23

Multiscale Representations: Fractals, Self-Similar Random Processes and Wavelets

Marie Farge
École Normale Supérieure

Kai Schneider
Aix-Marseille University

Olivier Pannekoucke
*National Centre
for Meteorological Research*

Romain Nguyen van yen
École Normale Supérieure

23.1 Introduction

Many growth processes that shape the human environment generate structures over a wide range of scales, e.g., trees, rivers, lightning 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 that give rise to chaotic behavior, 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 modeling environmental fluid flows be solved, thanks to the development of numerical methods and the advent of supercomputers. 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 that presents structures that look 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 toward small scales but on the contrary remains quite similar to which 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 polygons, 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 and of dissipation, 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. [28]

Fractals can be traced back to the discovery of continuous nondifferentiable functions, e.g., the Weierstrass function, and nonrectifiable curves, e.g., the Sierpinski gasket. Measure theory, as developed in particular by Felix Hausdorff at the end of the nineteenth century, and integration theory, as redesigned by Henri Lebesgue and others at the beginning of the twentieth century, together with the study of recursive sequences in the complex plane, by Pierre Fatou and Gaston Julia, were all precursors of fractals, although a different terminology was used those years. Only when computer graphics became widely available in the 1960s was one able to visualize fractals and wonder about their apparent complexity. Although the mathematical tools were already there, it is Benoît Mandelbrot, while working for the IBM Research Center in Yorktown (USA), who popularized fractals and named them in the 1970s. Actually, before he started talking about fractals, Mandelbrot was a specialist of the theory 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 [38]. It was Mandelbrot who gave in 1968 the name "fractional Brownian motion" [42] to the self-similar stochastic processes proposed by Kolmogorov in 1940 [33], which are generalizations to 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 1980s by Jean Morlet and Alex Grossmann. Jean Morlet was researching on oil exploration for the French company Elf, while Alex Grossmann was a specialist of coherent states in quantum mechanics and a member of the CPT (Centre de Physique Théorique) in Marseille (France) (see also [21] for more on the early history of wavelets). 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 led 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 chapter is to give researchers working in environmental fluid dynamics some mathematical tools to study the multiscale behavior 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 behavior. In this paper we will present several multiscale methods developed from three different view points: fractals, self-similar random processes, and wavelets. All of them are mathematical tools that do not have any explanatory power per se. They require the scientist who uses 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 induces an a priori interpretation, built in within it without the user being a ware of that. To avoid such a drawback, we will here limit ourselves to give definitions, expose methods, and illustrate their use on academic examples rather than from applications. We will justify this choice in the conclusion by showing how such misinterpretation already happened while applying either fractals or wavelets to study turbulence.

23.2 Principles

23.2.1 Fractals

23.2.1.1 Definition and History

To define what "fractal" means is quite a difficult endeavor since one finds in the literature different definitions. Here we use the following definition: a fractal is a geometrical object that is so convoluted, irregular, or fragmented that its length, surface, volume, as well as higher-dimensional generalizations of these measures, all equal either zero or infinity. Its boundary is a set of points, either connected or disconnected, which looks the same at different scales and tends to be space-filling. For instance, a fractal curve is not rectifiable, i.e., its length is infinite. If the points remain connected the boundary can be parametrized by a continuous but nondifferentiable function. Otherwise, the fractal is a dust of disconnected points that can only be parametrized by a measure. A fractal shape may look complicated although it is not, since it may have been generated by a simple iterative procedure. The difficulty is, given an observed complicated shape, can we infer the simple rule that 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 [44] and 2 years later in English [45], 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 [44]. Subsequently, Mandelbrot succeeded in gathering under the same name different mathematical objects that 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" [7].

An example of such entertaining mathematical object was the fractal curve known as the "snow flake," see Figure 23.1a, published in 1904 by Helge von Koch in the Swedish journal "Arkiv for Matematik" [32].

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 the set of all cs for which $0 \in K_c$ is connex, and they baptized it the Mandelbrot set M in order 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 piecewise regular curves (e.g., triangle and other polygons). He illustrated that with many examples [44,45] such as the length of the coast of Britain, fluctuations of stock exchange, flood data, etc.

23.2.1.2 Fractal Dimension

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

$$N(l) \underset{l \to 0}{\sim} l^{-d}, \qquad (23.1)$$

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 Equation 23.1 can be extended to more general sets, for which d is in general no more an integer. These sets are thus fractal sets and d is called their fractal dimension. A more rigorous definition of the fractal dimension relies on the Hausdorff dimension [24]. But the latter is less easy to compute from data, and, in all the examples we shall consider thereafter, 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 by Equation 23.1.

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 subsegment are retained; this construction is illustrated in Figure 23.1a. 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 subsegments of the 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 as follows:

$$d_C = \lim_{n \to \infty} \left(-\frac{\ln 2^n}{\ln 3^{-n}} \right) = \frac{\ln 2}{\ln 3}. \qquad (23.2)$$

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
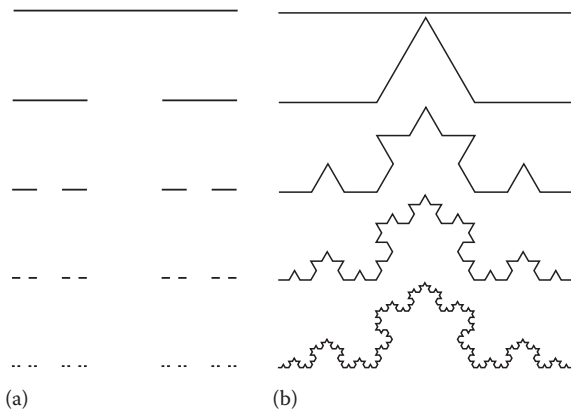

(a)   (b)

FIGURE 23.1 Illustration of the first four iteration steps leading to the Cantor dust (a) and to the von Koch snow flake (b).

replaced by four segments of length $l/3$ as illustrated in Figure 23.1b. 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 as follows:

$$d_K = \lim_{n \to \infty} \left(-\frac{\ln 4^n}{\ln 3^{-n}} \right) = \frac{\ln 4}{\ln 3}. \tag{23.3}$$

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.

23.2.1.3 Hölder Exponent and Singularity Spectrum

The fractal dimension was defined earlier 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

$$|f(x+l) - P_n(l)| \le K |l|^\alpha, \tag{23.4}$$

where n is the integer part of α (i.e., $n \le \alpha < n + 1$). 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 $P_n(l) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} l^k$, the Taylor expansion of f in x. The smaller the Hölder exponent, the stronger the singularity (as illustrated by the examples in Figure 23.2).
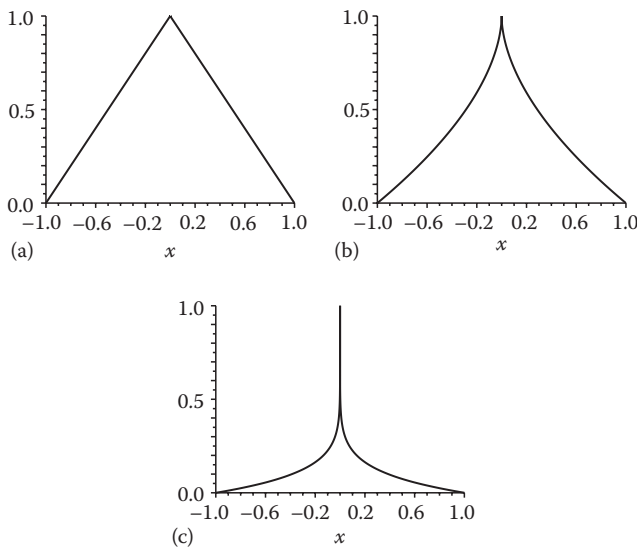
(a)

(b)

(c)

FIGURE 23.2 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).

Some functions, sometimes called multifractal functions [48], have a Hölder regularity that 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 that 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

$$|f(x+l) - f(x)| \sim l^{\alpha_x}, \tag{23.5}$$

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

$$N_{A_\alpha}(l) \sim l^{-d(\alpha)}. \tag{23.6}$$

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 Equation 23.6. 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 [48] the singularity spectrum $d(\alpha)$ appears as being the Legendre–Fenchel inverse transform of the multiscale exponent $\tau(q)$

$$d(\alpha) = \inf_\alpha \{q\alpha - \tau(q)\}. \tag{23.7}$$

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 as follows. In the Cantor set generation algorithm that we have described earlier, each interval was split into two pieces in a symmetric fashion at each iteration. Denoting by μ the characteristic function of the set obtained after n iterations of the procedure, one can easily show that the limit $f(x) = \lim_{n \to \infty} \int_0^x \chi_n(u) du$ exists, and the resulting function f, shown in Figure 23.3a, is called the Devil's staircase. 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. Therefore, in that case the singularity spectrum is reduced to the point $d(\ln 2/\ln 3) = \ln 2/\ln 3$.
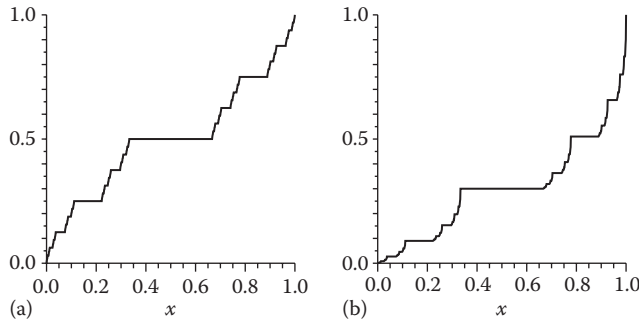
FIGURE 23.3 Illustration of the Devil's staircase with a homogeneous repartition of mass (a) and a heterogeneous repartition of mass where each left subsegment receives 30% of the mass (b).

More complex singularity spectra can be obtained by considering more general fractals similar to the Devil's staircase (see Figure 23.3b), which we do not detail here.

23.2.2 Self-Similar Random Processes

23.2.2.1 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

$$p(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \left(\frac{(\xi - \mu)^2}{2\sigma^2} \right), \tag{23.8}$$

where
  μ is the mean
  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 transform of its autocovariance function:

$$E(f) = \int_{\mathbb{R}} \langle \xi(\tau)\xi(0) \rangle e^{-\imath 2\pi\tau f} d\tau = \left\langle |\hat{\xi}(f)|^2 \right\rangle \text{ with}$$
$$\hat{\xi}(f) = \int \xi(t) e^{-\imath 2\pi t f} dt \quad \text{and} \quad \imath = \sqrt{-1}. \tag{23.9}$$

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 [4]. 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 nineteenth century observed the random motion of pollen suspension in water [37]. An extension of Brownian motion has been introduced by Kolmogorov in 1940 [33], of which a spectral representation was given by Hunt in 1951 [29] and that Mandelbrot proposed in 1968 to call fractional Brownian motion [42].

23.2.2.2 Brownian Motion

For Brownian motion the variance of the increments scales as

$$\langle |B(t) - B(\tau)|^2 \rangle = |t - \tau| \tag{23.10}$$

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 nonconstant 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 nonstationarity 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.

23.2.2.3 Fractional Brownian Motion

Fractional Brownian motion is a kind of self-similar Gaussian process that is nonstationary 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

$$B_H(t = 0) = 0 \tag{23.11}$$

and

$$\langle |B_H(t) - B_H(\tau)|^2 \rangle = |t - \tau|^{2H}, \tag{23.12}$$

where $0 < H < 1$ is an additional parameter called Hurst exponent [30]. 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

$$\langle B_H(t)B_H(\tau)\rangle = \frac{1}{2}(|t|^{2H} + |\tau|^{2H} - |t-\tau|^{2H}) \qquad (23.13)$$

Note that one given realization of fractional Brownian motion is not a fractal: the self-similarity is only fulfilled in the statistical sense. Indeed, Equation 23.12 implies that

$$\langle |B_H(\lambda t) - B_H(\lambda \tau)|^2\rangle = \lambda^{2H}\langle |B_H(t) - B_H(\tau)|^2\rangle. \qquad (23.14)$$

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 $2H + 1$,

$$E(f) = \frac{C_H}{f^{2H+1}}. \qquad (23.15)$$

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

$$B_H(t) = \int_{\mathbb{R}} \sqrt{E(f)}(e^{\imath 2\pi ft} - 1)d\xi(f), \qquad (23.16)$$

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 $(e^{\imath 2\pi ft} - 1)$ instead of $e^{\imath 2\pi ft}$ guarantees that $B_H(0) = 0$.

23.2.2.4 Multifractional Brownian Motion

Allowing for time (or space) varying Hurst exponents generalizes fractional Brownian motion, which is mono-fractal, to introduce stochastic multifractals. Such multifractional Brownian processes can be defined by generalizing the spectral representation, Equation 23.16, as follows:

$$B_\theta(t) = \int_{\mathbb{R}} \frac{e^{i2\pi ft} - 1}{|f|^{\theta(t)+1/2}} d\xi(f), \qquad (23.17)$$

where θ is a function $\theta: [0,1] \rightarrow]0,1[$ which can be seen as a local Hurst exponent of the process. Indeed, 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 $2 - \inf\{\theta(t), 0 \le t \le 1\}$.

Methods for synthesizing fractional Brownian motion are presented in Section 23.3.2.

23.2.3 Wavelets

23.2.3.1 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 on the spatial location of each wavenumber. This is due to the fact that, since the Fourier representation spreads time or space information among the phase of all Fourier coefficients, the energy spectrum (i.e., the modulus of the Fourier coefficients) does not carry any structural information in time or space. This is a major limitation of the classical way to analyze nonstationary 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 will only depend 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 $\hat{f}(k)$ defined as

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{-2\pi \imath kx}dx, \qquad (23.18)$$

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

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 analyze the signal f from both sides of the Fourier transform at the same time, due to the restriction $\Delta x \cdot \Delta k \ge C$, where Δx is the spatial support of $|f(x)|$ and Δk the spectral support of $|\hat{f}(k)|$, with C a constant that 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 losing 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 analyzed and processed.

In order to try to recover some space locality while using the Fourier transform, Gabor [23] has proposed the windowed Fourier transform, which consists of convolving the signal with a set of Fourier modes $e^{2\pi \imath kx}$ localized in a Gaussian envelope of constant width l_0. This transform allows then a space-wavenumber (or time-frequency) decomposition of the signal at a given scale l_0, which is kept fixed. But unfortunately, as shown by Balian [5], the bases constructed with such windowed Fourier modes cannot be orthogonal. In 1984, Grossmann and Morlet [25] 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 analyzing 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 the product $\Delta x \cdot \Delta k$ remains constant in the process of dilating and translating ψ. 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 analyze and synthesize signals or fields in higher dimensions [1,47]. In 1985, Meyer, while trying to prove the same kind of impossibility to build orthogonal bases as done by Balian [5] 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 [36]. In fact the Haar orthogonal basis, which was proposed in 1909 in the PhD thesis of Haar and published in 1910 [26], 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 often likes to build orthogonal wavelet bases in which the expansion of some signals of interest (depending on the application) are sparse, i.e., involve as few large wavelet coefficients as possible, while the rest are negligible. In particular, following Meyer's work, Daubechies has proposed in 1988 [12] orthogonal wavelet bases built with compactly supported functions defined by discrete Quadrature Mirror Filters (QMFs) of different lengths. The longer the filter, the sparser is the expansion of smooth signals, thanks to the higher number of vanishing moments of the wavelet, as detailed later. In 1989, Mallat has devised a fast algorithm [39] to compute the orthogonal wavelet transform using wavelets defined by QMF. Later Malvar [41] and Coifman and Meyer [8] have found a new kind of windows of variable width that 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 [9] have proposed the so-called wavelet packets, which, similarly to compactly supported wavelets, are wavepackets of prescribed number of vanishing moments, 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, for example, the Navier–Stokes equations, which are the fundamental equations 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, for example shocks. The "art" of predicting such nonlinear evolutions (a generic case being turbulent flows) consists in 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 modeled. A review of the different types of wavelet transforms and their applications to analyze and compute turbulent flows is given in [19,52].

23.2.3.2 Continuous Wavelet Transform

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

$$C_\psi = \int_0^\infty |\widehat{\psi}(k)|^2 \, dk < \infty, \qquad (23.19)$$

where $\widehat{\psi} = \int_{-\infty}^\infty f(x)e^{-2\pi i k x} dx$ is the Fourier transform of ψ. From (23.18), we see that for ψ to be admissible it should satisfy in particular $\widehat{\psi}(0) = 0$, i.e., the space average of ψ should vanish and only then can the wavelet transform be 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 $\widehat{\psi}(k)$ 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

$$\int_{-\infty}^\infty x^m \psi(x) dx = 0 \quad \text{for } m = 0, \ldots, M-1, \qquad (23.20)$$

namely that its first M moments vanish. In this case the wavelet analysis will enhance any quasi-singular behavior 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

$$\psi_{l,x}(x') = c(l)\psi\left(\frac{x'-x}{l}\right) \qquad (23.21)$$

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

$$\tilde{f}(l,x) = \langle f, \psi_{l,x} \rangle = \int_{-\infty}^{\infty} f(x')\psi_{l,x}^*(x')dx', \qquad (23.22)$$

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 are normalized in L^2-norm ($c(l) = l^{-1/2}$), then the squared wavelet coefficients correspond to the energy density of the signal whose evolution can be tracked in both space and scale. If the wavelets are normalized in L^1-norm ($c(l) = l^{-1}$), the coefficients are related to the previous ones by the following relation:

$$\tilde{f}_{L^1} = l^{-1/2}\tilde{f}_{L^2}. \qquad (23.23)$$

Note that to study the Hölder regularity of a function and estimate its singularity spectrum, one typically uses wavelet coefficients in L^1-norm (see Section 23.2.1.3).

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

$$f(x') = C_\psi^{-1}\int_{\infty}^{\infty}\int_{0^+}^{\infty} \tilde{f}(l,x)\psi_{l,x}(x')\frac{dl}{l^2}dx \qquad (23.24)$$

with C_ψ the constant of the admissibility condition given in Equation 23.19, which only depends on the chosen wavelet ψ.

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

$$\widetilde{\beta_1 f_1(x) + \beta_2 f_2(x)} = \beta_1\widetilde{f_1(x)} + \beta_2\widetilde{f_2(x)} \qquad (23.25)$$

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 that are partially lost by the orthogonal wavelet transform. Let us also mention that, due to the localization of wavelets in physical space, the behavior of the signal at infinity does not play any role. In contrast, the nonlocal nature of the trigonometric functions used for the Fourier transform does not allow us to locally analyze or process a signal with it.

Figure 23.4 shows six examples of wavelet analyses of academic signals using the complex-valued Morlet wavelet: a Dirac spike (a), a step function (b), a superposition of two cosine functions having different frequencies (c), a succession of two cosine functions having different frequencies (d),

a chirp (e), and 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 Figure 23.4a, e, and f, have similar flat Fourier and wavelet spectra (see Section 23.3.3.2), although the space-scale representation of the energy density in wavelet space exhibit very different behaviors.

The extension of the continuous wavelet transform to analyze signals in d dimensions is made possible by replacing the affine group by the Euclidean group including rotations. 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

$$\psi_{l,\vec{x},r}(x') = \frac{1}{l^{d/2}}\psi\left(r^{-1}\left(\frac{\vec{x}'-\vec{x}}{l}\right)\right) \qquad (23.26)$$

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

$$C_\psi = \int_0^{\infty}\left|\hat{\psi}(k)\right|^2\frac{d^dk}{|k|^d} < \infty \qquad (23.27)$$

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

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \qquad (23.28)$$

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

$$\widetilde{f(l,\vec{x},\theta)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\vec{x}')\psi_{l,\vec{x},\theta}^*(x')d\vec{x}', \qquad (23.29)$$

and the wavelet synthesis is

$$f(\vec{x}') = \frac{1}{C_\psi}\int_0^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{2\pi} \tilde{f}(l,\vec{x},\theta)\psi_{l,\vec{x},\theta}^*(\vec{x}')\frac{dl\,d\vec{x}\,d\theta}{l^3}. \qquad (23.30)$$

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

23.2.3.3 Orthogonal Wavelet Transform

Wavelets can also be used to construct discrete representations of various function spaces, called frames [11], by selecting a discrete
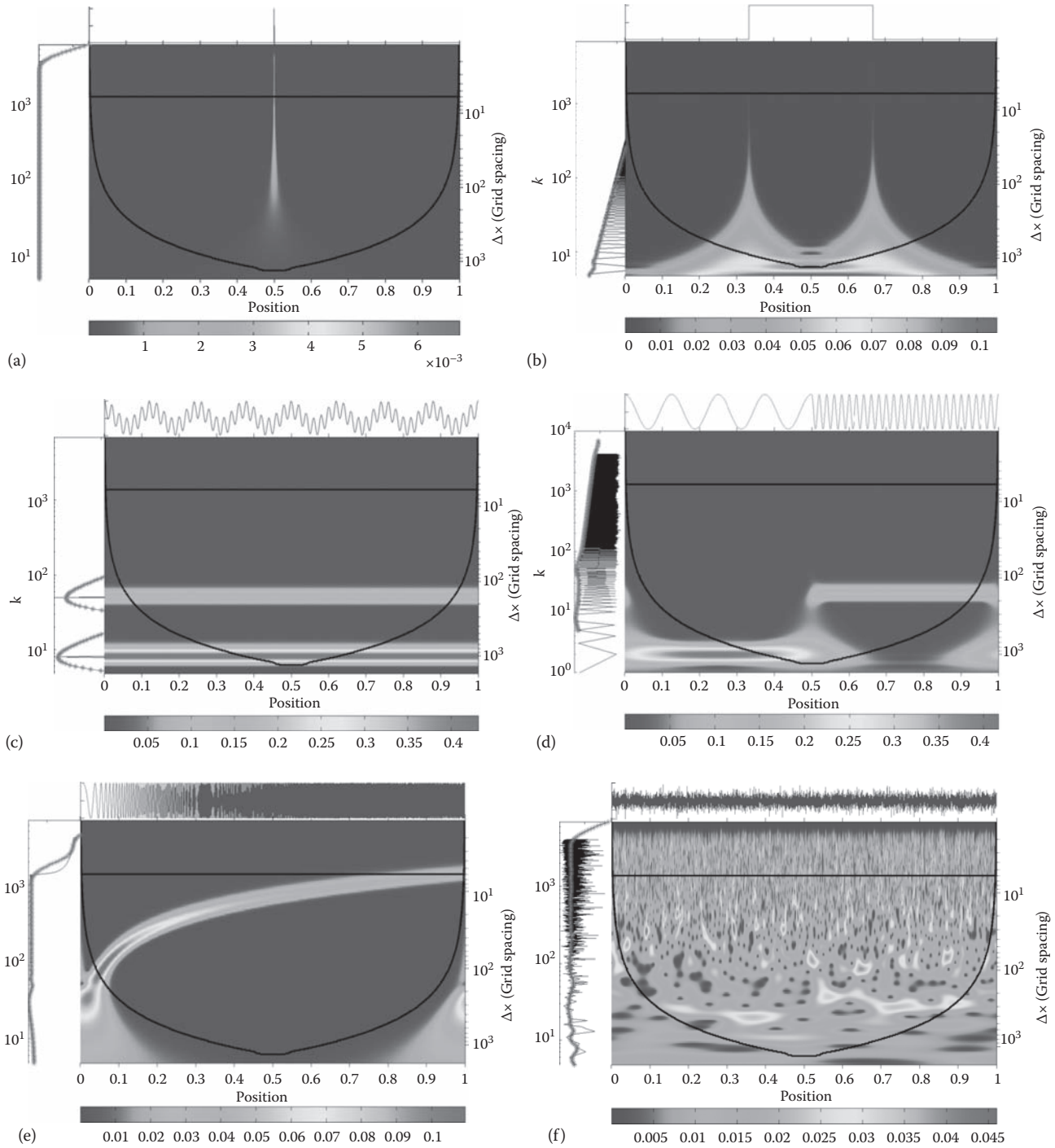
FIGURE 23.4 Examples of wavelet analyses of academic signals, namely a Dirac spike (a), a step function (b), the superposition of two cosine functions having different frequencies (c), the succession of two cosine functions of very different frequencies (d), a chirp (e), and finally one realization of a Gaussian white noise (f). The moduli 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 (grey crosses), as defined in Section 23.3.3.2, are also shown on the left, with the axes rotated by 90°.

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 discretized 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 visualized by plotting the modulus of the continuous wavelet coefficients of one realization of a white noise computed with a Morlet
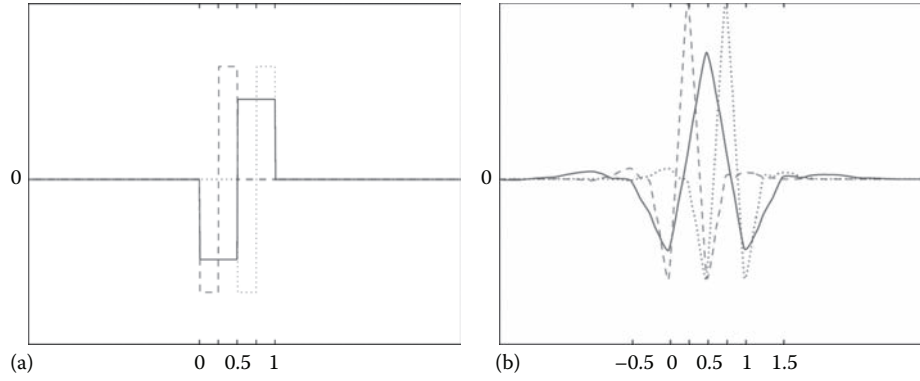
FIGURE 23.5 Orthogonal wavelets: Haar wavelet (a) and Coifman 12 wavelet (b). We have superposed one wavelet at scale $j = 0$ (solid line) and position $x = 0.5$) and two wavelets at the next smaller scale $j = 1$, located at position $x = 0.25$ (dashed line) and $x = 0.75$ (dotted line). They are mutually orthogonal, which can be directly seen for the Haar wavelet and which is much less obvious for the Coifman wavelet.

wavelet, see Figure 23.9b. 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 artifacts one encounters when denoising with the orthogonal wavelet transform, which does not preserve those invariances (Figure 23.5).

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 consist 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 earlier and we assume that it satisfies all the necessary conditions (see, e.g., [13]) so that the wavelets $\psi_{l,i}$ defined by Equation 23.21 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 expand 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 MRA as follows [13,19]:

For λ_0 belonging to the index set

$$\Lambda_0 = \{(j, \vec{\mu}, \vec{\tau}) \hat{u} j = 0, \dots, J-1, \quad \vec{\mu} \in \{0,1\}^3, \quad \vec{\tau} \in \{0, \dots, 2^j - 1\}^3\},$$

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 $\vec{\tau}$ are the 3D

equivalent to the scale and position parameters that we are already familiar with from the preceding discussion of the one-dimensional 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 nonvanishing 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 $\vec{\mu} \ne 0$, which we denote Λ. The wavelet coefficients and scaling coefficients of f_d are then simply defined by

$$\tilde{f}_\lambda^d = \langle f_d, \psi_\lambda \rangle$$
$$\bar{f}_\lambda^d = \langle f_d, \phi_\lambda \rangle,$$

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:

$$f_d = \bar{f}_{(0,0,0)} + \sum_{\lambda \in \Lambda} \tilde{f}_\lambda^d \psi_\lambda. \qquad (23.31)$$

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, $j = 0, \dots, J - 1$, while preserving some amount of space-locality thanks to the position index $\vec{\tau}$ and also some amount of directionality thanks to $\vec{\mu}$. Hence the expansion coefficients appearing in Equation 23.31 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, whose complexity is $O(N \log_2 N)$.

23.3 Methods of Analysis

23.3.1 Fractals

23.3.1.1 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 object is counted again. The earlier 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 Equation 23.1.

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, e.g., singular sets, the limit obtained with the box-counting algorithm corresponds to the Hausdorff dimension (box counting dimension $\geq$ Hausdorff dimension) and thus this technique is an efficient way for computing it.

23.3.1.2 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 functions $\{f_i\}_{i\in[1,N]}$ from $\mathbb{R}^d$ into itself which are contractions, i.e., 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 $C(\mathbb{R}^d)$ to itself, where $C(\mathbb{R}^d)$ denotes the set of all compact subsets of $\mathbb{R}^d$, defined by

$$F(A) = f_1(A) \cup \cdots \cup f_N(A), \qquad (23.32)$$

with $A \in C(\mathbb{R}^d)$. It can be shown that F itself is also a contraction defined into $C(\mathbb{R}^d)$ for the Hausdorff distance δ_H, that is $\delta_H(F(A),F(B)) \leq c\delta_H(A,B)$, 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 $(C(\mathbb{R}^d), \delta_H)$, F admits a fixed point in $C(\mathbb{R}^d)$, 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 Figure 23.1. 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 $f_i(x) = c_i R_i x + b_i$ 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

$$\sum_i c_i^d = 1. \qquad (23.33)$$

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 earlier. Similarly, 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. A 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 $f_i(x) = A_i x + b_i$, where A_i is a matrix, then p_i can be defined as $p_i = \dfrac{|\det A_i|}{\sum_k |\det A_k|}$. 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 nonzero 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 [6]. We consider a compact ensemble 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 practical 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 S invariant up to a tolerance ε, i.e., $\delta_H(S, F(S)) \leq \varepsilon$, then the limit ensemble A_F associated to the IFS satisfies

$$\delta_H(S,A_F) \leq \frac{\varepsilon}{1-c}, \qquad (23.34)$$

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.

23.3.1.3 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 Figure 23.6b, we show the singularity spectrum of the function f plotted in Figure 23.6a. 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 Figure 23.6a. If a white noise with a weak standard deviation of $\sigma = 0.01$ is added, see Figure 23.6c, then the signal becomes more irregular leading to a singularity spectrum truncated at a Hölder exponent close to $\alpha = 0.5$, as seen on Figure 23.6d. Moreover, the support of the singularities becomes larger since the fractal dimension $d(\alpha)$ for $\alpha = 0.5$ for the noise-free signal, in Figure 23.6a, is close to $\alpha = 0.5$,

see Figure 23.6b, while for the noisy signal in Figure 23.6c it is close to $\alpha = 1$, see Figure 23.6d. This effect is reinforced with a more intense noise of standard deviation $\sigma = 0.1$, see Figure 23.6e 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 Figure 23.6.

23.3.2 Self-Similar Random Processes

23.3.2.1 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
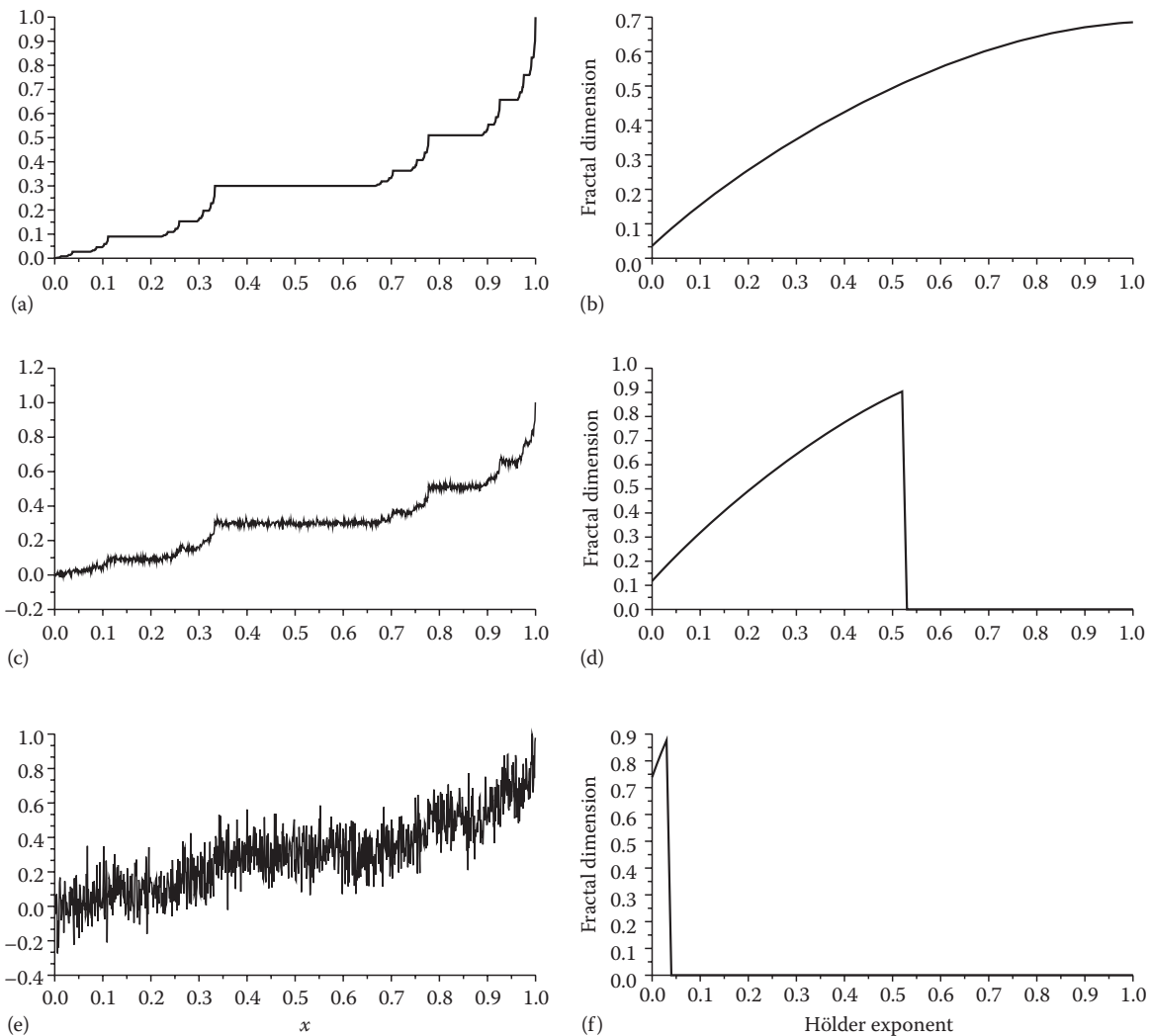

FIGURE 23.6 Singularity spectrum of a function f (a) (the Devil's staircase) and its noisy versions perturbed with a white noise of standard deviation $\sigma = 0.01$ (c) and $\sigma = 0.1$ (e). The corresponding singularity spectra are shown on the right column. Without noise (b), with noise of standard deviation $\sigma = 0.01$ (d) and $\sigma = 0.1$ (f).

the step size $\delta t = 1/N$, N being the number of sampling points, is given by

$$V_N = \sum_{k=0}^{N-1}\left[B_H\left(\frac{k}{N} + \frac{1}{N}\right) - B_H\left(\frac{k}{N}\right)\right]^2. \qquad (23.35)$$

This quadratic variation can be related to the Hurst exponent by

$$V_N = c\, N^{1-2H}, \qquad (23.36)$$

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

$$\frac{V_N}{V_{N/2}} = 2^{1-2H}, \qquad (23.37)$$

which leads thus to the Hurst exponent

$$H = \frac{1}{2}\left(1 - \log_2 \frac{V_N}{V_{N/2}}\right). \qquad (23.38)$$

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.

23.3.2.2 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 that allow the efficient generation of realizations with long-range dependence and with many scales without imposing a cutoff scale thanks to the vanishing moment property of the wavelets.

For synthesizing 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 $i,j = 1, \ldots, N$, 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 $i = 1, \ldots, N$. 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. Finally, let us recall that the Hausdorff dimension of the graph of B_H is $2 - H$.

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 and Majda) [17]. A related construction was proposed by Tafti and Unser [53].

An interesting technique from image processing, which was originally developed for generating artificial clouds in computer animations was proposed in [10]. Therewith intermittent scalar valued processes in two space dimensions can be efficiently generated, which have a given energy distribution, that could be self-similar. The resulting process is strictly band-limited.

23.3.2.3 Application to Fractional Brownian Motion

To illustrate the fractional Brownian motion, we show in Figure 23.7b three realizations of different fractional Brownian
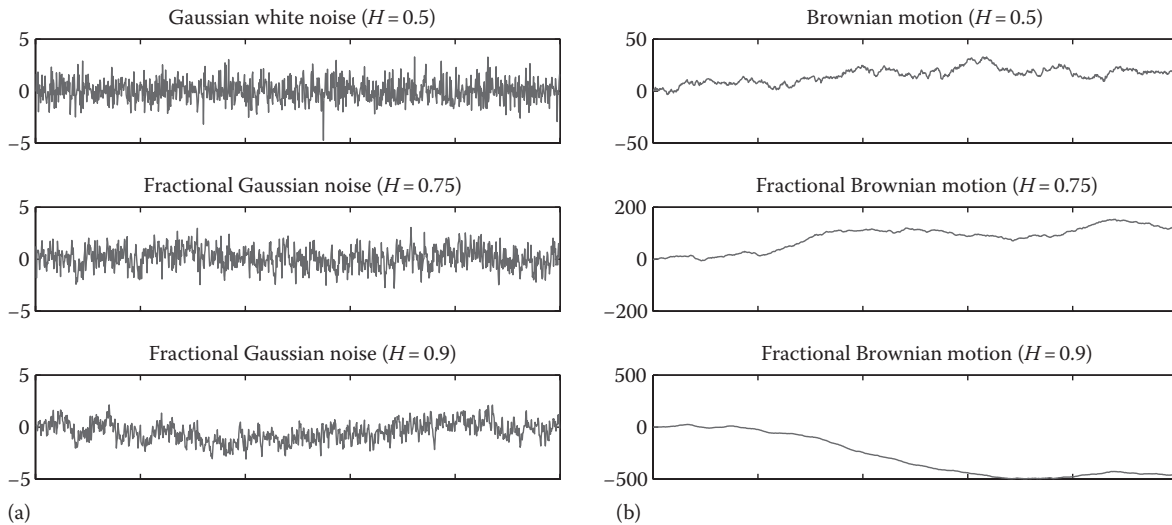

FIGURE 23.7 Sample trajectories of Gaussian fractional noise (a), and of fractional Brownian motion (b) for three different values of the Hurst exponent H. The Gaussian fractional noise (a) corresponds to increments of the fractional Brownian motion (b). The resolution is $N = 1024$.
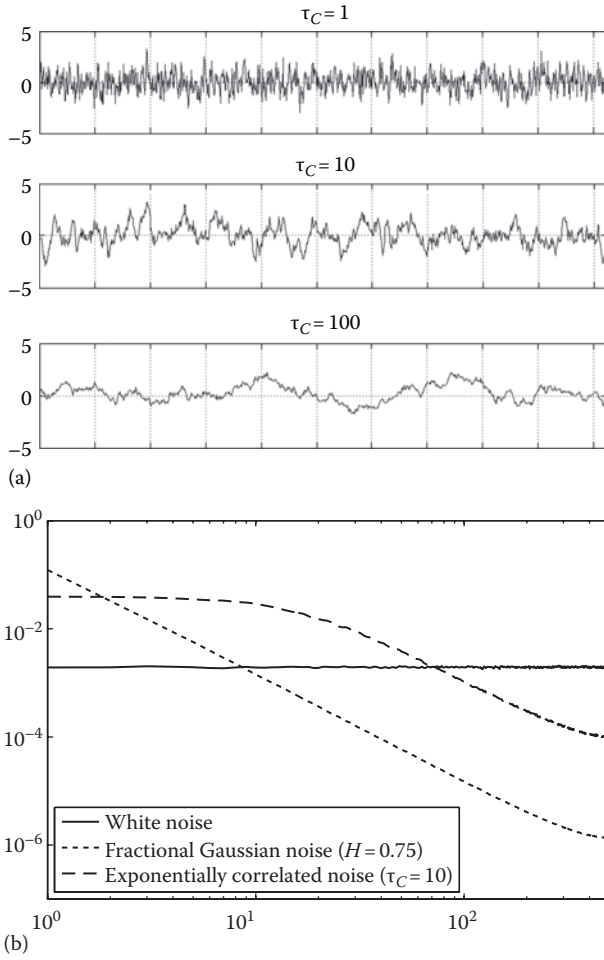
(a)

(b)

FIGURE 23.8 (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.

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 Figure 23.7a. We can observe that the regularity of the curves increases for larger values of H.

To model random process with short-range correlation we can suppose that the covariance function decays exponentially $\propto \exp(-t/\tau_c)$, with τ_c being the correlation time. The corresponding spectral density decays $\propto 4\tau_c/(1 + (f\tau_c)^2)$. Figure 23.8 shows examples for different values of τ_C (a) and different spectral densities (b). For increasing $\propto$ the apparent regularity of the trajectory increases, although the actual regularity of the underlying function remains the same.

23.3.3 Wavelets

23.3.3.1 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 behavior of the signal and detect for instance frequency modulation laws or quasi-singularities, even if they are superimposed. 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,

$$\psi(x) = (1 - x^2)e^{\frac{-x^2}{2}} \tag{23.39}$$

and its Fourier transform is

$$\hat{\psi}(k) = k^2 e^{\frac{-k^2}{2}} \tag{23.40}$$
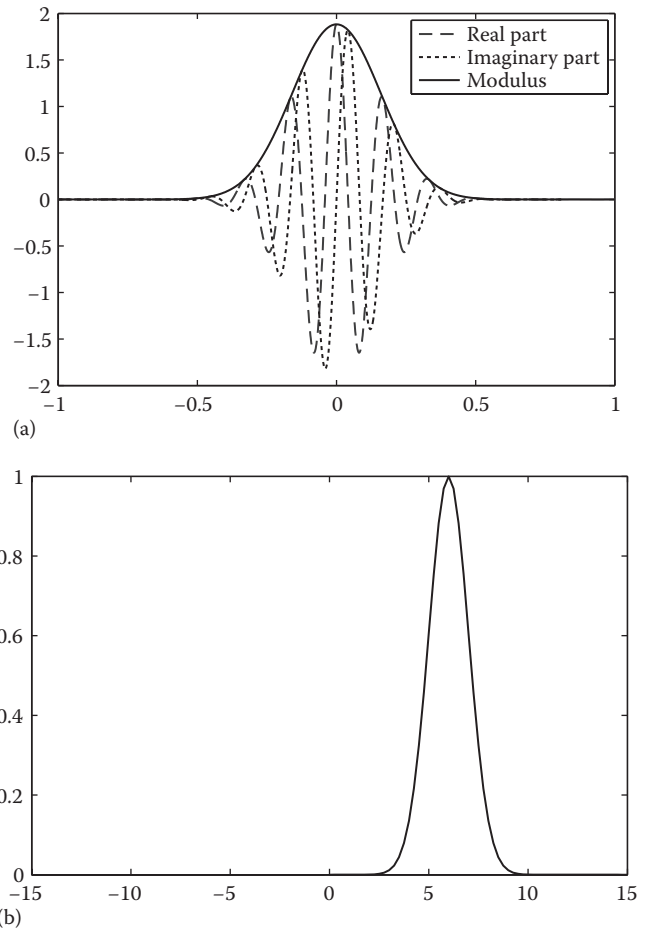

(a)

(b)

FIGURE 23.9 Localization of the Morlet wavelet in physical space (a) and in spectral space (b).

The most useful complex-valued wavelet is the Morlet wavelet (Figure 23.9):

$$\psi(x) = e^{ik_\psi x} e^{-\frac{x^2}{2}} \qquad (23.41)$$

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

$$k_\psi = \frac{\int_0^\infty k \,|\,\hat{\psi}(k)\,|\,dk}{\int_0^\infty |\,\hat{\psi}(k)\,|\,dk}. \qquad (23.42)$$

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 Equation 23.19 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., $\hat{\psi}(0)$, to zero, which gives

$$\hat{\psi}(k) = \begin{cases} e^{-\frac{(k-k_\psi)^2}{2}} & \text{for } k > 0, \\ 0 & \text{for } k \le 0. \end{cases}$$

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, and computational efficiency.

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 FWT in $24N$ multiplications (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 one 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 FWT [13,19,40]. 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 modeling 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 (Figure 23.10), and two self-similar random signals, which are fractional Brownian motions (FBM) having different Hurst exponent, i.e., $H = 0.25$ and $H = 0.75$ (Figure 23.11).
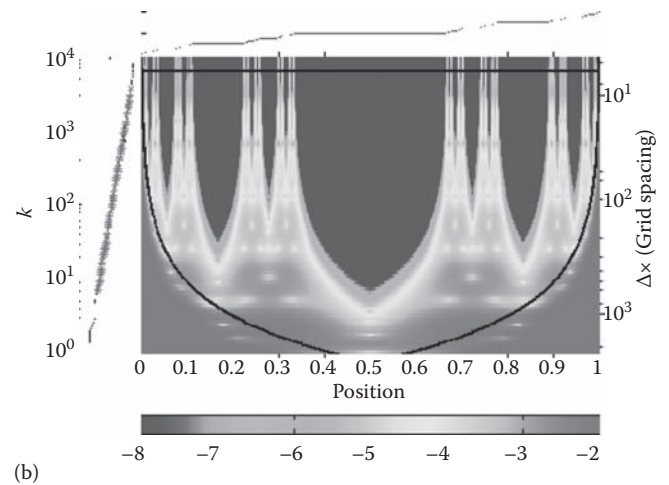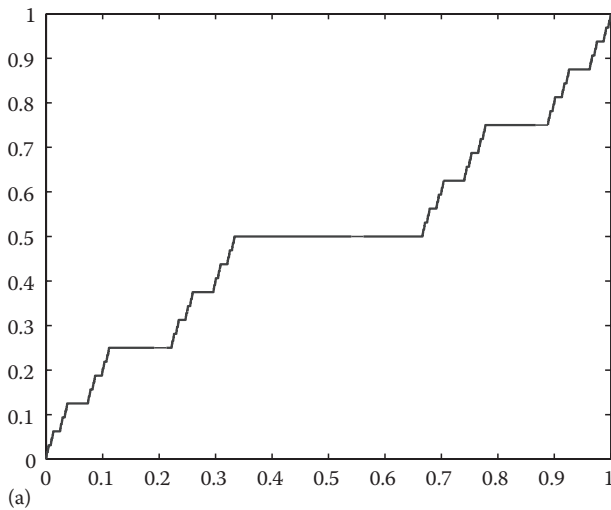

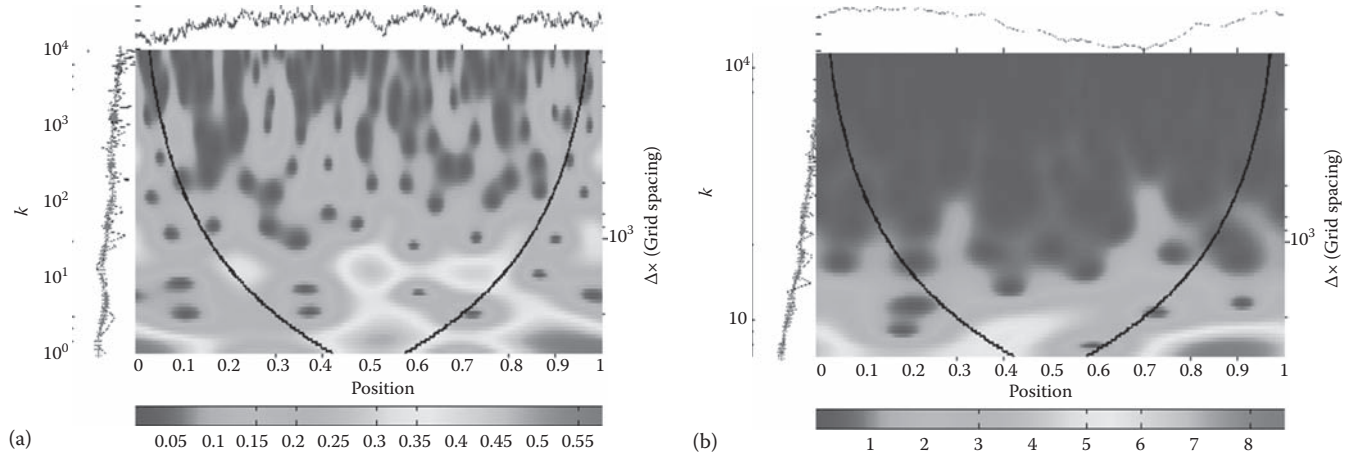**FIGURE 23.10** Devil's staircase (a) and its continuous wavelet analysis (b).

(a)

(b)

FIGURE 23.11 Continuous wavelet analysis of fractional Brownian motion with Hurst exponent $H = 0.25$ (a) and $H = 0.75$ (b).

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

$$\tilde{E}(k,x) = \frac{1}{C_\psi k_\psi} \left| \tilde{f}\left(\frac{k_\psi}{k}, x\right) \right|^2 \text{ for } k \geq 0, \qquad (23.43)$$

where

k_ψ is the centroid wavenumber of the analyzing wavelet ψ
C_ψ is defined by the admissibility condition given in Equation 23.19

By measuring $\tilde{E}(k, x)$ 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 23.3.3.5) 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 Equation 23.43 over all positions,

$$\tilde{E}(k) = \int_{-\infty}^{\infty} \tilde{E}(k,x)dx \qquad (23.44)$$

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.

23.3.3.3 Relation to Fourier Spectrum

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

$$\tilde{E}(k) = \frac{1}{C_\psi k_\psi} \int_0^\infty E(k') \left| \hat{\psi}\left(\frac{k_\psi k'}{k}\right) \right|^2 dk', \qquad (23.45)$$

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., $\frac{\Box k}{k} =$ constant. The wavelet energy spectrum thus yields a stabilized Fourier energy spectrum.

Considering, for example, the Marr wavelet given in Equation 23.39, 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 Equation 23.43, only the zeroth-order moment is vanishing. However higher mth-order moments are very small $\left(\propto k_\psi^m e^{(-k_\psi^2/2)}\right)$, 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

$$\hat{\psi}_n(k) = \alpha_n \exp\left(-\frac{1}{2}\left(k^2 + \frac{1}{k^{2n}}\right)\right), \quad n \geq 1, \qquad (23.46)$$

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.

23.3.3.4 Relation to Structure Functions

Structure functions, classically used to analyze nonstationary random processes, e.g., turbulent velocity fluctuations, have some limitations that 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 earlier 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 behavior 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

$$\psi^\delta(x) = \delta(x+1) - \delta(x). \tag{23.47}$$

We thus obtain

$$f(x+a) - f(x) = \tilde{f}_{x,a} = \langle f, \psi^\delta_{x,a} \rangle, \tag{23.48}$$

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

$$S_p(a) = \int \left(\tilde{f}_{x,a} \right)^p dx. \tag{23.49}$$

As already mentioned earlier the drawback of the DoD wavelet is that it has only one vanishing moment, its mean. Consequently the exponent of the pth-order structure function in the case of a power-law behavior 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 23.45 yields a relation between the global wavelet spectrum $\tilde{E}(k)$ and the Fourier spectrum $E(k)$ for a given wavelet ψ. Taking the Fourier transform of the DoD wavelet, we get $\hat{\psi}^\delta(k) = e^{\iota k} - 1 = e^{\frac{\iota k}{2}}(e^{\frac{\iota k}{2}} - e^{-\iota k2})$ and therefore we have $|\hat{\psi}^\delta(k)|^2 = 2(1 - \cos k)$. The relation between the Fourier and the wavelet spectrum thus becomes

$$\tilde{E}(k) = \frac{1}{C_\psi k} \int_0^\infty E(k') \left(2 - 2\cos\left(\frac{k_\psi k'}{k} \right) \right) dk', \tag{23.50}$$

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

$$\tilde{E}(k) = \frac{1}{C_\psi k} S_2(a). \tag{23.51}$$

Using now the result of Section 23.3.3.2 that for a Fourier spectrum that behaves like $k^{-\alpha}$ for $k \to \infty$, the wavelet spectrum only yields $\tilde{E}(k) \propto k^{-\alpha}$ 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.

23.3.3.5 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 [27,31]. 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 Equation 23.21.

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

$$\left[\tilde{f}(l, x_0) \right]_{l \to 0} \leq l^{m+1/2} l^{1/2} = l^m, \tag{23.52}$$

The factor $l^{1/2}$ comes from the fact that to study the scaling in x_0 of the function f we 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 Equation 23.21. More generally if f has Hölder regularity α at x_0 (see Section 23.2.1.3), then

$$\left[\tilde{f}(l, x_0) \right]_{l \to 0} \approx C e^{i\Phi} l^{1/2} \tag{23.53}$$

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 toward 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 Equation 23.52, 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 toward 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 Figure 23.10. We can then compute the singularity spectrum (see Section 23.2.1.3).

23.3.3.6 Intermittency Measures

Localized bursts of high-frequency activity define typically intermittent behavior. 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 artifacts 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 [19,50], which corresponds to the wavelet coefficients renormalized by the space-averaged energy at each scale, such that

$$I(l, \vec{x}) = \frac{|\tilde{f}(l, \vec{x})|^2}{\int_{-\infty}^{\infty} \tilde{f}(l, \vec{x})|^2 \, d^2\vec{x}}.$$ (23.54)

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 nonintermittent, while regions of larger values are intermittent.

Similar to the continuous wavelet transform the orthogonal wavelet transform allows to define intermittency measures, either local as shown earlier or global as illustrated in the following text. The space-scale information contained in the wavelet coefficients yields suitable global intermittency measures using scale-dependent moments and moment ratios [51]. For a signal f the moments of wavelet coefficients at different scales j are defined by

$$M_{p,j}(f) = 2^{-j} \sum_{i=0}^{2^j-1} \left(\tilde{f}_{j,i}\right)^p.$$ (23.55)

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 recovered 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:

$$Q_{p,q,j}(f) = \frac{M_{p,j}(f)}{(M_{q,j}(f))^{\frac{p}{q}}},$$ (23.56)

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 reflected in increasing $Q_{p,q,j}$ for increasing j (smaller scale) supposing $p > q$.

23.3.3.7 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 characterize 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 earlier 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 what remains after denoising, we need a model, not for 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 Equation 23.31 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 $\varepsilon = (2/d\sigma^2 \ln N)^{1/2}$. 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 (Figure 23.12a), 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 11.04 dB (Figure 23.12b). Applying the extraction method we recover a denoised version of the corrupted signal, which preserves the quasi-singularities (Figure 23.12c). It could be checked a posteriori that the incoherent contribution is spread, and therefore does not compress and has a Gaussian probability distribution.
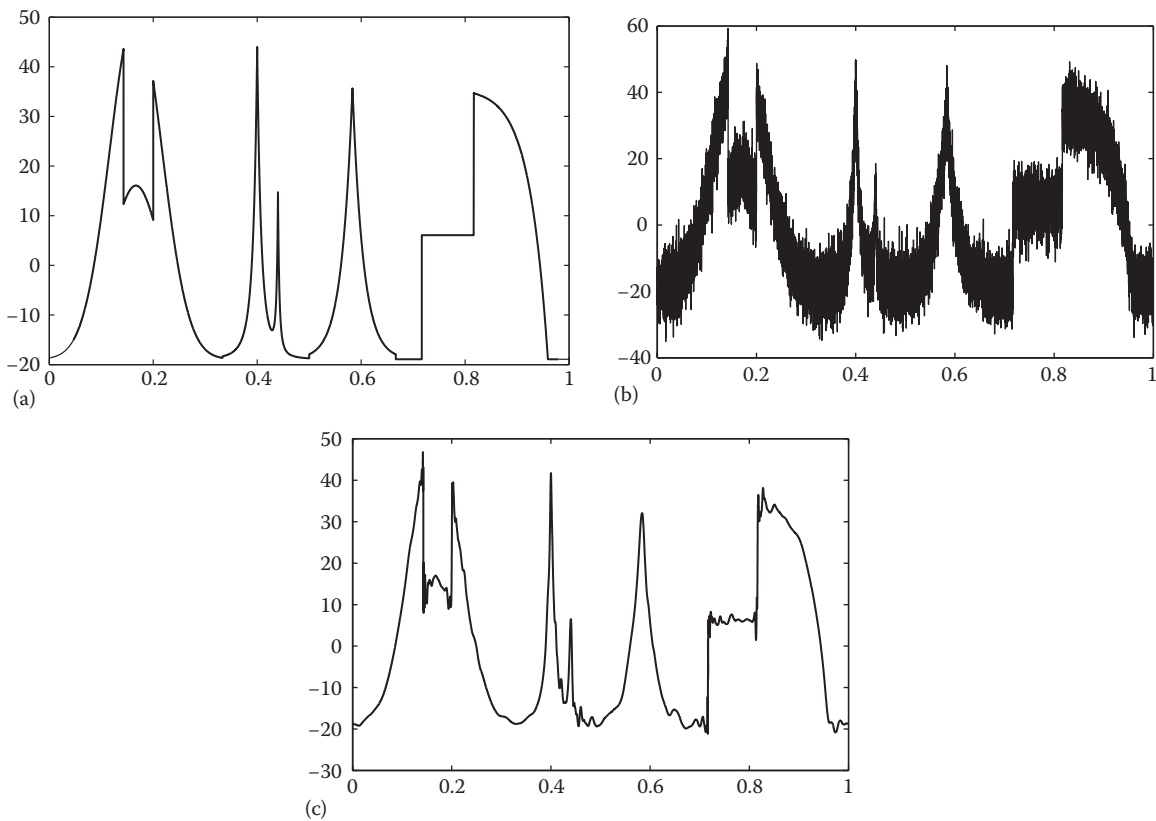

FIGURE 23.12 Academic example of denoising of a piecewise regular signal using the algorithm for coherent structure extraction. Original signal (a), same signal plus Gaussian white noise giving a signal to noise ratio (SNR) of 11.04 dB (b), denoised signal with SNR of 27.55 dB (c).

23.4 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, 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 therefore are 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 use of fractals and wavelets, as they were applied to study turbulence. Kolmogorov's statistical theory of homogeneous and isotropic turbulence [34] assumes that there exists an energy cascade from large to small scales, which is modeled 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 realizations, 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 [49].

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 that contains it. But keeping such a fractal picture to describe three-dimensional whirls that 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 [35].

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 multifractals (for a review of them see [22]).

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
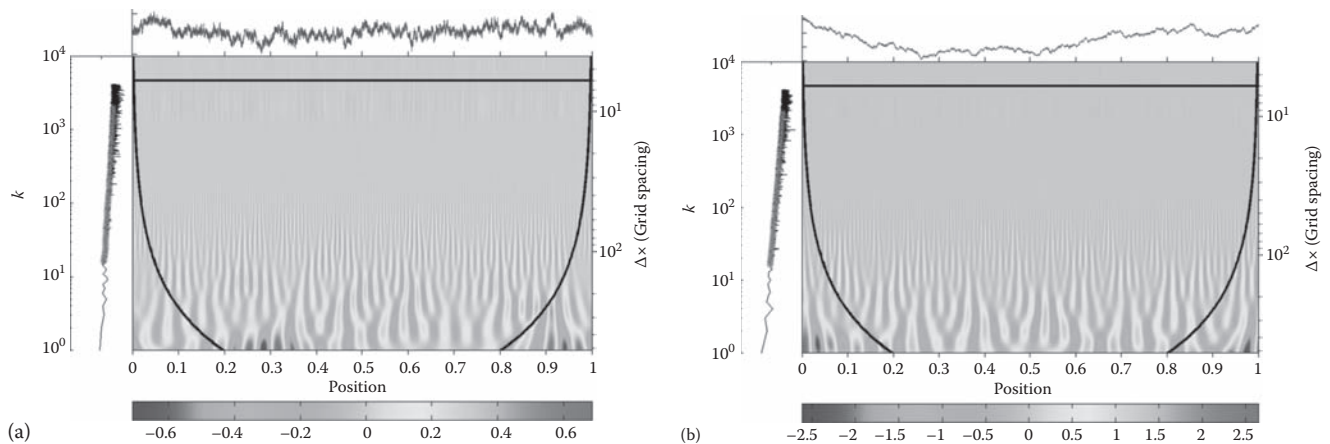

(a)

(b)

FIGURE 23.13 (See color insert.) Real part of the continuous wavelet analysis of fractional Brownian motion with Hurst exponent $H = 0.25$ (a) and of classical Brownian motion (b).

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 Figure 23.13), which proves that the correlation is among the wavelets but not within the signal itself. Unfortunately, in the case of turbulent signals this pattern has been misinterpreted as the evidence of whirls breaking in a paper published by *Nature* in 1989 under the title "Wavelet analysis reveals the multifractal nature of the Richardson's cascade" [2]. If the authors had used an orthogonal wavelet transform instead of a real-valued continuous wavelet transform they would not have observed and correlation.

Later 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"[43]. 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 1980s, 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?" [46]

Let the future tell us the answer...

Acknowledgments

References

1. J. P. Antoine, R. Murenzi, P. Vandergheynst, and S. T. Ali, 2004. *Two-Dimensional Wavelets and Their Relatives*. Cambridge University Press, Cambridge, U.K.

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. L. Bachelier, 1900. *Théorie de la Spéculation*. Gauthier-Villars, Paris, France.

5. R. Balian, 1981. Un principe d'incertitude en théorie du signal ou en mécanique quantique. *C. R. Acad. Sci. Paris*, **292**, série 2, 1357–1361.

6. M. Barnsley, 1988. *Fractals Everywhere*. Academic Press, Boston, MA.

7. B. Burke Hubbard, 1998. *The World According to Wavelets*. A. K. Peters, Wellesley, MA.

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

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

10. R. L. Cook and T. DeRose, 2005. Wavelet noise. *ACM Trans. Graph.*, **24**(3), 803–811.

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. *Commun. Pure Appl. Math.*, **41**(7), 909–996

13. I. Daubechies, 1992. *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA.

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. Sci. 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), 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), 146–162.

18. M. Farge and 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. M. Farge, A. Grossmann, Y. Meyer, T. Paul, J.-C. Risset, G. Saracco, and B. Torrésani, 2012. Les ondelettes et le CIRM. *Gazette Math.*, **131**, 47–57.

22. U. Frisch, 1995. *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press, Cambridge, U.K.

23. D. Gabor, 1946. Theory of Communication. *J. Inst. Electr. Eng.*, **93** (3), 429–457.

24. J.-F. Gouyet, 1996. *Physics and Fractal Structures*. Springer-Verlag, Berlin, Germany.

25. 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.

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

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

28. Interview de John Hubbard, *Science Publique, France-Culture*, 19 October 2010.

29. G. A. Hunt, 1951. Random Fourier transforms. *Trans. Am. Math. Soc.*, **71**, 38–69.

30. H. E. Hurst, 1951. Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civil Eng.*, **116**, 770–808.

31. 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, pp. 247–252.

32. 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.

33. A. N. Kolmogorov, 1940. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C. R. Dokl. Acad. Sci. SSSR*, **26**, 115–118.

34. 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. Roy. Soc. Ser. A Math. Phys. Sci.*, **434**, 9–13 (1991).

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

36. P. G. Lemarié and Y. Meyer, 1986. Ondelettes et bases Hilbertiennes. *Rev. Mat. Ibero Am.*, **2**, 1–18.

37. P. Lévy, 1954. *Le mouvement Brownien.* Fascicule CXXVI of Mémorial des Sciences Mathématiques. Gauthier-Villars, Paris, France.

38. P. Lévy, 1965. *Processus stochastique et mouvement Brownien.* Fascicule CXXVI of Mémorial des sciences mathématiques. Gauthier-Villars, Paris, France.

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

40. S. Mallat, 1998. *A Wavelet Tour of Signal Processing.* Academic Press.

41. H. Malvar, 1990. Lapped transforms for multiresolution signal decomposition: the wavelet decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.*, **11**, 674–693.

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

43. Interview de Benoît Mandelbrot. *A voix nue, France-Culture*, October 1990.

44. B. B. Mandelbrot, 1975. *Les Objets Fractals*. Flammarion, Paris, France.

45. B. B. Mandelbrot, 1977. *Fractals: Form, Chance and Dimension*. Freeman, San Fransisco, CA.

46. Y. Meyer, 1990. Ondelettes et applications. *J. Ann. Soc. Math. Société Mathématique de France (SMF)*, *Paris*, 1–15.

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

48. G. Parisi and U. Frisch, 1985. On the singularity structure of fully developed turbulence. *Turbulence and Predictability in Geophysical Fluid Dynamics* (eds. M. Ghil, R. Benzi and G. Parisi), pp. 84–87, North-Holland, Amsterdam, the Netherlands.

49. L. F. Richardson, 1922. *Weather Prediction by Numerical Process*. Cambridge University Press, Cambridge, U.K.

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

51. 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, Hackensack, NJ, pp. 302–328.

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

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