q(l)$ associated to the cover $\mathcal{B}_l$ of the domain is defined by $Z_q(l) = \sum_{B_l \in \mathcal{B}_l} |f(x + l) - f(x)|^q$. Note that it is sometimes called partition function by analogy with statistical physics. Contributions of boxes containing an α-singularity are given by $|f(x + l) - f(x)|^q \sim l^{q\alpha}$, while the number of such boxes is given by Eq. (6). Hence, the moment function can be approximated by $Z_q(l) \sim \sum_h l^{q\alpha - d(\alpha)}$. Since, l is assumed to be small, the leading contribution in Z_q is given by the term of minimum exponent $q\alpha - d(\alpha)$. It follows that the moment function is approximated by $Z_q(l) \approx l^{\tau(q)}$, where $\tau(q) = \inf_\alpha \{q\alpha - d(\alpha)\}$ is the multiscale exponent. Hence, as shown in [46] the singularity spectrum $d(\alpha)$ appears as being the Legendre-Fenchel inverse transform of the multiscale exponent $\tau(q)$

$$d(\alpha) = \inf_\alpha \{q\alpha - \tau(q)\}. \tag{7}$$

For instance, the singularity spectrum of the Riemann function $f(x) = \sum_{n=1}^{\infty} \frac{\sin n^2 x}{n^2}$, is $d(\alpha) = 4h - 2$ if $\alpha \in [1/2, 3/4]$ and $d(3/2) = 0$. Another example is given by the Devil's staircase, related to the Cantor set. In the Cantor set generation algorithm that we have described earlier, each interval was split into two pieces in a symmetric

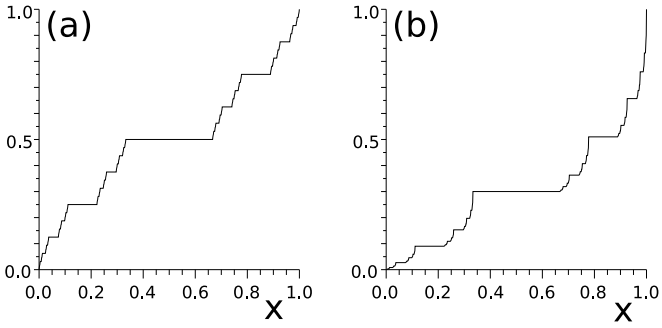


Figure 3: Illustration of the Devil's staircase with an homogeneous repartition of mass (a) and an heterogeneous repartition of mass where each left subsegment receives 30% of the mass (b).

fashion at each iteration. Denoting by μ the limit in the sense of distributions of the characteristic function of the set as a result of the iterative process, the associated function $f(x) = \int_0^x \mu(du)$ is called the Devil's staircase, as illustrated in Fig. 3(a). It can be shown that each singularity of f is of the same Hölder regularity $\alpha = \ln 2/\ln 3$, and the support of these singularities is the Cantor set. Thus, the singularity spectrum is in that case reduced to the point $d(\ln 2/\ln 3) = \ln 2/\ln 3$.

More complex singularity spectra can be obtained by considering more general fractals similar to the Devil's staircase (see Fig. 3(b)), which we do not detail here.

2.2 Self-similar random processes

2.2.1 Definition and history

Stochastic fractals, sometimes also called fractal noise, are self-similar random processes, which yield models for many applications, *e.g.*, turbulent velocity fields. The self-similarity of a stochastic process is only satisfied in the statistical sense and hence a given realization is not necessarily self-similar. One can distinguish between scalar or vector valued random processes in one or higher space dimensions. For the sake of simplicity we restrict ourselves in the following to scalar valued processes in one space dimension, which typically corresponds to time t or space x. The simplest ones are Gaussian random processes.

Denoting by $\xi(t)$ a Gaussian random process that we assume to be stationary (*i.e.*, all its statistics are invariant by translation), its one-point probability distribution function (pdf) is given by

$$p(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(\xi-\mu)^2}{2\sigma^2}\right), \tag{8}$$

where μ is the mean and σ the standard deviation. In the following we suppose that the mean vanishes since we are only interested in the fluctuations. The process $\xi(t)$ is then characterized by its autocovariance function, defined as $\langle\xi(\tau)\xi(0)\rangle$, where $\langle\cdot\rangle$ denotes the expectation, computed either from ensemble, time or space averages. Equivalently it can be characterized by its energy spectrum defined as the Fourier trans-

form of its autocovariance function, $E(f) = \int_{\mathbb{R}} \langle \xi(\tau)\xi(0)\rangle \, e^{-\iota 2\pi\tau f} \, d\tau = \left\langle |\widehat{\xi}(f)|^2 \right\rangle$ with $\widehat{\xi}(f) = \int \xi(t) e^{-\iota 2\pi t f} dt$ and $\iota = \sqrt{-1}$. The energy spectrum yields the spectral distribution of energy , and summing over all frequencies thus yields the total energy.

A simple example of a Gaussian process is the Wiener process, also called Brownian motion, which was proposed in 1900 by Louis Bachelier as a model to describe market price fluctuations. Its mathematical properties were studied in 1923 by Norbert Wiener who called it the fundamental random function. The nomenclature 'Brownian' is due to Paul Lévy who named the Wiener process Brownian motion in memory of the Scottish botanist Richard Brown, who in the beginning of the 19th century observed the random motion of pollen suspension in water. An extension of Brownian motion has been introduced by Kolmogorov in 1940 [32], a spectral representation was given by Hunt in 1951 [28], and Mandelbrot proposed in 1968 to call it 'fractional Brownian motion' [40].

2.2.2 Brownian motion

For Brownian motion the variance of the increments scales as

$$\langle |B(t) - B(\tau)|^2 \rangle = |t - \tau| \tag{9}$$

and the Hölder regularity of the trajectories is 1/2. The formal derivative of a Wiener process is called a Gaussian white noise. It is stationary and uncorrelated, *i.e.*, its autocovariance function is $\langle \xi(\tau)\xi(0)\rangle = \delta(\tau)$, where δ is the Dirac distribution, or equivalently its energy spectrum is constant, $E(f) = 1$. The constant spectrum means that all frequencies f have the same weight, and hence the noise is called white by analogy with white light. Correlated Gaussian processes have non constant spectra and they are called colored noise. Power-law spectra $E(f) \propto f^{\beta}$ are of particular interest as the processes are statistically self-similar, *i.e.*, $\langle \xi(\lambda\tau)\xi(0)\rangle = \lambda^{\alpha}\langle \xi(\tau)\xi(0)\rangle$. However such processes are not necessarily stationary and, in order to recover stationarity, we consider their increments. Due to non stationarity the energy spectrum can only be defined formally and can no more be integrated (due to infrared divergence). For example, the generalized energy spectrum of Brownian motion satisfies the power law $E(f) \propto 1/f^2$. Brownian motion thus belongs to the class of so-called $1/f$ processes, which have been studied for many applications.

2.2.3 Fractional Brownian motion

Fractional Brownian motion is a kind of self-similar Gaussian process which is non stationary and whose energy spectrum follows a power law. A given realization of such a noise is almost everywhere singular and has the same Hölder regularity at all points, *i.e.*, it is mono-fractal.

The fractional Brownian motion $B_H(t)$ is the Gaussian process with zero mean such that

$$B_H(t = 0) = 0 \tag{10}$$

and

$$\langle |B_H(t) - B_H(\tau)|^2 \rangle = |t - \tau|^{2H}, \tag{11}$$

where $0 < H < 1$ is an additional parameter called Hurst exponent [29]. Here H determines the regularity of the trajectories. The smaller H, the lower the regularity. For $H < 1/2$ the increments of the process are correlated, while for $H > 1/2$ they are anti-correlated. For $H = 1/2$ we get the classical Brownian motion. In all cases the process is said to be long-range dependent.

The covariance function of B_H is given by

$$\langle B_H(t)B_H(\tau)\rangle = \frac{1}{2}(|t|^{2H} + |\tau|^{2H} - |t-\tau|^{2H}) \tag{12}$$

Note that one given realization of fractional Brownian motion is not a fractal: the self-similarity is only fullfilled in the statistical sense. Indeed, Eq. (11) implies that

$$\langle |B_H(\lambda t) - B_H(\lambda\tau)|^2\rangle = \lambda^{2H}\langle |B_H(t) - B_H(\tau)|^2\rangle. \tag{13}$$

However, it can be shown that a given trajectory has the pointwise Hölder regularity $H = \alpha$ almost surely and is almost (besides for a set of measure zero) nowhere differentiable.

The self-similarity of the fractional Brownian motion $B_H(t)$ implies for the energy spectrum a power law behavior with exponent $2H + 1$,

$$E(f) = \frac{C_H}{f^{2H+1}}. \tag{14}$$

Gaussian processes, and thus also fractional Brownian motion, can be represented in Fourier space using the Cramer representation

$$B_H(t) = \int_{\mathbb{R}} \sqrt{E(f)}(e^{\iota 2\pi ft} - 1)d\xi(f), \tag{15}$$

where $d\xi(f)$ is an orthogonal Gaussian increment process with $\langle d\xi(f)d\xi(f')\rangle = \delta(f - f')$, which means that the measure corresponds to Gaussian white noise. The term $(e^{\iota 2\pi ft} - 1)$ instead of $e^{\iota 2\pi ft}$ guarantees that $B_H(0) = 0$.

2.2.4 Multi-fractional Brownian motion

Allowing for time (or space) varying Hurst exponents generalizes fractional Brownian motion, which is mono-fractal, to introduce stochastic multi-fractals. Their construction is based on the spectral representation of stochastic processes. The starting point is a function $\theta : [0, 1] \rightarrow]0, 1[$ and the corresponding multi-fractional Brownian process can be defined using the spectral representation,

$$B_\theta(t) = \int_{\mathbb{R}} \frac{e^{i2\pi ft} - 1}{|f|^{\theta(t)+1/2}} d\xi(f). \tag{16}$$

The pointwise Hölder regularity of $B_\theta(t)$ is almost surely equal to $\theta(t)$ and the Hausdorff dimension of the graph of B_θ is $2 - \inf\{\theta(t), 0 \leq t \leq 1\}$.

Methods for synthezising fractional Brownian motion are presented in section 3.2.

2.3 Wavelets

2.3.1 Definition and history

In a signal the useful information is often carried by both its frequency content and its time evolution, or by both its wavenumber content and its space evolution. Unfortunately the spectral analysis does not give information on the instant of emission of each frequency or the spatial location of each wavenumber. This is due to the fact that the Fourier representation spreads the time or space information among the phase of all Fourier coefficients, therefore the energy spectrum (*i.e.*, the modulus of the Fourier coefficients) loses any structural information in time or space. This is a major limitation of the classical way to analyze non stationary signals or inhomogeneous fields. A more appropriate representation should combine these two complementary descriptions.

From now on we will consider a signal $f(x)$ which only depends on space. The theory is the same for a signal $f(t)$ which depends on time, except that the wavenumber k should in that case be replaced by the frequency ν, and the spatial scale l by the time scale or duration τ. Any function $f \in L^2(\mathbb{R})$ also has a spectral representation $\widehat{f}(k)$ defined as

$$\widehat{f}(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi\iota kx} dx, \tag{17}$$

where $\iota = \sqrt{-1}$.

However, there is no perfect representation due to the limitation resulting from the Fourier's uncertainty principle (also called Heisenberg's uncertainty principle when it is used in quantum mechanics). One thus cannot perfectly analyse the signal f from both sides of the Fourier transform at the same time, due to the restriction $\Delta x \cdot \Delta k \geq C$, where Δx is the spatial support of $|f(x)|$ and Δk the spectral support of $|\widehat{f}(k)|$, with C a constant which depends on the chosen normalization of the Fourier transform.Due to the uncertainty principle there is always a compromise to be made in order to have, either a good spatial resolution Δx at the price of a poor spectral resolution Δk, or a good spectral resolution Δk while loosing the space resolution Δx, as it is the case with the Fourier transform. These two representations, in space or in wavenumber, are the most commonly used in practice because they allow to construct orthogonal bases onto which one projects the signal to be analysed and processed.

In order to try to recover some space locality while using the Fourier transform, Gabor [22] has proposed the windowed Fourier transform, which consists of convolving the signal with a set of Fourier modes $e^{2\pi\iota kx}$ localized in a Gaussian envelope of constant width l_0. This transform allows then a time-frequency (or space-wavenumber) decomposition of the signal at a given scale l_0, which is kept fixed. But unfortunately, as shown by Balian [4], the bases constructed with such windowed Fourier modes cannot be orthogonal. In 1984 Grossmann and Morlet [24] have proposed a new transform, the so called wavelet transform, which consists of convolving the signal with a set of wave packets, called wavelets, of different widths l and locations x. To analyze the signal $f(x)$, we generate the family of analysing wavelets $\psi_{l,x}$ by dilation (scale parameter l) and translation (position parameter x) of a given function ψ which oscillates with a characteristic wavenumber k_ψ in such a way that its mean remains zero. The wavelet transform thus allows a space-scale decomposition of the signal f given by its wavelet

coefficients $\widetilde{f}_{l,x}$. The wavelet representation yields the best compromise in view of the Fourier uncertainty principle, because it adapts the space-wavenumber resolution $\Delta x \cdot \Delta k$ to each scale l. In fact it gives for the large scales a good spectral resolution Δk but a poor spatial resolution Δx, while, on the contrary, it gives a good spatial resolution Δx with a poor spectral resolution Δk for the small scales.

In 1989 the continuous wavelet transform has been extended to analyse and synthesize signals or fields in higher dimensions [45, 1]. In 1985 Meyer, while trying to prove the same kind of impossibility to build orthogonal bases as done by Balian [4] in the case of the windowed Fourier transform, has been quite surprised to discover an orthogonal wavelet basis built with spline functions, now called the Meyer-Lemarié wavelet basis [35]. In fact the Haar orthogonal basis, which was proposed in 1909 in the PhD thesis of Haar and published in 1910 [25], is now recognized as the first orthogonal wavelet basis known, but the functions it uses are not regular, which limits its application. In practice one likes to build orthogonal wavelet bases using functions having sufficient regularity, depending on the application. In particular, following Meyer's work, Daubechies has proposed in 1988 [12] new orthogonal wavelet bases built with compactly supported functions defined by discrete Quadrature Mirror Filters (QMFs) of different lengths. The longer the filter, the more regular the associated functions. In 1989 Mallat has devised a fast algorithm [37] to compute the orthogonal wavelet transform using wavelets defined by QMF. Later Malvar [39], Coifman and Meyer [7] have found a new kind of windows of variable width which allows the construction of orthogonal adaptive local cosine bases which have then been used to design the MP3 format for sound compression. The elementary functions of such bases, called Malvar's wavelets, are parametrized by their position x, their scale l (width of the window) and their wavenumber k (proportional to the number of oscillations inside each window). In the same spirit, Coifman, Meyer and Wickerhauser [8] have proposed the so-called wavelet packets which, similarly to compactly supported wavelets, are wavepackets of prescribed regularity defined by discrete QMFs, from which one can construct orthogonal bases.

The Fourier representation is well suited to solve linear equations, for which the superposition principle holds and whose generic solutions either persist at a given scale, or spread to larger scales. In contrast, the superposition principle does not hold anymore for nonlinear equations, *e.g.*, the Navier-Stokes equations which is the fundamental equation of fluid dynamics. In this case the equations can no more be decomposed as a sum of simpler equations which can be solved separately. Generically the time evolution of their solutions involves a wide range of scales and could even lead to finite time singularities, *e.g.*, shocks. The 'art' of predicting the evolution of such nonlinear evolution, the generic case being turbulent flows, consists of disentangling the nonlinear from the linear dynamical components: the former should be deterministically computed while the latter could, either be discarded or their effect be statistically modelled. A review of the different types of wavelet transforms and their applications to analyse and compute turbulent flows is given in [19, 50].

2.3.2 Continuous wavelet transform

The only condition a real-valued function $\psi(x) \in L^2(\mathbb{R})$ or a complex-valued function $\psi(x) \in L^2(\mathbb{C})$ should satisfy to be called wavelets is the admissibility condition

$$C_\psi = \int_0^\infty \left|\widehat{\psi}(k)\right|^2 \frac{dk}{|k|} \; < \; \infty, \tag{18}$$

where $\hat{\psi} = \int_{-\infty}^{\infty} f(x) e^{-2\pi\imath kx} dx$ its the Fourier transform of ψ. If ψ is admissible its mean is zero, therefore $\widehat{\psi}(k=0) = 0$, and only then the wavelet transform is invertible. The wavelet ψ may also have other properties, such as being well-localized in physical space $x \in \mathbb{R}$ (fast decay of f for $|x|$ tending to ∞) and smooth, *i.e.*, well-localized in spectral space (fast decay of $\hat{\psi}(k)$ for $|k|$ tending to ∞). For several applications, in particular to study deterministic fractals or random processes, one also wishes that $\widehat{\psi}(k)$ decays rapidly near 0, or equivalently that the wavelet has enough cancellations such that

$$\int_{-\infty}^{\infty} x^m \psi(x) dx = 0 \quad \text{for} \quad m = 0, ..., M-1 \quad , \tag{19}$$

namely that its first M moments vanish. In this case the wavelet analysis will enhance any quasi-singular behaviour of the signal by hiding all its polynomial behavior up to degree m.

One then generates a family of wavelets by dilatation (or contraction), with the scale parameter $l \in \mathbb{R}^+$, and translation, with the location parameter $x \in \mathbb{R}$, of the so-called mother wavelet and obtains

$$\psi_{l,x}(x') = c(l)\, \psi\left(\frac{x'-x}{l}\right) \tag{20}$$

where $c(l) = l^{-1/2}$ corresponds to all wavelets being normalized in the L^2-norm, *i.e.*, they have the same energy, while for $c(l) = l^{-1}$ all wavelets are normalized in the L^1-norm.

The continuous wavelet transform of a function $f \in L^2(\mathbb{R})$ is the inner product of f with the analyzing wavelets $\psi_{l,x}$, which yields the wavelet coefficients

$$\widetilde{f}(l,x) = \langle f, \psi_{l,x} \rangle = \int_{-\infty}^{\infty} f(x') \psi^*_{l,x}(x')\, dx', \tag{21}$$

with ψ^* denoting the complex-conjugate of ψ. The continuous wavelet coefficients measure the fluctuations of f at scale l and around position x. If the analyzing wavelets have been normalized in L^2-norm, then the squared wavelet coefficients correspond to the energy density of the signal whose evolution can be tracked in both space and scale. Note that the wavelet coefficients written in L^1-norm are related to the wavelet coefficients written in L^2-norm by

$$\tilde{f}_{L^1} = l^{-1/2} \tilde{f}_{L^2}. \tag{22}$$

To study the Hölder regularity of a function and estimate its singularity spectrum, one prefers to use wavelet coefficients in L^1-norm (see section 2.1.3).

The function f can be reconstructed without any loss as the inner-product of its wavelet coefficients $\widetilde{f}$ with the analyzing wavelets $\psi_{l,x}$:

$$f(x') = {C_\psi}^{-1} \int_\infty^\infty \int_{0^+}^\infty \widetilde{f}(l,x)\psi_{l,x}(x') \, \frac{dl}{l^2} \, dx \tag{23}$$

with C_ψ the constant of the admissibility condition given in Eq. (18), which only depends on the chosen wavelet ψ.

Like the Fourier transform, the wavelet transform is linear, *i.e.*, we have

$$\widetilde{\beta_1 f_1(x) + \beta_2 f_2(x)} = \beta_1 \widetilde{f_1(x)} + \beta_2 \widetilde{f_2(x)} \tag{24}$$

with $\beta_1, \beta_2 \in \mathbb{R}$, and it is also an isometry, *i.e.*, it conserves the inner-product (Plancherel's theorem), and in particular the energy (Parseval's identity). The continuous wavelet transform is also covariant by translation and by dilation, both properties which are partially lost by the orthogonal wavelet transform. Let us also mention that, due to the localization of wavelets in physical space, the behaviour of the signal at infinity does not play any role. In contrast, the non local nature of the trigonometric functions used for the Fourier transform does not allow us to locally analyse or process a signal with it.

Fig. 4 shows six examples of wavelet analyses of academic signals using the complex-valued Morlet wavelet: a Dirac spike (a), a step function (b), the superposition of two cosine functions having different frequencies (c), succession of two cosine functions having different frequencies (d), a chirp (e), a Gaussian white noise (f). The modulus of the wavelet coefficients is plotted as a function of position x on abscissa and the log of the scale l on ordinate. The curved black lines delimitate the region where the coefficients are not influenced by left and right boundaries, which correspond to the spatial support of the wavelets localized in $x = 0$ and $x = 1$. The horizontal straight black line indicates the scale below which the wavelet coefficients are aliased, due to undersampling of the wavelets at small scales. Note in particular that three signals, namely Fig. 4 (a), (e) and (f), have similar flat Fourier and wavelet spectra (see Sec. 3.3.2), although the space-scale representation of the energy density in wavelet space exhibit very different behaviours.

The extension of the continuous wavelet transform to analyse signals in d dimensions is made possible by replacing the affine group by the Euclidean group with rotation. One thus generates the d-dimensional wavelet family $\psi_{l,r,\vec{x}}$ with l the dilation factor, R the rotation matrix in $\mathbb{R}^d$ and $\vec{x}$ the translation such that:

$$\psi_{l,\vec{x},r}(x') = \frac{1}{l^{d/2}} \, \psi\left(r^{-1}\left(\frac{\vec{x}' - \vec{x}}{l}\right)\right) \tag{25}$$

where the wavelet ψ should satisfies the admissibility condition which becomes in d-dimensions:

$$C_\psi = \int_0^\infty \left|\widehat{\psi}(k)\right|^2 \frac{d^d k}{|k|^d} \; < \; \infty \tag{26}$$

If we consider $d = 2$ then the rotation matrix $R(\theta)$ is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} . \tag{27}$$

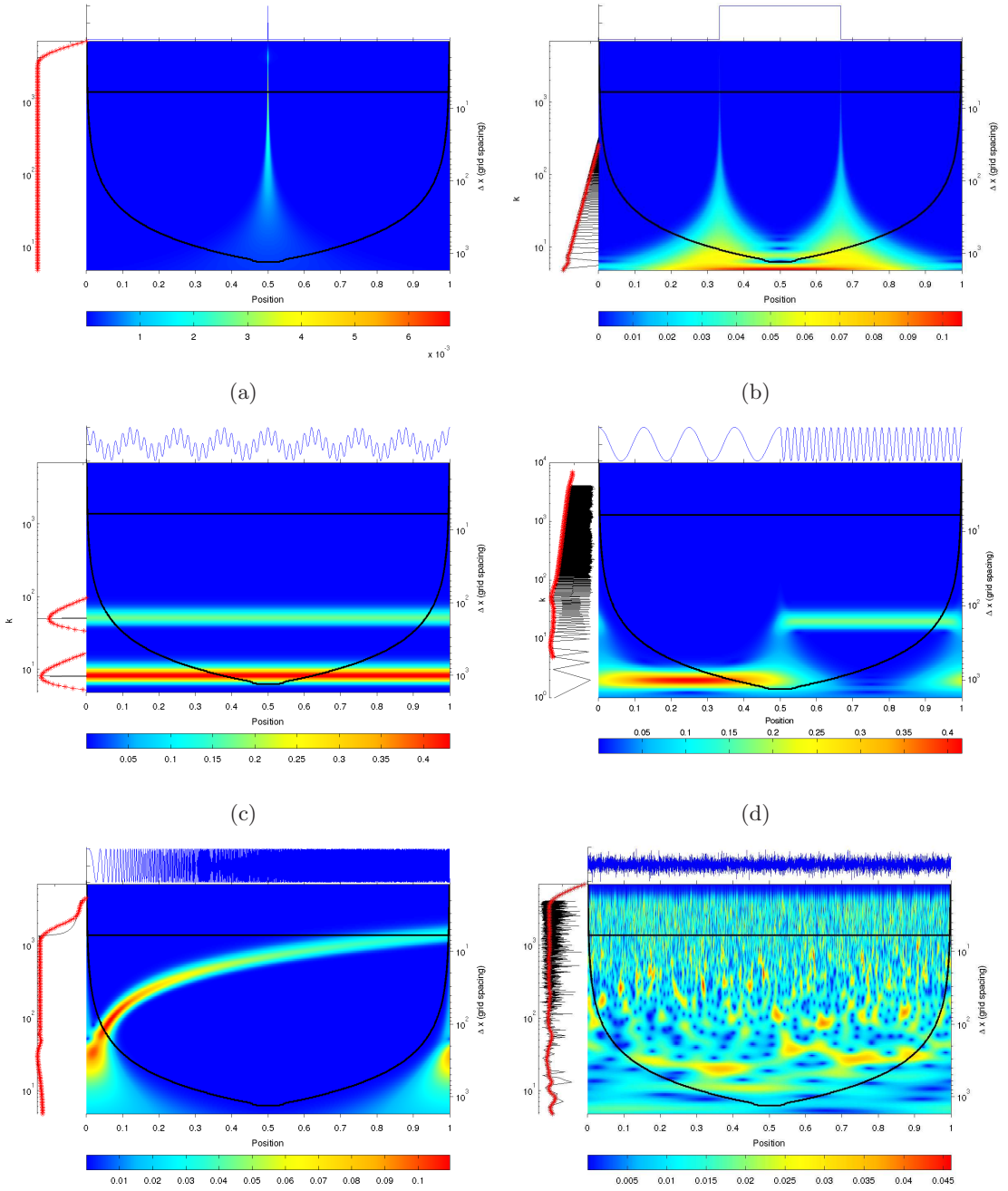


Figure 4: Examples of wavelet analyses of academic signals, namely a Dirac spike (a), a step function (b), the superposition of two cosine functions having different frequencies (c), the succession of two cosine functions of very different frequencies (d), a chirp (e), and finally one realization of a Gaussian white noise (f)). The modulus of the complex-valued Morlet wavelet coefficients are plotted as a function of position and scale. The original signal is plotted on the top. The Fourier spectrum (black curve) and the wavelet scalogram (red crosses), as defined in Sec. 3.3.2, are also shown on the left, with the axes rotated by 90 degrees.

The wavelet analysis of a two-dimensional scalar field $f(\vec{x})$ is

$$\widetilde{f(l,\vec{x},\theta)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\vec{x}')\psi^*_{l,\vec{x},\theta}(x')\,d\vec{x}' \quad , \tag{28}$$

and the wavelet synthesis is

$$f(\vec{x}') = \frac{1}{C_\psi} \quad \int_0^\infty \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{2\pi} \widetilde{f}(l,\vec{x},\theta)\,\psi^*_{l,\vec{x},\theta}(\vec{x}')\,\frac{dl\,d\vec{x}\,d\theta}{l^3} \quad . \tag{29}$$

In dimensions larger than two one needs $d-1$ angles to describe the rotation operator R.

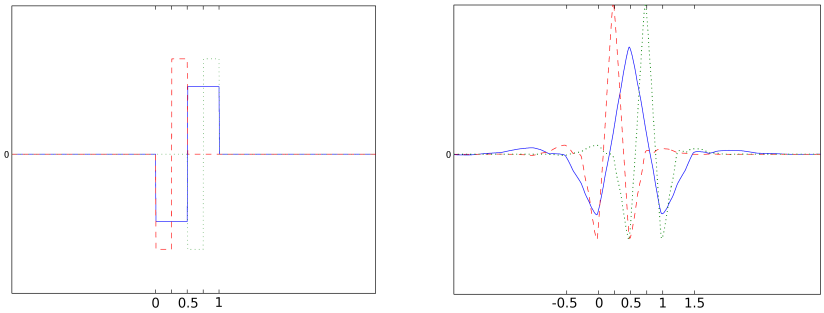
2.3.3 Orthogonal wavelet transform

Wavelets can also be used to construct discrete representations of various function spaces, called frames [11], by selecting a discrete subset of all their translations and dilations. Some special frames sampled on a dyadic grid $\lambda = (j,i)$, *i.e.*, for which the scale l has been discretised by octaves j and the position x by spatial steps $2^{-j}i$, constitute orthogonal wavelet bases. The main difference between the continuous and the orthogonal wavelet transform is that all orthogonal wavelet coefficients are decorrelated. This is not the case for the continuous wavelet coefficients which are redundant and correlated in both space and scale. These correlations can be visualised by plotting the modulus of the continuous wavelet coefficients of one realisation of a white noise computed with a Morlet wavelet, see Fig. 9 (b). The patterns one thus observes are due to the reproducing kernel of the continuous wavelet transform, which corresponds to the correlation between all the analyzing wavelets themselves. Note that the redundancy of the continuous wavelet transform is actually useful for algorithms such as edge and texture detection. Moreover, its translation and dilation invariance eliminates some artefacts one encounters when denoising with the orthogonal wavelet transform which does not preserve those invariances.

As a tutorial example, we explain the orthogonal wavelet decomposition of a three-dimensional vector field. For this we consider a square integrable vector-valued field $\vec{x} \to \vec{f}(\vec{x}) \in L^2(\mathbb{T}^3)$, where $\mathbb{T}^3 = (\mathbb{R}/\mathbb{Z})^3$ is the 3D torus and $\vec{x} = (x_1, x_2, x_3) \in \mathbb{T}^3$. Note that in practice the fact that f is defined on a torus simply means that periodic boundary conditions are assumed. The input data consists in discrete values of f sampled with a resolution $N_k = 2^J$ in each direction. N_k is thus the number of grid points and J is the number of octaves in each of the three directions, and the total number of grid points is thus $N = N_1 \times N_2 \times N_3 = 2^{3J}$. The mother wavelet is denoted ψ as above, and we assume that it satisfies all the necessary conditions (see, *e.g.*, [13]) so that the wavelets $\psi_{l,i}$ defined by Eq. (20) are pairwise orthogonal if (l,x) is sampled on the dyadic grid $\{(2^{-j}, 2^{-j}i) \mid j = 0,\dots,J-1,\ i = 0,\dots,2^j-1\}$. We also assume that the wavelet has been suitably periodized. To develop the components f_d of $\vec{f}$ (with $d = 1,2,3$) into an orthogonal wavelet series from the largest scale $l_{\max} = 2^0$ to the smallest scale $l_{\min} = 2^{-J+1}$, we need to construct a 3D multi-resolution analysis (MRA) as follows [13, 19].

For λ belonging to the index set

$$\Lambda_0 = \{(j,\vec{\mu},\vec{\imath}) \mid j = 0,\dots,J-1,\ \vec{\mu} \in \{0,1\}^3,\ \vec{\imath} \in \{0,\dots,2^j-1\}^3\} \ ,$$



(a) Haar wavelet (b) Coifman 12 wavelet

Figure 5: Orthogonal wavelets: Haar wavelet (left) and Coifman 12 wavelet (right). We have superposed one wavelet at scale $j = 0$ (blue) and position $x = 0.5$) and two wavelets at the next smaller scale $j = 1$, located at position $x = 0.25$ (red) and $x = 0.75$ (green). They are mutually orthogonal, which can be directly seen for the Haar wavelet and which is much less obvious for the Coifman wavelet.

define the 3D wavelet ψ_λ by

$$\psi_\lambda(x_1, x_2, x_3) = 2^{3j/2} \prod_{\substack{1 \le k \le 3 \\ \mu_k = 0}} \phi(2^j x_k - i_k) \prod_{\substack{1 \le k \le 3 \\ \mu_k = 1}} \psi(2^j x_k - i_k) \ ,$$

where ϕ is the scaling function (also called father wavelet) associated to ψ [13]. Here, the parameters j and $\vec{\imath}$ are the 3D equivalent to the scale and positions parameters that we are already familiar with from the preceding discussion of the 1D continuous wavelet transform. The new parameter, $\vec{\mu}$, provides an additional degree of freedom which is necessary to represent 3D data without loss of information. It controls the directions of oscillation of the wavelet. For example, if $\vec{\mu} = (1, 0, 0)$, the wavelet is oscillatory (*i.e.*, it has vanishing mean) in the first direction, whereas it has nonvanihsing mean in the two others directions. If $\vec{\mu} = (0, 0, 0)$, ψ_λ is the 3D equivalent to a scaling function, in which case we shall denote it ϕ_λ, following the classical convention. The wavelets are thus indexed by the subset of Λ_0 whose elements satisfy $\vec{\mu} \neq 0$, which we denote Λ. The wavelet coefficients and scaling coefficients of f_d are then simply defined by

$$\begin{aligned} \widetilde{f}^d_\lambda &= \langle f_d, \psi_\lambda \rangle \\ \overline{f}^d_\lambda &= \langle f_d, \phi_\lambda \rangle \ , \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2(\mathbb{R}^3)$.

Now we have all the ingredients to write down the wavelet series of f_d:

$$f_d = \overline{f}_{(0,0,0)} + \sum_{\lambda \in \Lambda} \widetilde{f}^d_\lambda \psi_\lambda \ . \tag{30}$$

The first term is a constant which is in fact the mean value of f, and the sum over λ contains all the oscillations of f at finer and finer scales, $j = 0, \ldots, J-1$, while preserving some amount of space-locality thanks to the position index $\vec{\imath}$, and also some

amount of directionality thanks to $\vec{\mu}$. Hence the expansion coefficients appearing in Eq. (30) can be used to compute directional and/or scale-wise statistics of $\vec{f}$, as we shall see further down. Importantly, there exists a fast wavelet algorithm with $O(N)$ complexity, where N denotes the number of wavelet coefficients used in the computation. It is thus asymptotically even faster than the FFT (Fast Fourier Transform), whose complexity is $O(N \log_2 N)$.

3 Methods of analysis

3.1 Fractals

3.1.1 Estimation of the fractal dimension

The box-counting algorithm is a simple method to compute the fractal dimension of a given object (a set of points S in Euclidean space $\mathbb{R}^d$, for example a curve in two dimensions or an iso-surface in three dimensions) by counting the number of boxes (squares in two dimensions, cubes in three dimensions, ...) which cover the object. First the object is overlaid with an equidistant Cartesian grid of size ℓ. Then the number of boxes with side length ℓ covering the object is counted which yields $N(\ell)$. Subsequently the grid size ℓ is reduced (*e.g.*, by a factor 2), a refined grid is overlaid and the number of boxes covering the objet is counted again. The above procedure is repeated until the finest resolution of the object is obtained. Finally, the number of boxes $N(\ell)$ covering the object is plotted against the inverse grid size $1/\ell$ in log-log representation. A straight line is fitted to the curve thus obtained and the slope of the curve yields the fractal dimension of the set S as defined by Eq. (1).

For a regular smooth curve (*e.g.*, a straight line in two or three dimensions) we can observe that the number of boxes covering the curve is proportional to the inverse of the grid size and hence its dimension is 1 which is equal to its topological dimension. For a smooth surface (*e.g.*, the surface of a sphere in three dimensions) we find that the number of boxes increases quadratically with the inverse grid size which yields its topological dimension of two. For fractals the obtained dimension differs from its topological one.

Besides pathological cases, the limit obtained with the box counting algorithm corresponds to the Hausdorff dimension and thus this technique is an efficient way for computing it.

3.1.2 Synthesis of fractal sets

Now we discuss a method to generate a fractal set of points based on iterated functions, recursively applied. An iterated function system (IFS) is a set of contractions $\{f_i\}_{i\in[1,N]}$ from $\mathbb{R}^d$ into itself such that there exists for each i a constant c_i such that $0 < c_i < 1$ with $|f_i(x) - f_i(y)| \leq c_i|x - y|$. The Hutchinson function F associated to the IFS is the transformation from $\mathcal{C}(\mathbb{R}^d)$ to itself, where $\mathcal{C}(\mathbb{R}^d)$ denotes the set of all compact subsets of $\mathbb{R}^d$, defined by

$$F(A) = f_1(A) \cup \cdots \cup f_N(A), \tag{31}$$

with $A \in \mathcal{C}(\mathbb{R}^d)$. It can be shown that F itself is also a contraction defined into $\mathcal{C}(\mathbb{R}^d)$ for the Hausdorff distance δ_H, that is $\delta_H\left(F(A), F(B)\right) \leq c\,\delta_H(A, B)$, where $\delta_H(A, B) =$

$\max\{\sup_{x\in A}\inf_{y\in B}|x-y|, \sup_{y\in B}\inf_{x\in A}|x-y|\}$ and $c=\max\{c_i\}$. Because of the completeness of the metric space $\big(\mathcal{C}(\mathbb{R}^d),\delta_H\big)$, F admits a fixed point in $\mathcal{C}(\mathbb{R}^d)$, and this fixed point is a compact limit ensemble A_F, obtained as $A_F=\lim_{n\to\infty}F^n(A)$, where A is an arbitrary initial compact set, and A_F verifies $A_F=F(A_F)$.

As illustration for an IFS, we consider the IFS $\{f_1,f_2\}$ defined on the real line $\mathbb{R}$ by $f_1(x)=x/3$ and $f_2(x)=x/3+2/3$. These functions are contractions with ratio 1/3. When applying these two contractions to the segment $[0,1]$, we obtain the algorithm for generating the Cantor set, as illustrated in Fig. 1. The Cantor set is thus the limit ensemble of the IFS $\{f_1,f_2\}$. In the particular case where the IFS is made of disconnected or just-touching affine functions $f_i(x)=c_iR_ix+b_i$ where $0<c_i<1$ is the magnitude, R_i the rotation matrix and b_i the translation, then the fractal dimension d of the limit set is linked to the similitude magnitude c_i by the relation

$$\sum_i c_i^d=1. \tag{32}$$

By applying this relation to the Cantor set, we obtain the equation $2(1/3)^d=1$, whose solution is the fractal dimension $d=\ln 2/\ln 3$ already found above. Similarily, the von Koch curve can be obtained from an IFS of four similitudes of magnitude 1/3, so that its fractal dimension satisfies $4(1/3)^d=1$, leading to the known result $d=\ln 4/\ln 3$.

To construct the limit ensemble, a direct solution is to start from a simple compact set and to make it evolve by using the Hutchinson function associated to the IFS. However this solution is computationally costly, since we have to deal with sets. An more efficient alternative is to use a random procedure as we will describe now. From a single point $A=\{x_0\}$ which is a compact set, a recursive process is generated so that $x_{n+1}=w_n$ where w_n is randomly chosen within the list $\{f_i(x_n)\}$ where $f_i(x_n)$ is sampled with probability p_i. If $f_i(x)=A_ix+b_i$, where A_i is a matrix, then p_i can be defined as $p_i=\frac{|\det A_i|}{\sum_k|\det A_k|}$. The intuitive reason of this choice for p_i is that the volume of the unit square transformed by f_i is $|\det A_i|$. When the determinant is zero, p_i is set to a small value compared to the other non zero determinants, and then normalized to ensure the probability normalization $\sum_i p_i=1$.

Another possibility to construct a fractal set of points from an existing set of points, is given by the collage theorem [5]. We consider a compact ensemble $\mathcal{S}$ of $\mathbb{R}^d$ and $\varepsilon>0$. The idea is to be able to reconstruct this ensemble from an IFS strategy, which would be easy if an IFS generating the pattern was known exactly. However, in pratical applications the generating system is unknown. The collage theorem states that, if one finds an IFS $\{f_i\}_{i\in[1,N]}$ such that the Hutchinson function F leaves $\mathcal{S}$ invariant up to a tolerance ε, *i.e.*, $\delta_H\left(\mathcal{S},F(\mathcal{S})\right)\leq\varepsilon$, then the limit ensemble A_F associated to the IFS satisfies

$$\delta_H(\mathcal{S},A_F)\leq\frac{\varepsilon}{1-c}, \tag{33}$$

where c is the contraction ratio of F.

Even if this theorem does not lead to a constructive method to determine an appropriate IFS, it provides a useful way for building fractal sets from a given set of points. In practice, the IFS can be looked for within a reduced class of contractions. For instance, one can try to estimate the smallest set of similitudes required to ensure a given tolerance ε.

3.1.3 Singularity spectrum

As an illustration of the singularity spectrum and its limitations, we compute the singularity spectrum of a function f and we compare its singularity spectrum when noise is added.

In Fig. 6(b), we show the singularity spectrum of the function f plotted in Fig. 6(a). The support of the spectrum is the whole interval $(0, 1)$, and the fractal dimension of the Hölder exponent close to $\alpha = 1$ is about $d = 0.7$. It is larger than the fractal dimension of stronger singularities (having small Hölder exponents). Hence, the support where the signal is regular is larger than the one where it is irregular, as seen in Fig. 6(a). If a white noise with a weak standard deviation of $\sigma = 0.01$ is added, see Fig. 6(c), then the signal becomes more irregular leading to a singularity spectrum truncated at an Hölder exponent closed to $\alpha = 0.5$, as seen on Fig. 6(d). Moreover, the support of the singularities becomes larger since the fractal dimension $d(\alpha)$ for $\alpha = 0.5$ for the noise-free signal, in Fig. 6(a), is close to $\alpha = 0.5$, see Fig. 6(b), while for the noisy signal in Fig. 6(c) it is close to $\alpha = 1$, see Fig. 6(d). This effect is reinforced with a more intense noise of standard deviation $\sigma = 0.1$, see Fig. 6(e) and (f).

This illustrates that the computation of the singularity spectrum is sensitive to the amount of noise present in the signal. Thus adding white noise to a signal reduces the regularity since large Hölder exponents disappear as the amount of noise increases, as seen in Fig. 6.

3.2 Self-similar random processes

3.2.1 Analysis

The Hurst exponent H of a stochastic process can be estimated by considering the quadratic variation of a given realization, *e.g.*, observed data. For fractional Brownian motion $B_H(t)$ with $t \in [0, 1]$ the quadratic variation V_N associated to the step size $\delta t = 1/N$, N being the number of sampling points, is given by

$$V_N = \sum_{k=0}^{N-1} \left[B_H\left(\frac{k}{N} + \frac{1}{N}\right) - B_H\left(\frac{k}{N}\right) \right]^2 \quad . \tag{34}$$

This quadratic variation can be related to the Hurst exponent by

$$V_n = c\, n^{1-2H}, \tag{35}$$

where c is a constant. Moreover the quadratic variation of the dyadically subsampled data, taking only one out of two values of $B_H(k/N)$, is $V_{N/2}$. It follows that

$$\frac{V_N}{V_{N/2}} = 2^{1-2H}, \tag{36}$$

which leads thus to the Hurst exponent

$$H = \frac{1}{2}\left(1 - \log_2 \frac{V_N}{V_{N/2}}\right) \quad . \tag{37}$$

Hence this relation can be used to estimate H from the data. It only requires to compute the quadratic variation of both the data and the dyadically subsampled data.

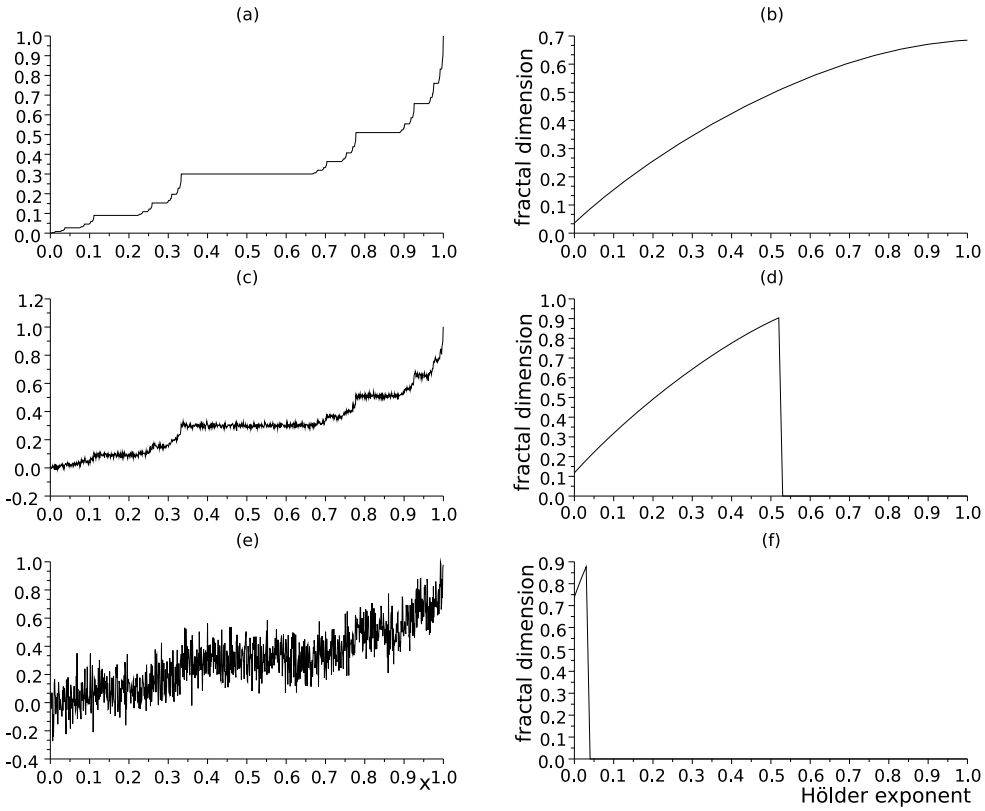


Figure 6: Singularity spectrum of a function f (a) (the Devil's staircase) and its noisy versions perturbed with a white noise of standard deviation $\sigma = 0.01$ (c) and $\sigma = 0.1$ (e). The corresponding singularity spectra are shown on the right column. Without noise (b), with noise of standard deviation $\sigma = 0.01$ (d) and $\sigma = 0.1$ (f).

3.2.2 Synthesis

Different approaches are available for the synthesis of self-similar random processes which are typically either based on the spectral representation of stochastic processes or construct the process in physical space using a decomposed covariance matrix. Additionally wavelet techniques have been developed which allow the efficient generation of realizations with long range dependence and with many scales without imposing a cut–off scale thanks to the vanishing moment property of the wavelets.

For synthezising fractional Brownian motion numerically one can either discretize the Cramer representation in a suitable way or generate it directly in physical space by applying the decomposed covariance matrix to Gaussian white noise.

For the latter the discrete covariance matrix $\Gamma_{i,j} = \langle B_H(t_i) B_H(\tau_j) \rangle$ for $i, j = 1, ..., N$, where N denotes the number of grid points, is first assembled. Then a Cholesky decomposition $\Gamma = LL^t$ is computed (where L is a lower triangular matrix with positive diagonal entries, and L^t is its transpose). Then, a vector of length N is constructed by taking one realization of Gaussian white noise with variance 1, *i.e.*, $\xi(t_i)$ for $i = 1, ..., N$. A realization of fractional Brownian motion is then obtained by multiplication of ξ with L,

$$B(t_i) = L_{ij} \xi(t_j)$$

where summation over j is assumed. For further details on generating Gaussian and also non Gaussian processes we refer to [10].

Different wavelet techniques for synthesizing fluctuating fields using self-similar random processes with a wide range of scales have been proposed. Elliot and Majda [16, 17] proposed a wavelet Monte-Carlo method to generate stochastic Gaussian processes with many scales for one dimensional scalar fields and for two dimensional divergent-free velocity fields. The fields thus obtained have a $k^{-5/3}$ scaling of the energy spectrum (which means that the increments grow as $l^{2/3}$) and thus correspond to fractional Brownian motion with a Hurst exponent $H = 2/3$. Applications were dealing with the simulation of particle dispersion (Elliot & Majda) [17]. A related construction was proposed by Tafti and Unser [51].

An interesting technique from image processing, which was originally developed for generating artificial clouds in computer animations was proposed in [9]. Therewith intermittent scalar valued processes in two space dimensions can be efficiently generated which have a given energy distribution which could be self-similar. The resulting process is stricly band-limited.

3.2.3 Application to fractional Brownian motion

To illustrate the fractional Brownian motion we show in Fig. 7 (right) three realizations of different fractional Brownian motion for $H = 0.5$ (corresponding to classical Brownian motion), $H = 0.75$ and 0.9. The corresponding increments, which are fractional Gaussian noise with different correlations, are shown in Fig. 7 (left). We can observe that the regularity of the curves increases for larger values of H.

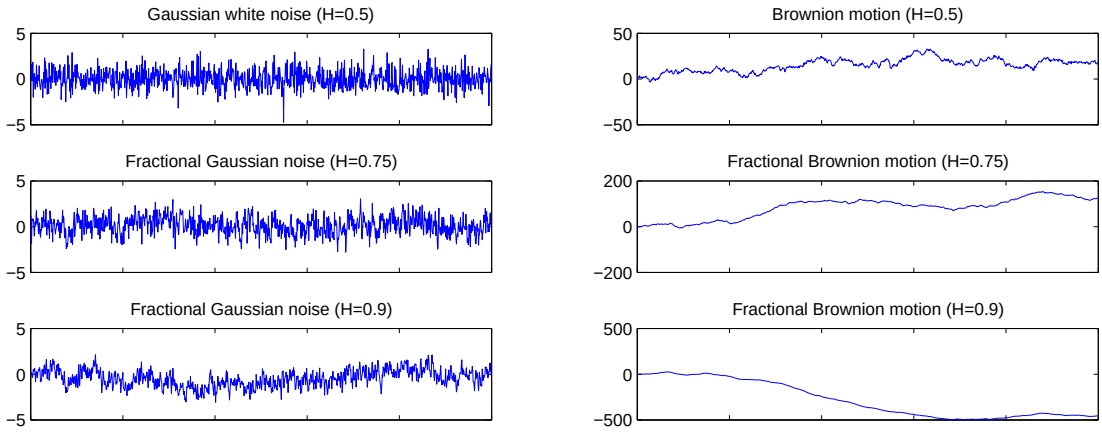


Figure 7: Sample trajectories of Gaussian fractional noise (left column), and of fractional Brownian motion (right column) for three different values of the Hurst exponent H. The Gaussian fractional noise (left column) corresponds to increments of the fractional Brownian motion (right column). The resolution is $N = 1024$.

To model random process with short range correlation we can suppose that the covariance function decays exponentially $\propto \exp(-t/\tau_c)$. The corresponding spectral density decays $\propto 4\tau_c/(1+(f\tau_c)^2)$. Fig. 8 shows examples for different values of τ_c (left) and different spectral densities (right). For increasing α the apparent regularity of the trajectory increases, although the actual regularity of the underlying function remains the same.

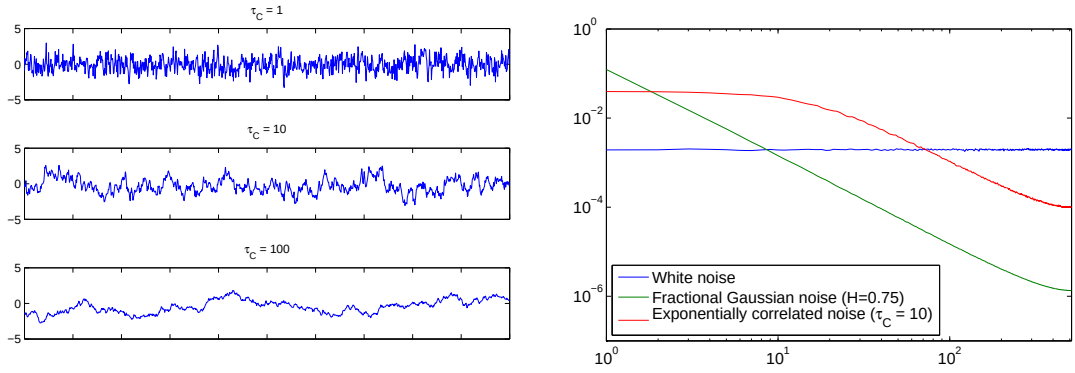
3.3 Wavelets

3.3.1 Wavelet analysis

The choice of the kind of wavelet transform one needs to solve a given problem is essential. Typically if the problem has to do with signal or image analysis, then the continuous wavelet transform should be preferred. The analysis benefits from the redundancy of the continuous wavelet coefficients which thus allows to continuously unfold the information content into both space and scale. The best is to choose a complex-valued wavelet, *e.g.*, the Morlet wavelet, since from the wavelet coefficients one can directly read off the space-scale behaviour of the signal and detect for instance frequency modulation laws or quasi-singularities, even if they are superposed. For this one plots the modulus and the phase of the wavelet coefficients in wavelet space, with a linear horizontal axis corresponding to the position x, and a logarithmic vertical axis corresponding to scale l, with the largest scale at the bottom and the smallest scale being at the top.

A classical real-valued wavelet is the Marr wavelet, also called 'Mexican hat', which is the second derivative of a Gaussian,

$$\psi(x) = (1-x^2)\, e^{\frac{-x^2}{2}} \tag{38}$$



(a) (b)

Figure 8: (a) Sample trajectories of Gaussian noise with exponentially decaying covariance. (b) Spectra averaged over 1000 realizations for three types of noise with identical variances, sampled on $N = 2^{14}$ points.

and its Fourier transform is

$$\widehat{\psi}(k) = k^2 \, e^{\frac{-k^2}{2}} \tag{39}$$

The most useful complex-valued wavelet is the Morlet wavelet,

$$\psi(x) = e^{\imath k_\psi x} \, e^{-\frac{x^2}{2}} \tag{40}$$

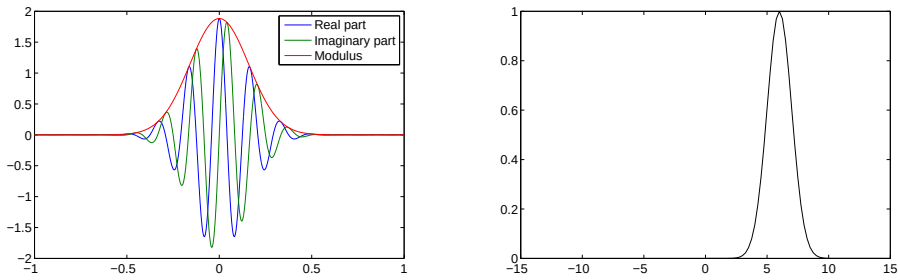
with the wavenumber k_ψ denoting the barycenter of the wavelet support in Fourier space given by

$$k_\psi = \frac{\int_0^\infty k|\widehat{\psi}(k)| \, dk}{\int_0^\infty |\widehat{\psi}(k)| \, dk} \quad . \tag{41}$$

The wavenumber k_ψ controls the number of oscillations inside the wavelet. Actually the Morlet wavelet does not *stricto sensus* respects the admissibility condition as defined in Eq. (18) since its mean is not zero. One should take $k_\psi > 5$ to insure that it vanishes up to the computer round-off errors. A better solution is to define the Morlet wavelet in Fourier space and enforce the admissibility condition by putting its mean, *i.e.*, $\widehat{\psi}(0)$, to zero which gives

$$\widehat{\psi}(k) \; = \; \left\{ \begin{array}{ll} e^{-\frac{(k-k_\psi)^2}{2}} & \text{for} \quad k > 0 \quad , \\ 0 \quad \text{for} \quad k \leq 0 \quad . \end{array} \right.$$

If the problem one would like to solve requires filtering or compressing a signal, an image or a vector field under study, then one should use the orthogonal wavelet transform to avoid the redundancy inherent to the continuous wavelet transform. In this case there is also a large collection of possible orthogonal wavelets and their choice depends on which properties one prefers, *e.g.*, compact-support, symmetry, smoothness, number of cancellations, computational efficiency.



(a) (b)

Figure 9: Localization of the Morlet wavelet in physical space (a) and in spectral space (b).

From our experience, we recommend the Coifman 12 wavelet, which is compactly supported, has four vanishing moments, is quasi-symmetric and is defined by a filter of length 12, which leads to a computational cost of the fast wavelet transform in $24N$ operations (since two filters are needed for the wavelet and the scaling function).

To analyze fluctuating signals or fields one should use the continuous wavelet transform with complex valued wavelets, since the modulus of the wavelet coefficients allows to read the evolution of the energy density in both space (or time) and scales. If one uses real-valued wavelets instead, the modulus of the wavelet coefficients will present the same oscillations as the analyzing wavelets and it will then become difficult to sort out features belonging to the signal or to the wavelet. In the case of complex-valued wavelets the quadrature between the real and the imaginary parts of the wavelet coefficients eliminates these spurious oscillations and this is why we recommend to use complex-valued wavelets, such as the Morlet wavelet. If ones wants to compress turbulent flows, and *a fortiori* to compute their evolution at a reduced cost compared to standard methods (finite difference, finite volume or spectral methods), one should use orthogonal wavelets. In this case there is no more redundancy of the wavelet coefficients and one has the same number of wavelet coefficients as the number of grid-points and one uses the fast wavelet transform [13, 19, 38]. The first application of wavelets to analyze turbulent flows has been published in 1988 [18]. Since then a long-term research program has been developed for analyzing, computing and modelling turbulent flows using either continuous wavelets or orthogonal wavelets, and also wavelet packets (one can download the corresponding papers from http://wavelets.ens.fr in 'Publications').

As an example we show the continuous wavelet transform, using the complex-valued Morlet wavelet, of several signals: a deterministic fractal which is the Devil's staircase (Fig. 10) and two self-similar random signals, which are, Fractional Brownian Motions (FBM) having different Hurst exponent, *i.e.*, $H = 0.25$ and $H = 0.75$ (Fig. 11).

3.3.2 Wavelet spectrum

Since the wavelet transform conserves energy and preserves locality in physical space, one can use it to extend the concept of the energy spectrum and define the local energy

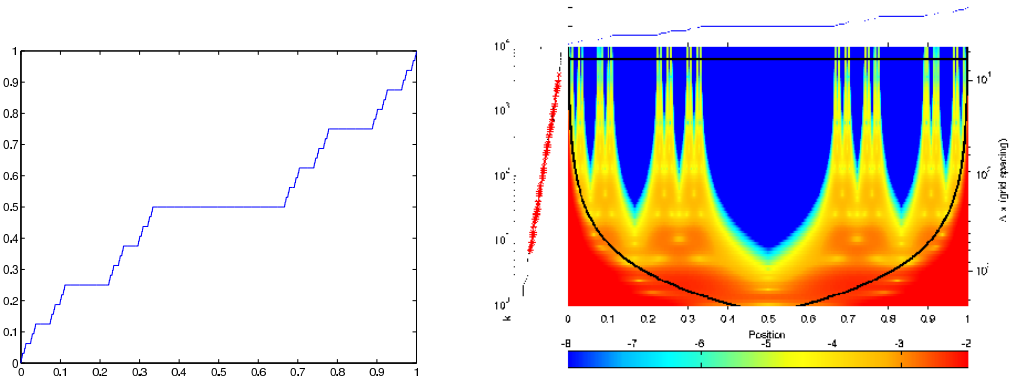


Figure 10: Devil's staircase (left), and its continuous wavelet analysis (right).

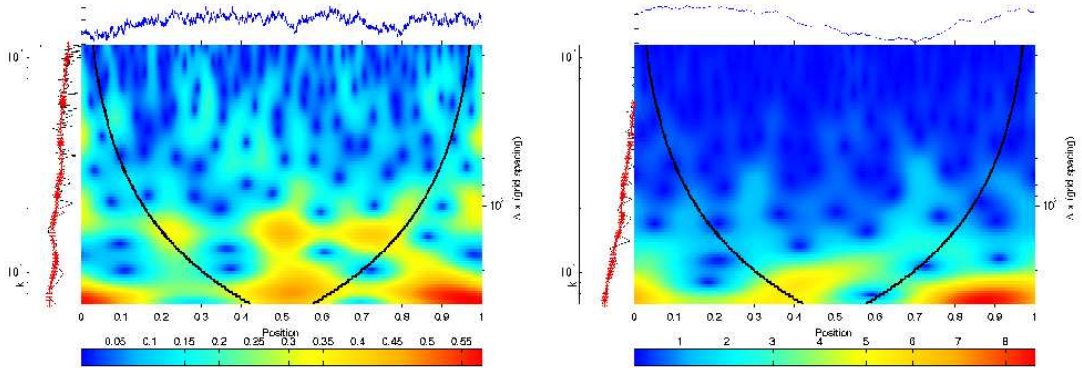


Figure 11: Continuous wavelet analysis of fractional Brownian motion with Hurst exponent $H = 0.25$ (left), and $H = 0.75$ (right).

spectrum of the function $f \in L^2(\mathbb{R})$, such that

$$\widetilde{E}(k,x) = \frac{1}{C_\psi k_\psi} \left| \widetilde{f}\left(\frac{k_\psi}{k}, x\right) \right|^2 \quad \text{for} \quad k \geq 0 \quad , \tag{42}$$

where k_ψ is the centroid wavenumber of the analyzing wavelet ψ and C_ψ is defined by the admissibility condition given in Eq. (18).

By measuring $\widetilde{E}(k,x)$ at different instants or positions in the signal, one estimates which elements in the signal contribute most to the global Fourier energy spectrum, that might suggest a way to decompose the signal into different components. One can split a given signal or field using the orthogonal wavelet transform into two orthogonal contributions (see section 3.3.5) and then plot the energy spectrum of each to exhibit their different spectral slopes and therefore their different correlation.

Although the wavelet transform analyzes the flow using localized functions rather than complex exponentials as for Fourier transform, one can show that the global wavelet energy spectrum approximates the Fourier energy spectrum provided the analyzing wavelet has enough vanishing moments. More precisely, the global wavelet spectrum, defined by integrating Eq. (42) over all positions,

$$\widetilde{E}(k) = \int_{-\infty}^{\infty} \widetilde{E}(k,x)\, dx \tag{43}$$

gives the correct exponent for a power-law Fourier energy spectrum $E(k)$ scaling as $k^{-\beta}$ if the analyzing wavelet has at least $M > \frac{\beta-1}{2}$ vanishing moments. Thus, the steeper the energy spectrum one would like to study, the more vanishing moments the analyzing wavelet should have. In practice one should choose first a wavelet with many vanishing moments and then reduce this number until the estimated slope varies. This will give the optimal wavelet to analyze the given function.

Relation to Fourier spectrum The wavelet energy spectrum $\widetilde{E}(k)$ is related to the Fourier energy spectrum $E(k)$ via,

$$\widetilde{E}(k) = \frac{1}{C_\psi k_\psi} \int_0^\infty E(k') \left| \widehat{\psi}\left(\frac{k_\psi k'}{k}\right) \right|^2 dk' \quad , \tag{44}$$

which shows that the wavelet spectrum is a smoothed version of the Fourier spectrum, weighted with the square of the Fourier transform of the wavelet ψ shifted at wavenumbers k. For increasing k, the averaging interval becomes larger, since wavelets are filters with constant relative bandwidth, *i.e.*, $\frac{\Delta k}{k}$ = constant. The wavelet energy spectrum thus yields a stabilized Fourier energy spectrum.

Considering for example the Marr wavelet given in Eq. (38), which is real-valued and has two vanishing moments only, the wavelet spectrum can estimate exponents of the energy spectrum for $\beta < 5$. In the case of the complex-valued Morlet wavelet given in Eq. (42), only the zeroth-order moment is vanishing. However higher m^{th}-order moments are very small ($\propto k_\psi^m e^{(-k_\psi^2/2)}$), provided that k_ψ is sufficiently large. For instance choosing $k_\psi = 6$ yields accurate estimates of the exponent of power-law energy spectra for at least $\beta < 7$.

There also exists a family of wavelets with an infinite number of vanishing moments

$$\widehat{\psi}_n(k) = \alpha_n \exp\left(-\frac{1}{2}\left(k^2 + \frac{1}{k^{2n}}\right)\right), \quad n \geq 1 \quad , \tag{45}$$

where α_n is a normalization factor. Wavelet spectra using this wavelet can thus correctly measure any power–law energy spectrum. This choice enables, in particular, the detection of the difference between a power–law energy spectrum and a Gaussian energy spectrum such that $E(k) \propto e^{-(k/k_0)^2}$. This is important in turbulence to determine the wavenumber after which the energy spectrum decays exponentially. The end of the inertial range, dominated by nonlinear interactions, and the beginning of the dissipative range, dominated by linear dissipation, can thus be detected.

Relation to structure functions Structure functions, classically used to analyze non stationary random processes, *e.g.*, turbulent velocity fluctuations, have some limitations which can be overcome using wavelet-based alternatives. Structure functions are defined by moments of increments of the random process. The latter can be interpreted as wavelet coefficients using a special wavelet, the difference of two Diracs (called DoD wavelet), which is very singular and has only one vanishing moment, namely its mean value. This unique vanishing moment of the DoD wavelet limits the adequacy of structure functions to analyze sufficiently smooth signals. Wavelets having more vanishing moments do not have this drawback.

For second order statistics, the classical energy spectrum, defined as the Fourier transform of the autocorrelation function is naturally linked to the second order structure function. Using the above relation of the wavelet spectrum to the Fourier spectrum a similar relation to second order structure functions can be derived. For structure functions yielding a power law behaviour the maximum exponent can be shown to be limited by the number of vanishing moments of the underlying wavelet.

The increments of a function $f \in L^2(\mathbb{R})$ are equivalent to its wavelet coefficients using the DoD wavelet

$$\psi^\delta(x) = \delta(x+1) - \delta(x) \quad . \tag{46}$$

We thus obtain

$$f(x+a) - f(x) = \widetilde{f}_{x,a} = \langle f, \psi^\delta_{x,a} \rangle \quad , \tag{47}$$

with $\psi^\delta_{x,a}(y) = 1/a[\delta(\frac{y-x}{a+1}) - \delta(\frac{y-x}{a})]$, where the wavelet is normalized with respect to the L^1-norm. The p-th order moment of the wavelet coefficients at scale a yields the p-th order structure function,

$$S_p(a) = \int (\widetilde{f}_{x,a})^p dx \quad . \tag{48}$$

As already mentioned above the drawback of the DoD wavelet is that it has only one vanishing moment, its mean. Consequently the exponent of the p-th order structure function in the case of a power law behaviour is limited by p, *i.e.*, if $S_p(a) \propto a^{\zeta(p)}$ then $\zeta(p) < p$. The detection of larger exponents necessitates the use of increments with a larger stencil, or wavelets with more vanishing moments.

We now focus on second order statistics, the case $p = 2$. Equation (44) yields a relation between the global wavelet spectrum $\widetilde{E}(k)$ and the Fourier spectrum $E(k)$ for a given wavelet ψ. Taking the Fourier transform of the DoD wavelet we get $\widehat{\psi^\delta}(k) = e^{\iota k} - 1 = e^{\frac{\iota k}{2}}(e^{\frac{\iota k}{2}} - e^{-\iota k 2})$ and therefore we have $|\widehat{\psi}^\delta(k)|^2 = 2(1 - \cos k)$. The relation between the Fourier and the wavelet spectrum thus becomes

$$\widetilde{E}(k) = \frac{1}{C_\psi k} \int_0^\infty E(k') \left(2 - 2\cos(\frac{k_\psi k'}{k})\right) dk' \quad , \tag{49}$$

and the wavelet spectrum can be related to the second order structure function by setting $a = k_\psi/k$

$$\widetilde{E}(k) = \frac{1}{C_\psi k} S_2(a) \quad . \tag{50}$$

Using now the result of section 3.3.2 that for a Fourier spectrum which behaves like $k^{-\alpha}$ for $k \to \infty$, the wavelet spectrum only yields $\widetilde{E}(k) \propto k^{-\alpha}$ if $\alpha < 2M + 1$, where M denotes the number of vanishing moments of the wavelet, we find for the structure function $S_2(a)$ that $S_2(a) \propto a^{\zeta(p)} = \left(\frac{k_\psi}{k}\right)^{\zeta(p)}$ for $a \to 0$ if $\zeta(2) \leq 2M$.

For the DoD wavelet we have $M = 1$, which explains why the second order structure function can only detect slopes smaller than 2, which corresponds to wavelet energy spectra with slopes being shallower than -3. This explains why the usual structure function gives spurious results for sufficiently smooth signals.

3.3.3 Detection and characterization of singularities

The possibility to evaluate the slope of the energy spectrum is an important property of the wavelet transform, related to its ability to characterize the regularity of the signal and detect isolated singularities [26, 30]. This is based on the fact that the local scaling of the wavelet coefficients is computed in L^1-norm, *i.e.*, with the normalization $c(l) = l^{-1}$ instead of $c(l) = l^{1/2}$ in Eq. (20).

If the function $f \in C^m(x_0)$, *i.e.*, if f is continuously differentiable in x_0 up to order m, then

$$[\tilde{f}(l, x_0)]_{l \to 0} \leq l^{m+1} l^{1/2} \tag{51}$$

The factor $l^{1/2}$ comes from the fact that to study the scaling in x_0 of the function f we must compute its wavelet coefficients in L^1-norm, instead of L^2, *i.e.*, with the normalization $c(l) = l^{-1}$ instead of $c(l) = l^{1/2}$ in Eq. (20).

If f has Hölder regularity α at x_0 (see Sec. 2.1.3), then

$$[\tilde{f}(l, x_0)]_{l \to 0} \approx C e^{i\Phi} l^\alpha l^{1/2} \tag{52}$$

Where Φ is the phase of the wavelet coefficients in x_0. The phases of the wavelet coefficients $\Phi(l, x)$ in wavelet coefficient space allow to localize the possible singularities of f since the lines of constant phase converge towards the locations of all the isolated singularities when $l \to 0$. If the function f presents few isolated singularities, their position x_0, their strength C, and their scaling exponent α can thus be estimated by the asymptotic behavior of $\tilde{f}(l, x_0)$, written in L^1-norm, in the limit l tending to zero. If, on the contrary, the modulus of the wavelet coefficient becomes zero at small scale around

x_0, then the function f is regular at x_0. This result is the converse of Eq. (51) but it only works for isolated singularities since it requires that in the vicinity of x_0 the wavelet coefficients remain smaller than those pointing towards x_0. Consequently its use is not applicable to signals presenting dense singularities. The scaling properties presented in this paragraph are independent of the choice of the analyzing wavelet ψ. Actually we recommend to use complex-valued wavelets since one thus obtains complex-valued wavelet coefficients whose phases locate the singularities while their moduli estimate the Hölder exponents of all isolated singularities, as illustrated in Fig. 10. We can then compute the singularity spectrum (see section 2.1.3).

3.3.4 Intermittency measures

Localized bursts of high frequency activity define typically intermittent behaviour. Localization in both physical space and spectral space is thus implied and a suitable basis for representing intermittency should reflect this dual localization. The Fourier representation yields perfect localization in spectral space, but global support in physical space. Filtering a fluctuating signal with an ideal high-pass Fourier filter implies some loss of spatial information in physical space. Strong gradients are smoothed out and spurious oscillations occur in the background. This comes from the fact that the modulus and phase of the discarded high-wavenumber Fourier modes have been lost. The artefacts of Fourier filtering lead to errors in estimating the flatness, and hence the signal's intermittency.

An intermittent quantity (*e.g.*, velocity derivative) contains rare but strong events (*i.e.*, bursts of intense activity), which correspond to large deviations reflected in 'heavy tails' of the probability distribution function of that quantity. Second-order statistics (*e.g.*, energy spectrum, second-order structure function) are not very sensitive to such rare events whose spatial support is too small to play a role in the integral. For higher-order statistics, however, these rare events become increasingly important, may eventually dominate and thus allow to detect intermittency. Of course, not for all problems intermittency is essential, *e.g.* second-order statistics are sufficient to measure dispersion (dominated by energy-containing scales), but not to calculate drag or mixing (dominated by vorticity production in thin boundary or shear layers).

Using the continuous wavelet transform we have proposed the local intermittency measure [19, 48] which corresponds to the wavelet coefficients renormalized by the space averaged energy at each scale, such that

$$I(l,\vec{x}) = \frac{|\tilde{f}(l,\vec{x})|^2}{\int_{-\infty}^{\infty} \tilde{f}(l,\vec{x})|^2 d^2\vec{x}} \quad . \tag{53}$$

It yields information on the spatial variance of energy as a function of scale and position. For regions where $I(l,\vec{x}) \approx 1$ the field is non intermittent while regions of larger values are intermittent.

Similarly to the continuous wavelet transform the orthogonal wavelet transform allows to define intermittency measures, either local as shown above, or global as illustrated below. The space-scale information contained in the wavelet coefficients yields

suitable global intermittency measures using scale-dependent moments and moment ratios [49]. For a signal f the moments of wavelet coefficients at different scales j are defined by

$$M_{p,j}(f) = 2^{-j} \sum_{i=0}^{2^j-1} \left(\widetilde{f}_{j,i}\right)^p \quad . \tag{54}$$

The scale distribution of energy, *i.e.*, the scalogram, is obtained from the second order moment of the orthogonal wavelet coefficients: $E_j = 2^{j-1} M_{2,j}$. The total energy is then recoved by the sum: $E = \sum_{j\geq 0} E_j$ thanks to the orthogonality of the decomposition.

Ratios of moments at different scales quantify the sparsity of the wavelet coefficients at each scale and thus measure the intermittency

$$Q_{p,q,j}(f) = \frac{M_{p,j}(f)}{(M_{q,j}(f))^{\frac{p}{q}}} \quad , \tag{55}$$

which correspond to quotient of norms computed in two different sequence spaces, l^p- and l^q-spaces. Typically, one chooses $q = 2$ to define statistical quantities as a function of scale. For $p = 4$ we obtain the scale dependent flatness $F_j = Q_{4,2,j}$ which equals 3 for a Gaussian white noise at all scales j. and indicates that a signal is not intermittent. Scale dependent skewness, hyperflatness and hyperskewness are defined for $p = 3, 5$ and 6, respectively. Intermittency of a signal is relfected in increasing $Q_{p,q,j}$ for increasing j (smaller scale) supposing $p > q$.

3.3.5 Extraction of coherent structures

To study fluctuating signals or fields we need to separate the rare and extreme events from the dense events, and then calculate their statistics independently for each one. For this we cannot use pattern recognition methods since there is no simple patterns to characterise them. Moreover there is no clear scale separation between the rare and the dense events and therefore a Fourier filter cannot disentangle them. Since the rare events are well localized in physical space, one might try to use an on-off filter defined in physical space to extract them. However, this approach changes the spectral properties by introducing spurious discontinuities, adding an artificial scaling (*e.g.*, k^{-2} in one dimension) to the energy spectrum. The wavelet representation can overcome these problems since it combines both physical and spectral localizations (bounded from below by the uncertainty principle).

We have proposed in 1999 [20] a better approach to extract rare events out of fluctuating signals or fields which is based on the orthogonal wavelet representation. We rely on the fact that rare events are localized while dense events are not, and we assume that the later are noise-like. From a mathematical viewpoint a noise cannot be compressed in any functional basis. Another way to say this is to observe that the shortest description of a noise is the noise itself. Note that one often calls 'noise' what actually is 'experimental noise', *i.e.*, something that one would like to discard although it may not be noise-like in the above mathematical sense. The problem of extracting the rare events has thus become the problem of denoising the signal or the field under study. Assuming that they are are what remains after denoising, we need a model, not for

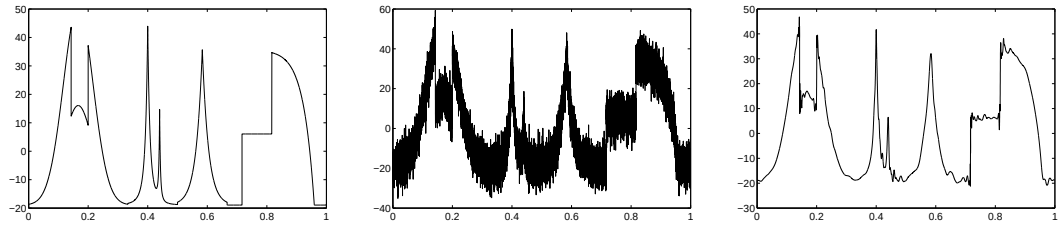


Figure 12: Academic example of denoising of a piecewise regular signal using the algorithm for coherent structure extraction. Original signal (left), same signal plus Gaussian white noise giving a signal to noise ratio (SNR) of $11.04dB$ (middle), denoised signal with SNR of $27.55dB$ (right).

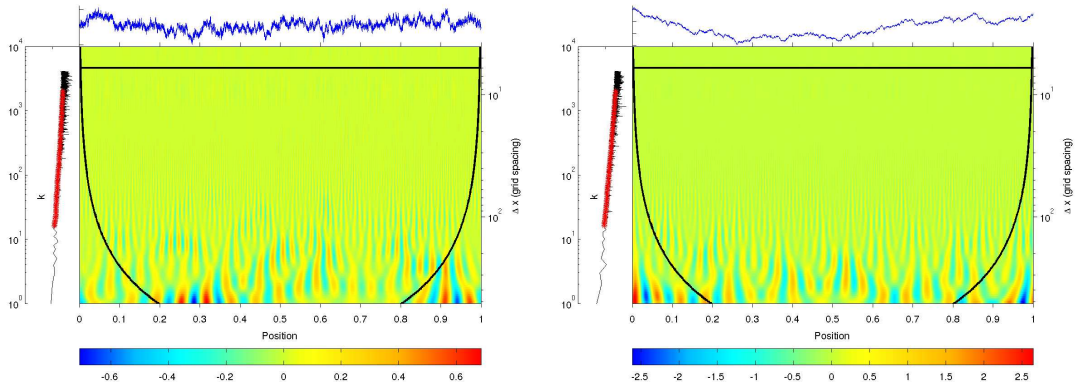


Figure 13: Real part of the continuous wavelet analysis of fractional Brownian motion with Hurst exponent $H = 0.25$ (left), and of classical Brownian motion (right).

the rare events, but for the noise. As a first guess, we choose the simplest model and suppose the noise to be additive, Gaussian and white, *i.e.*, uncorrelated.

We now describe the wavelet algorithm for extracting coherent structures out of a signal corrupted by a Gaussian noise with variance σ^2 and vanishing mean, sampled on N equidistant grid points. The noisy signal $f(x)$ is projected onto orthogonal wavelets using Eq. (30) to get $\widetilde{f}_\lambda$. Its wavelet coefficients are then split into two sets, those whose modulus is larger than a threshold ε that we call 'coherent', and those remaining that we call 'incoherent'. The threshold value, based on minmax statistical estimation [14], is $\varepsilon = (2/d\sigma^2 \ln N)^{1/2}$, where d is the space dimension. Note that besides the choice of the wavelet there is no adjustable parameter since σ^2 and N are known *a priori*. In case the variance of the noise is unknown, one estimates it recursively from the variance of the incoherent wavelet coefficients, as proposed in [3]. The convergence rate increases with the signal to noise ratio, namely if there is only noise it converges in zero iteration. The coherent signal f_C is reconstructed from the wavelet coefficients whose modulus is larger than ε and the incoherent signal f_I from the remaining wavelet coefficients. The two signals thus obtained, f_C and f_I, are orthogonal.

To illustrate the method we choose an academic signal (Fig. 12, left) which is a superposition of several quasi-singularities having different Hölder exponents, to which we have superimposed a Gaussian white noise yielding a signal to noise ratio of $11.04dB$ (Fig. 12, middle). Applying the extraction method we recover a denoised version of the corrupted signal which preserves the quasi-singularities (Fig. 12, right). It could be checked *a posteriori* that the incoherent contribution is spread, and therefore does not compress, and has a Gaussian probability distribution.

4 Recommendations

In the introduction we stated cautious remarks about the risk of misusing new mathematical tools, if one has not first gained enough practice on academic examples. The problem is the following. When doing research the questions one addresses are still open and there exist several competing theories, models and interpretations. Nothing being clearly fixed yet, neither the comprehension of the physical phenomenon under study, nor the practice of the new techniques in use, one runs the risk to perform a Rorschach's test rather than a rational analysis. Indeed, the interpretation of the results may reveal one's unconscious desire for a preferred explanation. Although it is a good thing to rely on one's intuition and have a preferred theory, one should be conscious of that risk, and make sure to avoid bias. Moreover, when a new technique is proposed, most of referees do not master it yet and are therefore not able to detect flaws in a submitted paper.

Let us take as example the case of turbulence, which has applications in everyday life and plays an important role in environmental fluid dynamics. For centuries, turbulence has been an open problem and thus a test ground for new mathematical techniques. Let us focus here on the case of fractals and wavelets, as they were applied to study turbulence. Kolmogorov's statistical theory of homogeneous and isotropic turbulence [33] assumes that there exists an energy cascade from large to small scales, which is modelled as a self-similar stochastic process whose spectrum scales as $k^{-5/3}$, where k is the wavenumber. Although this prediction only holds for an ensemble average of many flow realisations, many authors interpret the energy cascade as caused by the successive breakings of whirls into smaller and smaller ones, as if they were stones. This interpretation was inspired by a comment Lewis Fry Richardson made in 1922: *When making a drawing of a rising cumulus from a fixed point, the details change before the sketch was completed. We realize thus that: big whirls have little whirls that feed on their velocity, and little whirls have lesser whirls and so on to viscosity– in the molecular sense*[47]. We think that Richardson's quote has been misunderstood and turbulence misinterpreted. Indeed, his remark concerns the interface between a cumulus cloud and the surrounding clear air, which is a very convoluted two-dimensional surface developing into a three-dimensional volume. Such an interface may develop into a fractal since its topological dimension is lower than the dimension of the space which contains it. But keeping such a fractal picture to describe three-dimensional whirls which evolve inside a three-dimensional space does not make sense since both have the same topological dimension. In 1974 Kraichnan was already suspicious about this interpretation, when he wrote: *'The terms 'scale of motion' or 'eddy of size l' appear repeatedly in the treatment of the inertial range. One gets an impression of little, randomly distributed whirls in the fluid, with the fission of the whirls into smaller ones,*

after the fashion of Richardson's poem. This picture seems to be drastically in conflict with what can be inferred about the qualitative structures of high Reynolds numbers turbulence from laboratory visualization techniques and from plausible application of the Kelvin's circulation theorem' [34]. Unfortunately Kraichnan's viewpoint was not taken into account and, on the contrary, the picture of breaking whirls was even reinforced by the terminology fractals due to its Latin root *fractare* (to break). This gave rise to numerous models of turbulence which were based on fractals, and later on multi-fractals (for a review of them see [21]).

Let us now consider the use of wavelets to analyze turbulent flows and illustrate the risk of misinterpretation there too. If one performs the continuous wavelet analysis of any fluctuating signals, for example the temporal fluctuations of one velocity component of a three-dimensional turbulent flow, one should be very cautious, especially when using a real-valued wavelet. Indeed, for this class of noise-like signals one observes a tree-like pattern in the two-dimensional plot of their wavelet coefficients which is generic to the continuous wavelet transform and corresponds to its reproducing kernel [19]. When one performs the continuous wavelet transform of one realization of a Gaussian white noise one observes such a pattern (see Fig. 13), which proves that the correlation is among the wavelets but not in the signal itself. Unfortunately in the case of turbulent signals, this pattern has been interpreted as the evidence of whirls breaking in a paper published in 1989 by Nature under the title 'Wavelet analysis reveals the multifractal nature of the Richardson's cascade' [2].

Let Benoît Mandelbrot concludes: *'In the domain I know of, there are many words which are meaningless, that do not have any content, which have been created just to impress, to give the feeling that a domain exists when actually there is none. If one gives a name to a science, this science maybe does not exist. And, once more, due to the fierce discipline I was imposing to myself, I avoided that [...]. Therefore I have created the word "fractal" with much reflection. The idea was that of objects which are dispersed, which are broken into small pieces.'*[41] The question remains for us: are fractals a new science or only consist of refurbishing older concepts to launch a new fashion? In the same vein, Yves Meyer wrote: *Wavelets are fashionable and therefore excite curiosity and irritation. It is amazing that wavelets have appeared, almost simultaneously in the beginning of the* 80*'s, as an alternative to traditional Fourier analysis, in domains as diverse as speech analysis and synthesis, signal coding for telecommunications, (low-level) information, extraction process performed by the retinian system, fully-developed turbulence analysis, renormalization in quantum field theory, functional spaces interpolation theory... But this pretention for pluridisciplinarity can only be irritating, as are all "great syntheses" which allow one to understand and explain everything. Will wavelets soon join "catastrophe theory" or "fractals" in the bazaar of all-purpose systems?* [44] Let the future tells us the answer...

Acknowledgments

We thank Barbara Burke, John Hubbard and Rodrigo Pereira for useful comments. The authors are very thankful to CEMRACS (Centre d'Eté de Recherche Avancée en Calcul Scientifique) and CIRM (Centre International de Rencontres Mathématiques),

Marseille, France, for their hospitality while writing this paper. M.F. , R.N.V.Y. and O.P. acknowledge financial support from the ANR (Agence Nationale pour la Recherche) within the GeoFluids program. M.F. is grateful to the Wissenschaftskolleg zu Berlin for its hospitality while writing this paper. M.F., K.S. and R.N.V.Y. thankfully acknowledge financial support from the PEPS program of INSMI-CNRS. They also thank the Association CEA-EURATOM and the FRF2S (French Research Federation for Fusion Studies) for supporting their work within the framework of the EFDA (European Fusion Development Agreement) under contract V.3258.001. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] J. P. Antoine, R. Murenzi, P. Vandergheynst, and S. T. Ali, 2004. *Two-dimensional wavelets and their relatives*. Cambridge University Press.

[2] F. Argoul, A. Arnéodo, G. Grasseau, Y. Gagne, E. J. Hopfinger and U. Frisch, 1989. Wavelet analysis of turbulence reveals the multifractal nature of the Richardson cascade. *Nature*, **338**, 51–53.

[3] A. Azzalini, M. Farge and K. Schneider, 2005. Nonlinear wavelet thresholding : A recursive method to determine the optimal denoising threshold. *Appl. Comput. Harm. Analysis*, **18** (2), 177–185.

[4] R. Balian, 1981. Un principe d'incertitude en théorie du signal ou en mécanique quantique. *C.R.A.S. Paris*, **292**, srie 2, 1357–1361.

[5] M. Barnsley, 1988. *Fractals everywhere*. Academic Press.

[6] B. Burke Hubbard, 1998. *The world according to wavelets*. A. K. Peters.

[7] R.R. Coifman and Y. Meyer, 1991. Remarques sur l'analyse de Fourier à fenêtre. *C. R. Acad. Sci. Paris*, **312**, serie I, 259–261.

[8] R.R. Coifman and V. M. Wickerhauser, 1992, Entropy based algorithms for best basis selection. *IEEE Trans. Information Theory*, **32**, 712–718.

[9] R. L. Cook, T. DeRose, 2005. Wavelet noise. *ACM Transactions on Graphics*, **24**(3), 803–811.

[10] M. S. Crouse and R. G. Baraniuk, 1999. Fast, exact synthesis of Gaussian and non Gaussian long range dependent processes. *IEEE Trans. Inform. Theory*, submitted.

[11] I. Daubechies, A. Grossmann and Y. Meyer, 1986. Painless nonorthogonal expansions. *J. Math. Phys.*, **27**, 1271–1283.

[12] I. Daubechies, 1988 Orthonormal bases of compactly supported wavelets *Comm. Pure Appl. Math.*, **41**(7), 909–996

[13] I. Daubechies, 1992. *Ten lectures on wavelets.* SIAM, Philadelphia.

[14] D. Donoho and I. Johnstone, 1994. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, **81**(3), 425–455.

[15] A. Douady and J. H. Hubbard, 1982. Iterations des polynomes quadratiques complexes, *C. R. Acad. Sc. Paris*, **294**, 123–126.

[16] F.W. Elliott, and A.J. Majda, 1994. A wavelet Monte-Carlo method for turbulent diffusion with many spatial scales. *J. Comput. Phys.*, **113**(1), pp. 82–111.

[17] F.W. Elliott, and A.J. Majda, 1995. A new algorithm with plane waves and wavelets for random velocity fields with many spatial scales. *J. Comput. Phys.*, **117**(1), pp. 146–162.

[18] M. Farge & G. Rabreau, 1988. Transformée en ondelettes pour détecter et analyser les structures cohérentes dans les écoulements turbulents bidimensionnels. *C. R. Acad. Sci. Paris*, **2**(307), 1479–1486.

[19] M. Farge, 1992. Wavelet transforms and their applications to turbulence. *Ann. Rev. Fluid Mech.*,**24**, 395–457.

[20] M. Farge, K. Schneider and N. Kevlahan, 1999. Non-Gaussianity and Coherent Vortex Simulation for two-dimensional turbulence using an orthonormal wavelet basis. *Phys. Fluids*, **11**(8), 2187–2201.

[21] U. Frisch, 1995. *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press.

[22] D. Gabor, 1946. Theory of Communication. *J. Inst. Electr. Engin.*, **93** (3), 429–457.

[23] J.-F. Gouyet, 1996. *Physics and fractal structures. Springer-Verlag*, Berlin, and Masson, Paris.

[24] A. Grossmann and J. Morlet, 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Appl. Anal.*, **15**(4), 723–736.

[25] A. Haar, 1910. Zur Theorie der orthogonalen Funktionensysteme. *Math. Ann.*, **69**(3), 331–371

[26] M. Holschneider, 1988. On the wavelet transform of fractal objects. *J. Stat. Phys.*, **50**(5/6), 963–996.

[27] Interview de John Hubbard, *Science Publique, France-Culture*, October 2010.

[28] G. A. Hunt., 1951. Random Fourier transforms. *Trans. Amer. Math. Soc.*, **71**, 38–69.

[29] H. E. Hurst, 1951. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. of Civil Engineers*, **116**, 770–808.

[30] S. Jaffard, 1989. Construction of wavelets on open sets. *First International Conference on Wavelets, Marseille, 14-18 December 1987* (eds. J.M. Combes, A. Grossmann and P. Tchamitchian), Springer, 247–252.

[31] H. von Koch, 1906. *Sur une courbe continue sans tangente, obtenue par une construction géomètrique élémentaire pour l'étude de certaines questions de la théorie des courbes planes.* Acta Math., **30**, 145–174.

[32] A. N. Kolmogorov, 1940. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C. R. (Doklady) Acad. Sci. SSSR*,**26**, 115–118.

[33] A. N. Kolmogorov, 1941. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proc. USSR Acad. Sci.*, **30**, 299-303 (Russian), translation in *Proc. Royal Soc., Series A: Mathematical and Physical Sciences*, **434**, 9-13 (1991).

[34] R. H. Kraichnan, 1974. On Kolmogorov's inertial-range theories. *J. Fluid Mech.*, **62**, 305–330

[35] P.G. Lemarié and Y. Meyer, 1986. Ondelettes et bases Hilbertiennes. *Rev. Mat. IberoAm.*, **2**, 1–18.

[36] P. Lévy. *Le mouvement brownien.* Fascicule CXXVI of Mémorial des sciences mathématiques. Gauthier-Villars, Paris, 1954.

[37] S. Mallat, 1989. Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$. *Trans. Amer. Math. Soc.*, **315**, 69–87.

[38] S. Mallat, 1998. *A wavelet tour of signal processing.* Academic Press.

[39] Malvar H., 1990. Lapped transforms for multiresolution signal decomposition: the wavelet decomposition *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11, 674–693.

[40] B. B. Mandelbrot and J. W. van Ness, 1968. Fractional Brownian motions, fractional noises and applications *SIAM Rev.*, **10**(4), 422–437.

[41] Interview de Benoît Mandelbrot. *A voix nue, France-Culture*, October 1990.

[42] B. B. Mandelbrot, 1975. *Les objets fractals.* Flammarion, Paris.

[43] B. B. Mandelbrot, 1977. *Fractals: Form, chance and dimension.* Freeman, San Fransisco.

[44] Y. Meyer, 1990. Ondelettes et applications. *J. Ann. Soc. Math. SMF, Paris*, 1–15.

[45] Murenzi, R., 1990. *Ondelettes multidimensionnelles et application l'analyse d'images.* Thèse de Doctorat de l'Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 1990.

[46] G. Parisi and U. Frisch, 1985. On the singularity structure of fully developed turbulence. *Turbulence and Predictability in Geophysical Fluid Dynamics* (ed. M. Ghil, R. Benzi and G. Parisi), 84–87, North-Holland.

[47] L. F. Richardson, 1922. *Weather prediction by numerical process.* Cambridge University Press.

[48] J. Ruppert-Felsot, M. Farge and P. Petitjeans, 2009. Wavelet tools to study intermittency: application to vortex bursting. *J. Fluid Mech.*, 636, 427–453.

[49] K. Schneider, M. Farge and N. Kevlahan, 2004. Spatial intermittency in two-dimensional turbulence. *Woods Hole Mathematics, Perspectives in Mathematics and Physics* (eds. N. Tongring and R.C. Penner), World Scientific, 302–328.

[50] K. Schneider and O. Vasilyev., 2.010 Wavelet methods in computational fluid dynamics. *Annu. Rev. Fluid Mech.*, **42**, 473–503.

[51] P. Tafti and M. Unser, 2010. Fractional Brownian vector fields. SIAM Multiscale Model. Sim., **8**(5), 1645–1670.