CHAPTER

Wavelet Zoom

6

A wavelet transform can focus on localized signal structures with a zooming procedure that progressively reduces the scale parameter. Singularities and irregular structures often carry essential information in a signal. For example, discontinuities in images may correspond to occlusion contours of objects in a scene. The wavelet transform amplitude across scales is related to the local signal regularity and Lipschitz exponents. Singularities and edges are detected from wavelet transform local maxima at multiple scales. These maxima define a geometric scale–space support from which signal and image approximations are recovered.

Nonisolated singularities appear in highly irregular signals such as multifractals. The wavelet transform takes advantage of multifractal self-similarities to compute the distribution of their singularities. This singularity spectrum characterizes multifractal properties. Throughout this chapter wavelets are real functions.

6.1 LIPSCHITZ REGULARITY

To characterize singular structures, it is necessary to precisely quantify the local regularity of a signal $f(t)$. Lipschitz exponents provide uniform regularity measurements over time intervals, but also at any point v. If f has a singularity at v, which means that it is not differentiable at v, then the Lipschitz exponent at v characterizes this singular behavior.

Section 6.1.1 relates the uniform Lipschitz regularity of f over $\mathbb{R}$ to the asymptotic decay of the amplitude of its Fourier transform. This global regularity measurement is useless in analyzing the signal properties at particular locations. Section 6.1.3 studies zooming procedures that measure local Lipschitz exponents from the decay of the wavelet transform amplitude at fine scales.

6.1.1 Lipschitz Definition and Fourier Analysis

The Taylor formula relates the differentiability of a signal to local polynomial approximations. Suppose that f is m times differentiable in $[v-h, v+h]$. Let p_v be

the Taylor polynomial in the neighborhood of v:

$$p_v(t) = \sum_{k=0}^{m-1} \frac{f^{(k)}(v)}{k!} (t-v)^k. \tag{6.1}$$

The Taylor formula proves that the approximation error

$$\varepsilon_v(t) = f(t) - p_v(t)$$

satisfies

$$\forall t \in [v-h, v+h], \quad |\varepsilon_v(t)| \leq \frac{|t-v|^m}{m!} \sup_{u \in [v-h, v+h]} |f^m(u)|. \tag{6.2}$$

The mth-order differentiability of f in the neighborhood of v yields an upper bound on the error $\varepsilon_v(t)$ when t tends to v. The Lipschitz regularity refines this upper bound with noninteger exponents. Lipschitz exponents are also called *Hölder* exponents in mathematics literature.

Definition 6.1: *Lipschitz.*

- A function f is pointwise Lipschitz $\alpha \geq 0$ at v, if there exists $K > 0$ and a polynomial p_v of degree $m = \lfloor \alpha \rfloor$ such that

$$\forall t \in \mathbb{R}, \quad |f(t) - p_v(t)| \leq K |t-v|^\alpha. \tag{6.3}$$

- A function f is uniformly Lipschitz α over $[a, b]$ if it satisfies (6.3) for all $v \in [a, b]$ with a constant K that is independent of v.
- The Lipschitz regularity of f at v or over $[a, b]$ is the supremum of the α such that f is Lipschitz α.

At each v the polynomial $p_v(t)$ is uniquely defined. If f is $m = \lfloor \alpha \rfloor$ times continuously differentiable in a neighborhood of v, then p_v is the Taylor expansion of f at v. Pointwise Lipschitz exponents may vary arbitrarily from abscissa to abscissa. One can construct multifractal functions with nonisolated singularities, where f has a different Lipschitz regularity at each point. In contrast, uniform Lipschitz exponents provide a more global measurement of regularity, which applies to a whole interval. If f is uniformly Lipschitz $\alpha > m$ in the neighborhood of v, then one can verify that f is necessarily m times continuously differentiable in this neighborhood.

If $0 \leq \alpha < 1$, then $p_v(t) = f(v)$ and the Lipschitz condition (6.3) becomes

$$\forall t \in \mathbb{R}, \quad |f(t) - f(v)| \leq K |t-v|^\alpha.$$

A function that is bounded but discontinuous at v is Lipschitz 0 at v. If the Lipschitz regularity is $\alpha < 1$ at v, then f is not differentiable at v and α characterizes the singularity type.

Fourier Condition

The uniform Lipschitz regularity of f over $\mathbb{R}$ is related to the asymptotic decay of its Fourier transform. Theorem 6.1 can be interpreted as a generalization of Theorem 2.5.

Theorem 6.1. A function f is bounded and uniformly Lipschitz α over $\mathbb{R}$ if

$$\int_{-\infty}^{+\infty} |\hat{f}(\omega)|\,(1+|\omega|^{\alpha})\,d\omega < +\infty. \tag{6.4}$$

Proof. To prove that f is bounded, we use the inverse Fourier integral (2.8) and (6.4), which shows that

$$|f(t)| \leqslant \int_{-\infty}^{+\infty} |\hat{f}(\omega)|\,d\omega < +\infty.$$

Let us now verify the Lipschitz condition (6.3) when $0 \leqslant \alpha \leqslant 1$. In this case, $p_v(t) = f(v)$ and the uniform Lipschitz regularity means that there exists $K > 0$ such that for all $(t, v) \in \mathbb{R}^2$

$$\frac{|f(t)-f(v)|}{|t-v|^{\alpha}} \leqslant K.$$

Since

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}(\omega)\,\exp(i\omega t)\,d\omega, \\ \frac{|f(t)-f(v)|}{|t-v|^{\alpha}} &\leqslant \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\hat{f}(\omega)|\,\frac{|\exp(i\omega t)-\exp(i\omega v)|}{|t-v|^{\alpha}}\,d\omega. \end{aligned} \tag{6.5}$$

For $|t-v|^{-1} \leqslant |\omega|$,

$$\frac{|\exp(i\omega t)-\exp(i\omega v)|}{|t-v|^{\alpha}} \leqslant \frac{2}{|t-v|^{\alpha}} \leqslant 2\,|\omega|^{\alpha}.$$

For $|t-v|^{-1} \geqslant |\omega|$,

$$\frac{|\exp(i\omega t)-\exp(i\omega v)|}{|t-v|^{\alpha}} \leqslant \frac{|\omega|\,|t-v|}{|t-v|^{\alpha}} \leqslant |\omega|^{\alpha}.$$

Cutting the integral (6.5) in two for $|\omega| < |t-v|^{-1}$ and $|\omega| \geqslant |t-v|^{-1}$ yields

$$\frac{|f(t)-f(v)|}{|t-v|^{\alpha}} \leqslant \frac{1}{2\pi}\int_{-\infty}^{+\infty} 2\,|\hat{f}(\omega)|\,|\omega|^{\alpha}\,d\omega = K.$$

If (6.4) is satisfied, then $K < +\infty$ so f is uniformly Lipschitz α.

Let us extend this result for $m = \lfloor \alpha \rfloor > 0$. We proved in (2.42) that (6.4) implies that f is m times continuously differentiable. One can verify that f is uniformly Lipschitz α over $\mathbb{R}$ if and only if $f^{(m)}$ is uniformly Lipschitz $\alpha - m$ over $\mathbb{R}$. The Fourier transform of $f^{(m)}$ is $(i\omega)^m \hat{f}(\omega)$. Since $0 \leqslant \alpha - m < 1$, we can use our previous result, which proves that $f^{(m)}$ is uniformly Lipschitz $\alpha - m$, and thus that f is uniformly Lipschitz α. ■

The Fourier transform is a powerful tool for measuring the minimum global regularity of functions. However, it is not possible to analyze the regularity of f at a particular point v from the decay of $|\hat{f}(\omega)|$ at high frequencies ω. In contrast, since wavelets are well localized in time, the wavelet transform gives Lipschitz regularity over intervals *and* at points.

6.1.2 Wavelet Vanishing Moments

To measure the local regularity of a signal, it is not so important to use a wavelet with a narrow frequency support, but vanishing moments are crucial. If the wavelet has n vanishing moments, then we show that the wavelet transform can be interpreted as a multiscale differential operator of order n. This yields a first relation between the differentiability of f and its wavelet transform decay at fine scales.

Polynomial Suppression

The Lipschitz property (6.3) approximates f with a polynomial p_v in the neighborhood of v:

$$f(t) = p_v(t) + \varepsilon_v(t) \quad \text{with} \quad |\varepsilon_v(t)| \leq K\,|t-v|^{\alpha}. \tag{6.6}$$

A wavelet transform estimates the exponent α by ignoring the polynomial p_v. For this purpose, we use a wavelet that has $n > \alpha$ *vanishing moments*:

$$\int_{-\infty}^{+\infty} t^k\,\psi(t)\,dt = 0 \quad \text{for} \quad 0 \leq k < n.$$

A wavelet with n vanishing moments is orthogonal to polynomials of degree $n-1$. Since $\alpha < n$, the polynomial p_v has degree at most $n-1$. With the change of variable $t' = (t-u)/s$, we verify that

$$W p_v(u,s) = \int_{-\infty}^{+\infty} p_v(t)\,\frac{1}{\sqrt{s}}\,\psi\left(\frac{t-u}{s}\right)dt = 0. \tag{6.7}$$

Since $f = p_v + \varepsilon_v$,

$$Wf(u,s) = W\varepsilon_v(u,s). \tag{6.8}$$

Section 6.1.3 explains how to measure α from $|Wf(u,s)|$ when u is in the neighborhood of v.

Multiscale Differential Operator

Theorem 6.2 proves that a wavelet with n vanishing moments can be written as the nth-order derivative of a function θ; the resulting wavelet transform is a multiscale differential operator. We suppose that ψ has a fast decay, which means that for any decay exponent $m \in \mathbb{N}$ there exists C_m such that

$$\forall t \in \mathbb{R}, \quad |\psi(t)| \leq \frac{C_m}{1+|t|^m}. \tag{6.9}$$

Theorem 6.2. A wavelet ψ with a fast decay has n vanishing moments if and only if there exists θ with a fast decay such that

$$\psi(t) = (-1)^n\,\frac{d^n\theta(t)}{dt^n}. \tag{6.10}$$

As a consequence

$$Wf(u,s) = s^n\,\frac{d^n}{du^n}(f \star \bar{\theta}_s)(u), \tag{6.11}$$

with $\bar{\theta}_s(t) = s^{-1/2}\theta(-t/s)$. Moreover, ψ has no more than n vanishing moments if and only if $\int_{-\infty}^{+\infty} \theta(t)\, dt \neq 0$.

Proof. The fast decay of ψ implies that $\hat{\psi}$ is $\mathbf{C}^\infty$. This is proved by setting $f = \hat{\psi}$ in Theorem 2.5. The integral of a function is equal to its Fourier transform evaluated at $\omega = 0$. The derivative property (2.22) implies that for any $k < n$,

$$\int_{-\infty}^{+\infty} t^k\, \psi(t)\, dt = (i)^k \hat{\psi}^{(k)}(0) = 0. \tag{6.12}$$

We can therefore make the factorization

$$\hat{\psi}(\omega) = (-i\omega)^n\, \hat{\theta}(\omega), \tag{6.13}$$

and $\hat{\theta}(\omega)$ is bounded. The fast decay of θ is proved with an induction on n. For $n = 1$,

$$\theta(t) = \int_{-\infty}^{t} \psi(u)\, du = \int_{t}^{+\infty} \psi(u)\, du,$$

and the fast decay of θ is derived from (6.9). We then similarly verify that increasing the order of integration by 1 up to n maintains the fast decay of θ.

Conversely, $|\hat{\theta}(\omega)| \leqslant \int_{-\infty}^{+\infty} |\theta(t)|\, dt < +\infty$, because θ has a fast decay. The Fourier transform of (6.10) yields (6.13), which implies that $\hat{\psi}^{(k)}(0) = 0$ for $k < n$. It follows from (6.12) that ψ has n vanishing moments.

To test whether ψ has more than n vanishing moments, we compute with (6.13)

$$\int_{-\infty}^{+\infty} t^n\, \psi(t)\, dt = (i)^n\, \hat{\psi}^{(n)}(0) = (-i)^n\, n!\, \hat{\theta}(0).$$

Clearly, ψ has no more than n vanishing moments if and only if $\hat{\theta}(0) = \int_{-\infty}^{+\infty} \theta(t)\, dt \neq 0$.

The wavelet transform (4.32) can be written

$$Wf(u,s) = f \star \bar{\psi}_s(u) \quad \text{with} \quad \bar{\psi}_s(t) = \frac{1}{\sqrt{s}}\, \psi\left(\frac{-t}{s}\right). \tag{6.14}$$

We derive from (6.10) that $\bar{\psi}_s(t) = s^n\, \frac{d^n \bar{\theta}_s(t)}{dt^n}$. Commuting the convolution and differentiation operators yields

$$Wf(u,s) = s^n f \star \frac{d^n \bar{\theta}_s}{dt^n}(u) = s^n\, \frac{d^n}{du^n}(f \star \bar{\theta}_s)(u).$$

■

If $K = \int_{-\infty}^{+\infty} \theta(t)\, dt \neq 0$, then the convolution $f \star \bar{\theta}_s(t)$ can be interpreted as a weighted average of f with a kernel dilated by s. So (6.11) proves that $Wf(u,s)$ is an nth-order derivative of an averaging of f over a domain proportional to s. Figure 6.1 shows a wavelet transform calculated with $\psi = -\theta'$, where θ is a Gaussian. The resulting $Wf(u,s)$ is the derivative of f averaged in the neighborhood of u with a Gaussian kernel dilated by s.

Since θ has a fast decay, one can verify that

$$\lim_{s \to 0} \frac{1}{\sqrt{s}}\, \bar{\theta}_s = K\, \delta,$$

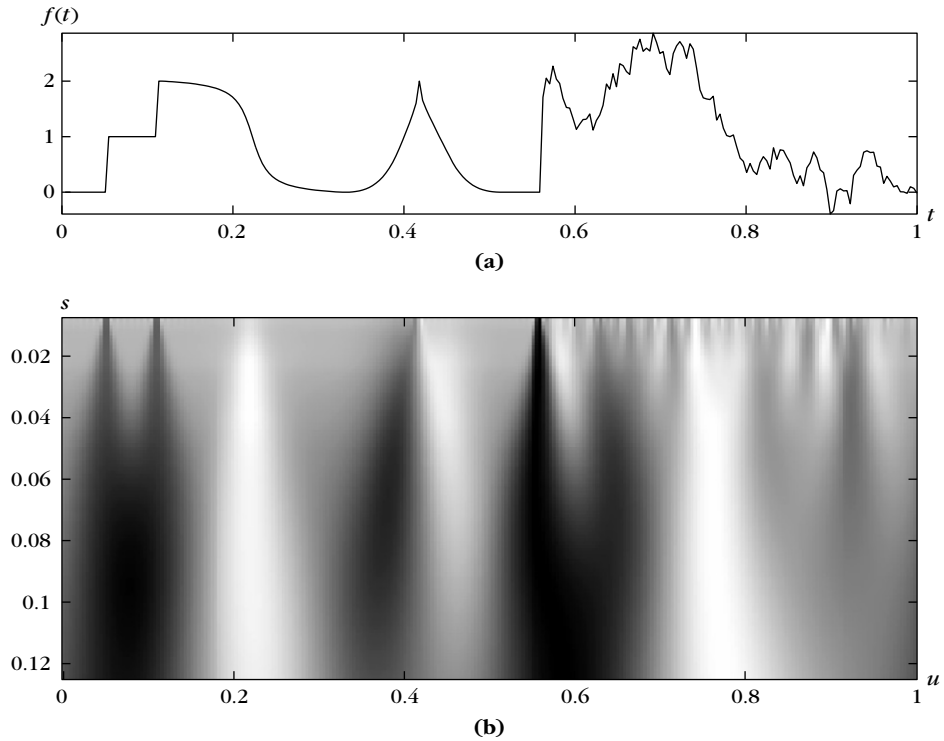

FIGURE 6.1

Wavelet transform $Wf(u,s)$ calculated with $\psi = -\theta'$ where θ is a Gaussian, for the signal f shown in **(a)**. Position parameter u and scale s vary, respectively, along the horizontal and vertical axes. **(b)** Black, gray, and white points correspond to positive, zero, and negative wavelet coefficients. Singularities create large-amplitude coefficients in their cone of influence.

in the sense of the weak convergence (A.30). This means that for any ϕ that is continuous at u,

$$\lim_{s \to 0} \phi \star \frac{1}{\sqrt{s}} \bar{\theta}_s(u) = K\, \phi(u).$$

If f is n times continuously differentiable in the neighborhood of u, then (6.11) implies that

$$\lim_{s \to 0} \frac{Wf(u,s)}{s^{n+1/2}} = \lim_{s \to 0} f^{(n)} \star \frac{1}{\sqrt{s}} \bar{\theta}_s(u) = K f^{(n)}(u). \tag{6.15}$$

In particular, if f is $\mathbf{C}^n$ with a bounded nth-order derivative, then $|Wf(u,s)| = O(s^{n+1/2})$. This is a first relation between the decay of $|Wf(u,s)|$ when s decreases and the uniform regularity of f. Finer relations are studied in the next section.

6.1.3 Regularity Measurements with Wavelets

The decay of the wavelet transform amplitude across scales is related to the uniform and pointwise Lipschitz regularity of the signal. Measuring this asymptotic decay is equivalent to zooming into signal structures with a scale that goes to zero. We suppose that the wavelet ψ has n vanishing moments and is $\mathbf{C}^n$ with derivatives that have a fast decay. This means that for any $0 \leq k \leq n$ and $m \in \mathbb{N}$ there exists C_m such that

$$\forall t \in \mathbb{R}, \quad |\psi^{(k)}(t)| \leq \frac{C_m}{1+|t|^m}. \tag{6.16}$$

Theorem 6.3 relates the uniform Lipschitz regularity of f on an interval to the amplitude of its wavelet transform at fine scales.

Theorem 6.3. If $f \in \mathbf{L}^2(\mathbb{R})$ is uniformly Lipschitz $\alpha \leq n$ over $[a, b]$, then there exists $A > 0$ such that

$$\forall (u, s) \in [a, b] \times \mathbb{R}^+, \quad |Wf(u, s)| \leq A\, s^{\alpha+1/2}. \tag{6.17}$$

Conversely, suppose that f is bounded and that $Wf(u, s)$ satisfies (6.17) for an $\alpha < n$ that is not an integer. Then f is uniformly Lipschitz α on $[a+\varepsilon, b-\varepsilon]$, for any $\varepsilon > 0$.

Proof. This theorem is proved with minor modifications in the proof of Theorem 6.4. Since f is Lipschitz α at any $v \in [a, b]$, Theorem 6.4 shows in (6.20) that

$$\forall (u, s) \in \mathbb{R} \times \mathbb{R}^+, \quad |Wf(u, s)| \leq A\, s^{\alpha+1/2}\left(1 + \left|\frac{u-v}{s}\right|^{\alpha}\right).$$

For $u \in [a, b]$, we can choose $v = u$, which implies that $|Wf(u, s)| \leq A\, s^{\alpha+1/2}$. We verify from the proof of (6.20) that the constant A does not depend on v because the Lipschitz regularity is uniform over $[a, b]$.

To prove that f is uniformly Lipschitz α over $[a+\varepsilon, b-\varepsilon]$, we must verify that there exists K such that for all $v \in [a+\varepsilon, b-\varepsilon]$ we can find a polynomial p_v of degree $\lfloor \alpha \rfloor$ such that

$$\forall t \in \mathbb{R}, \quad |f(t) - p_v(t)| \leq K\, |t-v|^{\alpha}. \tag{6.18}$$

When $t \notin [a+\varepsilon/2, b-\varepsilon/2]$, then $|t-v| \geq \varepsilon/2$, and since f is bounded, (6.18) is verified with a constant K that depends on ε. For $t \in [a+\varepsilon/2, b-\varepsilon/2]$, the proof follows the same derivations as the proof of pointwise Lipschitz regularity from (6.21) in Theorem 6.4. The upper bounds (6.26) and (6.27) are replaced by

$$\forall t \in [a+\varepsilon/2, b-\varepsilon/2], \quad |\Delta_j^{(k)}(t)| \leq K\, 2^{(\alpha-k)j} \quad \text{for } 0 \leq k \leq \lfloor \alpha \rfloor + 1. \tag{6.19}$$

This inequality is verified by computing an upper-bound integral similar to (6.25) but which is divided in two—$u \in [a, b]$ and $u \notin [a, b]$. When $u \in [a, b]$, the condition (6.21) is replaced by $|Wf(u, s)| \leq A\, s^{\alpha+1/2}$ in (6.25). When $u \notin [a, b]$, we just use the fact that $|Wf(u, s)| \leq \|f\|\, \|\psi\|$ and derive (6.19) from the fast decay of $|\psi^{(k)}(t)|$, by observing that $|t-u| \geq \varepsilon/2$ for $t \in [a+\varepsilon/2, b-\varepsilon/2]$. The constant K depends on A and ε but not on v. The proof then proceeds like the proof of Theorem 6.4, and since the resulting

constant K in (6.29) does not depend on v, the Lipschitz regularity is uniform over $[a-\varepsilon, b+\varepsilon]$. ■

The inequality (6.17) is really a condition on the asymptotic decay of $|Wf(u,s)|$ when s goes to zero. At large scales it does not introduce any constraint since the Cauchy-Schwarz inequality guarantees that the wavelet transform is bounded:

$$|Wf(u,s)| = |\langle f, \psi_{u,s}\rangle| \leq \|f\|\,\|\psi\|.$$

When the scale s decreases, $Wf(u,s)$ measures fine-scale variations in the neighborhood of u. Theorem 6.3 proves that $|Wf(u,s)|$ decays like $s^{\alpha+1/2}$ over intervals where f is uniformly Lipschitz α.

Observe that the upper bound (6.17) is similar to the sufficient Fourier condition of theorem (6.1), which supposes that $|\hat{f}(\omega)|$ decays faster than $\omega^{-\alpha}$. The wavelet scale s plays the role of a "localized" inverse frequency ω^{-1}. As opposed to the Fourier transform theorem (6.1), the wavelet transform gives a Lipschitz regularity condition that is localized over any finite interval and it provides a necessary condition that is nearly sufficient. When $[a,b]=\mathbb{R}$, then (6.17) is a necessary and sufficient condition for f to be uniformly Lipschitz α on $\mathbb{R}$.

If ψ has exactly n vanishing moments, then the wavelet transform decay gives no information concerning the Lipschitz regularity of f for $\alpha > n$. If f is uniformly Lipschitz $\alpha > n$, then it is $\mathbf{C}^n$ and (6.15) proves that $\lim_{s\to 0} s^{-n-1/2}\,Wf(u,s) = Kf^{(n)}(u)$ with $K \neq 0$. This proves that $|Wf(u,s)| \sim s^{n+1/2}$ at fine scales despite the higher regularity of f.

If the Lipschitz exponent α is an integer, then (6.17) is not sufficient to prove that f is uniformly Lipschitz α. When $[a,b]=\mathbb{R}$, if $\alpha = 1$ and ψ has two vanishing moments, then the class of functions that satisfy (6.17) is called the *Zygmund class* [44]. It is slightly larger than the set of functions that are uniformly Lipschitz 1. For example, $f(t) = t\,\log_e t$ belongs to the Zygmund class although it is not Lipschitz 1 at $t=0$.

Pointwise Lipschitz Regularity

The study of pointwise Lipschitz exponents with the wavelet transform is a delicate and beautiful topic that finds its mathematical roots in the characterization of Sobolev spaces by Littlewood and Paley in the 1930s. Characterizing the regularity of f at a point v can be difficult because f may have very different types of singularities that are aggregated in the neighborhood of v. In 1984, Bony [118] introduced the "two-microlocalization" theory, which refines the Littlewood-Paley approach to provide pointwise characterization of singularities that he used to study the solution of hyperbolic partial differential equations. These technical results became much simpler through the work of Jaffard [312] who proved that the two-microlocalization properties are equivalent to specific decay conditions on the wavelet transform amplitude. Theorem 6.4 gives a necessary and a sufficient condition on the wavelet transform for estimating the Lipschitz regularity of f at a point v. Remember that the wavelet ψ has n vanishing moments and n derivatives having a fast decay.

Theorem 6.4: *Jaffard.* If $f \in \mathbf{L}^2(\mathbb{R})$ is Lipschitz $\alpha \leqslant n$ at v, then there exists A such that

$$\forall (u,s) \in \mathbb{R} \times \mathbb{R}^+, \quad |Wf(u,s)| \leqslant A\, s^{\alpha+1/2} \left(1 + \left|\frac{u-v}{s}\right|^{\alpha}\right). \tag{6.20}$$

Conversely, if $\alpha < n$ is not an integer and there exist A and $\alpha' < \alpha$ such that

$$\forall (u,s) \in \mathbb{R} \times \mathbb{R}^+, \quad |Wf(u,s)| \leqslant A\, s^{\alpha+1/2} \left(1 + \left|\frac{u-v}{s}\right|^{\alpha'}\right), \tag{6.21}$$

then f is Lipschitz α at v.

Proof. The necessary condition is relatively simple to prove but the sufficient condition is much more difficult.

Proof of (6.20). Since f is Lipschitz α at v, there exists a polynomial p_v of degree $\lfloor \alpha \rfloor < n$ and K such that $|f(t) - p_v(t)| \leqslant K|t-v|^{\alpha}$. Since ψ has n vanishing moments, we saw in (6.7) that $Wp_v(u,s) = 0$, and thus

$$\begin{aligned} |Wf(u,s)| &= \left| \int_{-\infty}^{+\infty} \big(f(t) - p_v(t)\big)\, \frac{1}{\sqrt{s}}\, \psi\Big(\frac{t-u}{s}\Big)\, dt \right| \\ &\leqslant \int_{-\infty}^{+\infty} K\, |t-v|^{\alpha}\, \frac{1}{\sqrt{s}} \left| \psi\Big(\frac{t-u}{s}\Big) \right| dt. \end{aligned}$$

The change of variable $x = (t-u)/s$ gives

$$|Wf(u,s)| \leqslant \sqrt{s} \int_{-\infty}^{+\infty} K\, |sx + u - v|^{\alpha}\, |\psi(x)|\, dx.$$

Since $|a+b|^{\alpha} \leqslant 2^{\alpha}\, (|a|^{\alpha} + |b|^{\alpha})$,

$$|Wf(u,s)| \leqslant K\, 2^{\alpha} \sqrt{s} \left(s^{\alpha} \int_{-\infty}^{+\infty} |x|^{\alpha}\, |\psi(x)|\, dx + |u-v|^{\alpha} \int_{-\infty}^{+\infty} |\psi(x)|\, dx \right),$$

which proves (6.20).

Proof of (6.21). The wavelet reconstruction formula (4.37) proves that f can be decomposed in a Littlewood-Paley–type sum

$$f(t) = \sum_{j=-\infty}^{+\infty} \Delta_j(t) \tag{6.22}$$

with

$$\Delta_j(t) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{2^j}^{2^{j+1}} Wf(u,s)\, \frac{1}{\sqrt{s}}\, \psi\left(\frac{t-u}{s}\right) \frac{ds}{s^2}\, du. \tag{6.23}$$

Let $\Delta_j^{(k)}$ be its kth-order derivative. To prove that f is Lipschitz α at v we shall approximate f with a polynomial that generalizes the Taylor polynomial

$$p_v(t) = \sum_{k=0}^{\lfloor \alpha \rfloor} \left(\sum_{j=-\infty}^{+\infty} \Delta_j^{(k)}(v) \right) \frac{(t-v)^k}{k!}. \tag{6.24}$$

If f is n times differentiable at v, then p_v corresponds to the Taylor polynomial; however, this is not necessarily true. We shall first prove that $\sum_{j=-\infty}^{+\infty} \Delta_j^{(k)}(v)$ is finite by getting upper bounds on $|\Delta_j^{(k)}(t)|$. These sums may be thought of as a generalization of pointwise derivatives.

To simplify the notation, we denote by K a generic constant that may change value from one line to the next but that does not depend on j and t. The hypothesis (6.21) and the asymptotic decay condition (6.16) imply that

$$\begin{aligned} |\Delta_j(t)| &= \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{2^j}^{2^{j+1}} A\, s^\alpha \left(1 + \left|\frac{u-v}{s}\right|^{\alpha'}\right) \frac{C_m}{1+|(t-u)/s|^m} \frac{ds}{s^2}\, du \\ &\leqslant K \int_{-\infty}^{+\infty} 2^{\alpha j} \left(1 + \left|\frac{u-v}{2^j}\right|^{\alpha'}\right) \frac{1}{1+\left|(t-u)/2^j\right|^m} \frac{du}{2^j}. \end{aligned} \tag{6.25}$$

Since $|u-v|^{\alpha'} \leqslant 2^{\alpha'}(|u-t|^{\alpha'} + |t-v|^{\alpha'})$, the change of variable $u' = 2^{-j}(u-t)$ yields

$$|\Delta_j(t)| \leqslant K\, 2^{\alpha j} \int_{-\infty}^{+\infty} \frac{1 + |u'|^{\alpha'} + \left|(v-t)/2^j\right|^{\alpha'}}{1+|u'|^m}\, du'.$$

Choosing $m = \alpha' + 2$ yields

$$|\Delta_j(t)| \leqslant K\, 2^{\alpha j} \left(1 + \left|\frac{v-t}{2^j}\right|^{\alpha'}\right). \tag{6.26}$$

The same derivations applied to the derivatives of $\Delta_j(t)$ yield

$$\forall k \leqslant \lfloor \alpha \rfloor + 1, \quad |\Delta_j^{(k)}(t)| \leqslant K\, 2^{(\alpha-k)j} \left(1 + \left|\frac{v-t}{2^j}\right|^{\alpha'}\right). \tag{6.27}$$

At $t = v$, it follows that

$$\forall k \leqslant \lfloor \alpha \rfloor, \quad |\Delta_j^{(k)}(v)| \leqslant K\, 2^{(\alpha-k)j}. \tag{6.28}$$

This guarantees a fast decay of $|\Delta_j^{(k)}(v)|$ when 2^j goes to zero, because α is not an integer so $\alpha > \lfloor \alpha \rfloor$. At large scales 2^j, since $|Wf(u,s)| \leqslant \|f\|\, \|\psi\|$ with the change of variable $u' = (t-u)/s$ in (6.23), we have

$$|\Delta_j^{(k)}(v)| \leqslant \frac{\|f\|\, \|\psi\|}{C_\psi} \int_{-\infty}^{+\infty} |\psi^{(k)}(u')|\, du' \int_{2^j}^{2^{j+1}} \frac{ds}{s^{3/2+k}},$$

therefore $|\Delta_j^{(k)}(v)| \leqslant K 2^{-(k+1/2)j}$. Together with (6.28) this proves that the polynomial p_v defined in (6.24) has coefficients that are finite.

With the Littlewood-Paley decomposition (6.22), we compute

$$|f(t) - p_v(t)| = \left| \sum_{j=-\infty}^{+\infty} \left(\Delta_j(t) - \sum_{k=0}^{\lfloor \alpha \rfloor} \Delta_j^{(k)}(v)\, \frac{(t-v)^k}{k!} \right) \right|.$$

The sum over scales is divided in two at 2^j such that $2^j \geqslant |t-v| \geqslant 2^{j-1}$. For $j \geqslant J$, we can use the classical Taylor theorem to bound the Taylor expansion of Δ_j:

$$I = \sum_{j=J}^{+\infty} \left| \Delta_j(t) - \sum_{k=0}^{\lfloor \alpha \rfloor} \Delta_j^{(k)}(v) \frac{(t-v)^k}{k!} \right|$$

$$\leqslant \sum_{j=J}^{+\infty} \frac{(t-v)^{\lfloor \alpha \rfloor + 1}}{(\lfloor \alpha \rfloor + 1)!} \sup_{h \in [t,v]} |\Delta_j^{\lfloor \alpha \rfloor + 1}(h)|.$$

Inserting (6.27) yields

$$I \leqslant K\, |t-v|^{\lfloor \alpha \rfloor + 1} \sum_{j=J}^{+\infty} 2^{-j(\lfloor \alpha \rfloor + 1 - \alpha)} \left| \frac{v-t}{2^j} \right|^{\alpha'},$$

and since $2^j \geqslant |t-v| \geqslant 2^{j-1}$, we get $I \leqslant K\, |v-t|^{\alpha}$.

Let us now consider the case $j < J$:

$$II = \sum_{j=-\infty}^{j-1} \left| \Delta_j(t) - \sum_{k=0}^{\lfloor \alpha \rfloor} \Delta_j^{(k)}(v)\, \frac{(t-v)^k}{k!} \right|$$

$$\leqslant K \sum_{j=-\infty}^{j-1} \left(2^{\alpha j} \left(1 + \left| \frac{v-t}{2^j} \right|^{\alpha'} \right) + \sum_{k=0}^{\lfloor \alpha \rfloor} \frac{(t-v)^k}{k!}\, 2^{j(\alpha-k)} \right)$$

$$\leqslant K \left(2^{\alpha J} + 2^{(\alpha-\alpha')J}\, |t-v|^{\alpha'} + \sum_{k=0}^{\lfloor \alpha \rfloor} \frac{(t-v)^k}{k!}\, 2^{j(\alpha-k)} \right),$$

and since $2^j \geqslant |t-v| \geqslant 2^{j-1}$, we get $II \leqslant K\, |v-t|^{\alpha}$. As a result,

$$|f(t) - p_v(t)| \leqslant I + II \leqslant K\, |v-t|^{\alpha}, \tag{6.29}$$

which proves that f is Lipschitz α at v. ■

Cone of Influence

To interpret more easily the necessary condition (6.20) and the sufficient condition (6.21), we shall suppose that ψ has a compact support equal to $[-C, C]$. The *cone of influence* of v in the scale–space plane is the set of points (u, s) such that v is included in the support of $\psi_{u,s}(t) = s^{-1/2}\, \psi((t-u)/s)$. Since the support of $\psi((t-u)/s)$ is equal to $[u - Cs, u + Cs]$, the cone of influence of v is defined by

$$|u - v| \leqslant Cs. \tag{6.30}$$

It is illustrated in Figure 6.2. If u is in the cone of influence of v, then $Wf(u,s) = \langle f, \psi_{u,s} \rangle$ depends on the value of f in the neighborhood of v. Since $|u-v|/s \leqslant C$, the conditions (6.20, 6.21) can be written as

$$|Wf(u,s)| \leqslant A'\, s^{\alpha + 1/2},$$

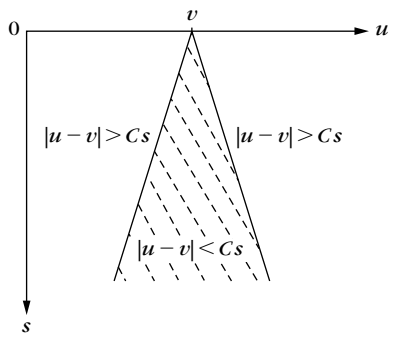

FIGURE 6.2

The cone of influence of an abscissa v consists of the scale–space points (u, s) for which the support of $\psi_{u,s}$ intersects $t = v$.

which is identical to the uniform Lipschitz condition (6.17) given by Theorem 6.3. In Figure 6.1, the high-amplitude wavelet coefficients are in the cone of influence of each singularity.

Oscillating Singularities

It may seem surprising that (6.20) and (6.21) also impose a condition on the wavelet transform outside the cone of influence of v. Indeed, this corresponds to wavelets of which the support does not intersect v. For $|u - v| > Cs$, we get

$$|Wf(u, s)| \leq A' s^{\alpha - \alpha' + 1/2} |u - v|^{\alpha}. \tag{6.31}$$

We shall see that it is indeed necessary to impose this decay when u tends to v in order to control the oscillations of f that might generate singularities.

Let us consider the generic example of a highly oscillatory function

$$f(t) = \sin \frac{1}{t},$$

which is discontinuous at $v = 0$ because of the acceleration of its oscillations. Since ψ is a smooth $\mathbf{C}^n$ function, if it is centered close to zero, then the rapid oscillations of $\sin t^{-1}$ produce a correlation integral $\langle \sin t^{-1}, \psi_{u,s} \rangle$ that is very small. With an integration by parts, one can verify that if (u, s) is in the cone of influence of $v = 0$, then $|Wf(u, s)| \leq A s^{2+1/2}$. This looks as if f is Lipschitz 2 at 0. However, Figure 6.3 shows high-energy wavelet coefficients outside the cone of influence of $v = 0$, which are responsible for the discontinuity. To guarantee that f is Lipschitz α, the amplitude of such coefficients is controlled by the upper bound (6.31).

To explain why high-frequency oscillations appear outside the cone of influence of v, we use the results of Section 4.4.3 on the estimation of instantaneous frequencies with wavelet ridges. The instantaneous frequency of $\sin t^{-1} = \sin \theta(t)$ is $|\theta'(t)| = t^{-2}$. Let ψ^a be the analytic part of ψ, defined in (4.47). The corresponding

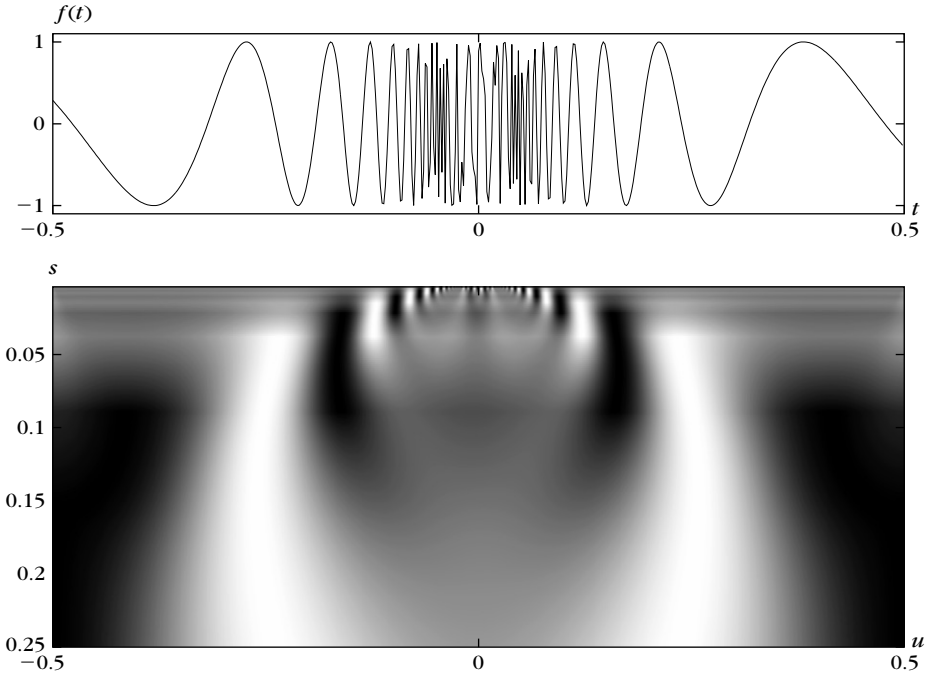

FIGURE 6.3

Wavelet transform of $f(t) = \sin(a\,t^{-1})$ calculated with $\psi = -\theta'$, where θ is a Gaussian. High-amplitude coefficients are along a parabola outside the cone of influence of $t = 0$.

complex analytic wavelet transform is $W^a f(u,s) = \langle f, \psi^a_{u,s} \rangle$. It was proved in (4.109) that for a fixed time u, the maximum of $s^{-1/2}|W^a f(u,s)|$ is located at the scale

$$s(u) = \frac{\eta}{\theta'(u)} = \eta\, u^2,$$

where η is the center frequency of $\hat{\psi}^a(\omega)$. When u varies, the set of points $(u, s(u))$ defines a *ridge* that is a parabola located outside the cone of influence of $v = 0$ in the plane (u,s). Since $\psi = \mathrm{Re}[\psi^a]$, the real wavelet transform is

$$Wf(u,s) = \mathrm{Re}[W^a f(u,s)].$$

The high-amplitude values of $Wf(u,s)$ are thus located along the same parabola ridge curve in the scale–space plane, which clearly appears in Figure 6.3. Real wavelet coefficients $Wf(u,s)$ change signs along the ridge because of the variations of the complex phase of $W^a f(u,s)$.

The example of $f(t) = \sin t^{-1}$ can be extended to general oscillating singularities [32]. A function f has an oscillating singularity at v if there exist $\alpha \geqslant 0$ and $\beta > 0$ such that for t in a neighborhood of v,

$$f(t) \sim |t - v|^{\alpha}\, g\left(\frac{1}{|t-v|^{\beta}}\right),$$

where $g(t)$ is a $\mathbf{C}^\infty$ oscillating function that has primitives bounded at any order. The function $g(t) = \sin t^{-1}$ is a typical example. The oscillations have an instantaneous frequency $\theta'(t)$ that increases to infinity faster than $|t|^{-1}$ when t goes to v. High-energy wavelet coefficients are located along the ridge $s(u) = \eta/\theta'(u)$, and this curve is necessarily outside the cone of influence $|u - v| \leqslant Cs$.

6.2 WAVELET TRANSFORM MODULUS MAXIMA

Theorems 6.3 and 6.4 prove that the local Lipschitz regularity of f at v depends on the decay at fine scales of $|Wf(u, s)|$ in the neighborhood of v. Measuring this decay directly in the time-scale plane (u, s) is not necessary. The decay of $|Wf(u, s)|$ can indeed be controlled from its local maxima values. Section 6.2.1 studies the detection and characterization of singularities from wavelet local maxima. Signal approximations are recovered in Section 6.2.2, from the scale–space support of these local maxima at dyadic scales.

6.2.1 Detection of Singularities

Singularities are detected by finding the abscissa where the wavelet modulus maxima converge at fine scales. A wavelet *modulus maximum* is defined as a point (u_0, s_0) such that $|Wf(u, s_0)|$ is locally maximum at $u = u_0$. This implies that

$$\frac{\partial Wf(u_0, s_0)}{\partial u} = 0.$$

This local maximum should be a strict local maximum in either the right or the left neighborhood of u_0 to avoid having any local maxima when $|Wf(u, s_0)|$ is constant. We call any connected curve $s(u)$ in the scale–space plane (u, s) along which all points are modulus maxima a *maxima line*. (See Figure 6.5b on page 218, which shows the wavelet modulus maxima of a signal.)

To better understand the properties of these maxima, the wavelet transform is written as a multiscale differential operator. Theorem 6.2 proves that if ψ has exactly n vanishing moments and a compact support, then there exists θ of compact support such that $\psi = (-1)^n \theta^{(n)}$ with $\int_{-\infty}^{+\infty} \theta(t)\, dt \neq 0$. The wavelet transform is rewritten in (6.11) as a multiscale differential operator

$$Wf(u, s) = s^n \frac{d^n}{du^n} (f \star \bar{\theta}_s)(u). \tag{6.32}$$

If the wavelet has only one vanishing moment, wavelet modulus maxima are the maxima of the first-order derivative of f smoothed by $\bar{\theta}_s$, as illustrated by Figure 6.4. These multiscale modulus maxima are used to locate discontinuities and edges in images. If the wavelet has two vanishing moments, the modulus maxima correspond to high curvatures. Theorem 6.5 proves that if $Wf(u, s)$ has no modulus maxima at fine scales, then f is locally regular.

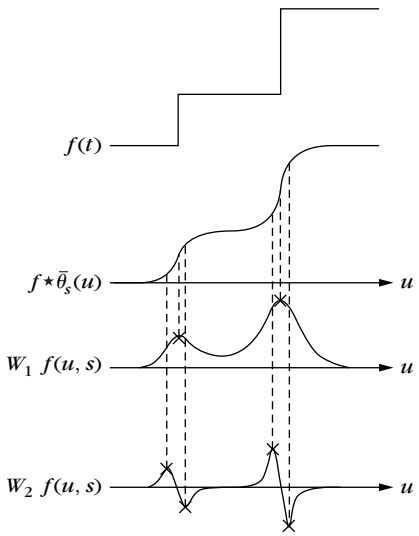

FIGURE 6.4

The convolution $f \star \bar{\theta}_s(u)$ averages f over a domain proportional to s. If $\psi = -\theta'$, then $W_1 f(u,s) = s\frac{d}{du}(f \star \bar{\theta}_s)(u)$ has modulus maxima at sharp variation points of $f \star \bar{\theta}_s(u)$. If $\psi = \theta''$, then the modulus maxima of $W_2 f(u,s) = s^2 \frac{d^2}{du^2}(f \star \bar{\theta}_s)(u)$ correspond to locally maximum curvatures.

Theorem 6.5: *Hwang, Mallat.* Suppose that ψ is $\mathbf{C}^n$ with a compact support, and $\psi = (-1)^n \theta^{(n)}$ with $\int_{-\infty}^{+\infty} \theta(t)dt \neq 0$. Let $f \in \mathbf{L^1}[a,b]$. If there exists $s_0 > 0$ such that $|Wf(u,s)|$ has no local maximum for $u \in [a,b]$ and $s < s_0$, then f is uniformly Lipschitz n on $[a+\varepsilon, b-\varepsilon]$, for any $\varepsilon > 0$.

This theorem is proved in [364]. It implies that f can be singular (not Lipschitz 1) at a point v only if there is a sequence of wavelet maxima points $(u_p, s_p)_{p\in\mathbb{N}}$ that converges toward v at fine scales:

$$\lim_{p\to+\infty} u_p = v \quad \text{and} \quad \lim_{p\to+\infty} s_p = 0.$$

These modulus maxima points may or may not be along the same maxima line. This result guarantees that all singularities are detected by following the wavelet transform modulus maxima at fine scales. Figure 6.5 gives an example where all singularities are located by following the maxima lines.

Maxima Propagation

For all $\psi = (-1)^n \theta^{(n)}$, we are not guaranteed that a modulus maxima located at (u_0, s_0) belongs to a maxima line that propagates toward finer scales. When s decreases, $Wf(u,s)$ may have no more maxima in the neighborhood of $u = u_0$.

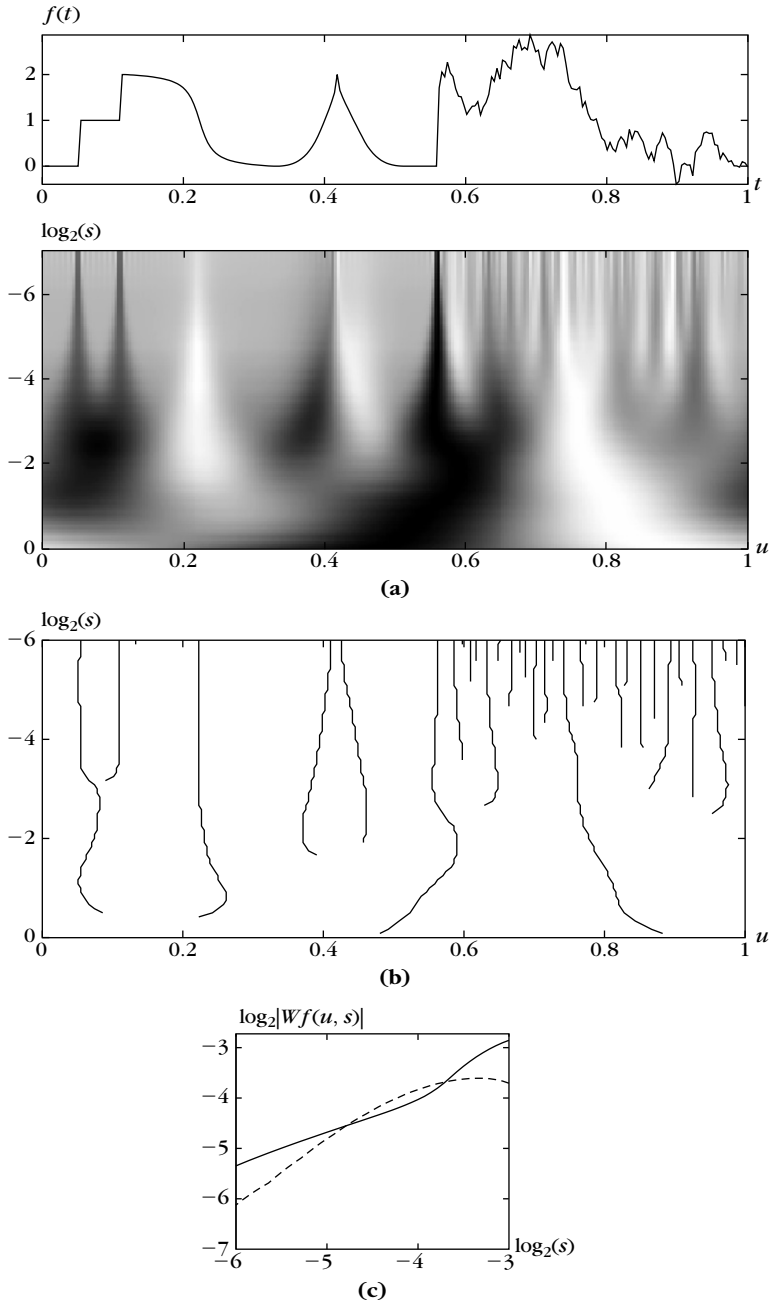

FIGURE 6.5

(a) Wavelet transform $Wf(u, s)$; the horizontal and vertical axes give u and $\log_2 s$, respectively. **(b)** Modulus maxima of $Wf(u, s)$. **(c)** The full line gives the decay of $\log_2 |Wf(u, s)|$ as a function of $\log_2 s$ along the maxima line that converges to the abscissa $t = 0.05$. The dashed line gives $\log_2 |Wf(u, s)|$ along the left maxima line that converges to $t = 0.42$.

Theorem 6.6 proves that this is never the case if θ is a Gaussian. The wavelet transform $Wf(u,s)$ can then be written as the solution of the heat-diffusion equation, where s is proportional to the diffusion time. The maximum principle applied to the heat-diffusion equation proves that maxima may not disappear when s decreases. Applications of the heat-diffusion equation to the analysis of multiscale averaging have been studied by several computer vision researchers [309, 330, 496].

Theorem 6.6: *Hummel, Poggio, Yuille.* Let $\psi = (-1)^n \theta^{(n)}$, where θ is a Gaussian. For any $f \in \mathbf{L}^2(\mathbb{R})$, the modulus maxima of $Wf(u,s)$ belong to connected curves that are never interrupted when the scale decreases.

Proof. To simplify the proof, we suppose that θ is a normalized Gaussian $\theta(t) = 2^{-1}\pi^{-1/2} \exp(-t^2/4)$ and that the Fourier transform is $\hat{\theta}(\omega) = \exp(-\omega^2)$. Theorem 6.2 proves that

$$Wf(u,s) = s^n f^{(n)} \star \theta_s(u), \tag{6.33}$$

where the nth derivative $f^{(n)}$ is defined in the sense of distributions. Let τ be the diffusion time. The solution of

$$\frac{\partial g(u,\tau)}{\partial \tau} = \frac{\partial^2 g(u,\tau)}{\partial u^2} \tag{6.34}$$

with initial condition $g(u,0) = g_0(u)$ is obtained by computing the Fourier transform with respect to u of (6.34):

$$\frac{\partial \hat{g}(\omega,u)}{\partial \tau} = -\omega^2 \hat{g}(\omega,\tau).$$

It follows that $\hat{g}(\omega,\tau) = \hat{g}_0(\omega) \exp(-\tau\omega^2)$ and thus,

$$g(u,\tau) = \frac{1}{\sqrt{\tau}} g_0 \star \theta_{\sqrt{\tau}}(u).$$

For $\tau = s^2$, setting $g_0 = f^{(n)}$ and inserting (6.33) yields $Wf(u,s) = s^{n+1} g(u,s^2)$. Thus, the wavelet transform is proportional to a heat diffusion with initial condition $f^{(n)}$.

The maximum principle for the parabolic heat equation [35] proves that a global maximum of $|g(u,s^2)|$ for $(u,s) \in [a,b] \times [s_0,s_1]$ is necessarily either on the boundary $u = a, b$ or at $s = s_0$. A modulus maxima of $Wf(u,s)$ at (u_1,s_1) is a local maxima of $|g(u,s^2)|$ for a fixed s and varying u. Suppose that a line of modulus maxima is interrupted at (u_1,s_1), with $s_1 > 0$. One can then verify that there exists $\varepsilon > 0$ such that a global maximum of $|g(u,s^2)|$ over $[u_1-\varepsilon, u_1+\varepsilon] \times [s_1-\varepsilon, s_1]$ is at (u_1,s_1). This contradicts the maximum principle, and thus proves that all modulus maxima propagate toward finer scales. ■

Derivatives of Gaussians are most often used to guarantee that all maxima lines propagate up to the finest scales. Chaining together maxima into maxima lines is also a procedure for removing spurious modulus maxima created by numerical errors in regions where the wavelet transform is close to zero.

Isolated Singularities

A wavelet transform may have a sequence of local maxima that converge to an abscissa v even though f is perfectly regular at v. This is the case of the maxima line of Figure 6.5 that converges to the abscissa $v = 0.23$. To detect singularities it is therefore not sufficient to follow the wavelet modulus maxima across scales. The Lipschitz regularity is calculated from the decay of the modulus maxima amplitude.

Let us suppose that for $s < s_0$ all modulus maxima that converge to v are included in a cone

$$|u - v| \leq Cs. \tag{6.35}$$

This means that f does not have oscillations that accelerate in the neighborhood of v. The potential singularity at v is necessarily isolated. Indeed, we can derive from Theorem 6.5 that the absence of maxima outside the cone of influence implies that f is uniformly Lipschitz n in the neighborhood of any $t \neq v$ with $t \in (v - Cs_0, v + Cs_0)$. The decay of $|Wf(u, s)|$ in the neighborhood of v is controlled by the decay of the modulus maxima included in the cone $|u - v| \leq Cs$. Theorem 6.3 implies that f is uniformly Lipschitz α in the neighborhood of v if and only if there exists $A > 0$ such that each modulus maximum (u, s) in the cone (6.35) satisfies

$$|Wf(u, s)| \leq A\, s^{\alpha + 1/2}, \tag{6.36}$$

which is equivalent to

$$\log_2 |Wf(u, s)| \leq \log_2 A + \left(\alpha + \frac{1}{2}\right) \log_2 s. \tag{6.37}$$

Thus, the Lipschitz regularity at v is the maximum slope of $\log_2 |Wf(u, s)|$ as a function of $\log_2 s$ along the maxima lines converging to v.

In numerical calculations, the finest scale of the wavelet transform is limited by the resolution of the discrete data. From a sampling at intervals N^{-1}, Section 4.3.3 computes the discrete wavelet transform at scales $s \geq \lambda N^{-1}$ where λ is large enough to avoid sampling coarsely the wavelets at the finest scale. The Lipschitz regularity α of a singularity is then estimated by measuring the decay slope of $\log_2 |Wf(u, s)|$ as a function of $\log_2 s$ for $2^j \geq s \geq \lambda N^{-1}$. The largest scale 2^j should be smaller than the distance between two consecutive singularities to avoid having other singularities influence the value of $Wf(u, s)$. The sampling interval N^{-1} must be small enough to measure α accurately. The signal in Figure 6.5(a) is defined by $N = 256$ samples. Figure 6.5(c) shows the decay of $\log_2 |Wf(u, s)|$ along the maxima line converging to $t = 0.05$. It has slope $\alpha + 1/2 \approx 1/2$ for $2^{-4} \geq s \geq 2^{-6}$. As expected, $\alpha = 0$ because the signal is discontinuous at $t = 0.05$. Along the second maxima line converging to $t = 0.42$ the slope is $\alpha + 1/2 \approx 1$, which indicates that the singularity is Lipschitz $1/2$.

When f is a function with singularities that are not isolated, finite resolution measurements are not sufficient to distinguish individual singularities. Section 6.4 describes a global approach that computes the singularity spectrum of multifractals by taking advantage of their self-similarity.

Smoothed Singularities

The signal may have important variations that are infinitely continuously differentiable. For example, at the border of a shadow the gray level of an image varies quickly but is not discontinuous because of the diffraction effect. The smoothness of these transitions is modeled as a diffusion with a Gaussian kernel that has a variance that is measured from the decay of wavelet modulus maxima.

In the neighborhood of a sharp transition at v, we suppose that

$$f(t) = f_0 \star g_\sigma(t), \tag{6.38}$$

where g_σ is a Gaussian of variance σ^2:

$$g_\sigma(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-t^2}{2\sigma^2}\right). \tag{6.39}$$

If f_0 has a Lipschitz α singularity at v that is isolated and nonoscillating, it is uniformly Lipschitz α in the neighborhood of v. For wavelets that are derivatives of Gaussians, Theorem 6.7 [367] relates the decay of the wavelet transform to σ and α.

Theorem 6.7. Let $\psi = (-1)^n \theta^{(n)}$ with $\theta(t) = \lambda \exp(-t^2/(2\beta^2))$. If $f = f_0 \star g_\sigma$ and f_0 is uniformly Lipschitz α on $[v-h, v+h]$, then there exists A such that

$$\forall (u,s) \in [v-h, v+h] \times \mathbb{R}^+, \quad |Wf(u,s)| \leq A s^{\alpha+1/2} \left(1 + \frac{\sigma^2}{\beta^2 s^2}\right)^{-(n-\alpha)/2}. \tag{6.40}$$

Proof. The wavelet transform can be written as

$$Wf(u,s) = s^n \frac{d^n}{du^n}(f \star \bar{\theta}_s)(u) = s^n \frac{d^n}{du^n}(f_0 \star g_\sigma \star \bar{\theta}_s)(u). \tag{6.41}$$

Since θ is a Gaussian, one can verify with a Fourier transform calculation that

$$\bar{\theta}_s \star g_\sigma(t) = \sqrt{\frac{s}{s_0}}\, \bar{\theta}_{s_0}(t) \quad \text{with} \quad s_0 = \sqrt{s^2 + \frac{\sigma^2}{\beta^2}}. \tag{6.42}$$

Inserting this result in (6.41) yields

$$Wf(u,s) = s^n \sqrt{\frac{s}{s_0}} \frac{d^n}{du^n}(f_0 \star \bar{\theta}_{s_0})(u) = \left(\frac{s}{s_0}\right)^{n+1/2} Wf_0(u, s_0). \tag{6.43}$$

Since f_0 is uniformly Lipschitz α on $[v-h, v+h]$, Theorem 6.3 proves that there exists $A > 0$ such that

$$\forall (u,s) \in [v-h, v+h] \times \mathbb{R}^+, \quad |Wf_0(u,s)| \leq A s^{\alpha+1/2}. \tag{6.44}$$

Inserting this in (6.43) gives

$$|Wf(u,s)| \leq A \left(\frac{s}{s_0}\right)^{n+1/2} s_0^{\alpha+1/2}, \tag{6.45}$$

from which we derive (6.40) by inserting the expression (6.42) of s_0. ■

This theorem explains how the wavelet transform decay relates to the amount of diffusion of a singularity. At large scales $s \gg \sigma/\beta$, the Gaussian averaging is not "felt" by the wavelet transform that decays like $s^{\alpha+1/2}$. For $s \leqslant \sigma/\beta$, the variation of f at v is not sharp relative to s because of the Gaussian averaging. At these fine scales, the wavelet transform decays like $s^{n+1/2}$ because f is $\mathbf{C}^\infty$.

The parameters K, α, and σ are numerically estimated from the decay of the modulus maxima along the maxima curves that converge toward v. The variance β^2 depends on the choice of wavelet and is known in advance. A regression is performed to approximate

$$\log_2 |Wf(u,s)| \approx \log_2(K) + \left(\alpha + \frac{1}{2}\right)\log_2 s - \frac{n-\alpha}{2}\log_2\left(1 + \frac{\sigma^2}{\beta^2 s^2}\right).$$

Figure 6.6 gives the wavelet modulus maxima computed with a wavelet that is a second derivative of a Gaussian. The decay of $\log_2 |Wf(u,s)|$ as a function of $\log_2 s$ is given along several maxima lines corresponding to smoothed and nonsmoothed singularities. The wavelet is normalized so that $\beta = 1$ and the diffusion scale is $\sigma = 2^{-5}$.

6.2.2 Dyadic Maxima Representation

Wavelet transform maxima carry the properties of sharp signal transitions and singularities. By recovering a signal approximation from these maxima, signal singularities can be modified or removed by processing the wavelet modulus maxima.

For fast numerical computations, the detection of wavelet transform maxima is limited to dyadic scales $\{2^j\}_{j\in\mathbb{Z}}$. Suppose that ψ is a dyadic wavelet, which means that there exist $A > 0$ and B such that

$$\forall \omega \in \mathbb{R} - \{0\}, \quad A \leqslant \sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j\omega)|^2 \leqslant B. \tag{6.46}$$

As a consequence of Theorem 5.11 on translation-invariant frames, it is proved in Section 5.2 that the resulting translation-invariant dyadic wavelet transform $\{Wf(u,2^j)\}_{j\in\mathbb{Z}}$ is complete and stable. This dyadic wavelet transform has the same properties as a continuous wavelet transform $Wf(u,s)$. All theorems of Sections 6.1.3 and 6.2 remain valid if we restrict s to the dyadic scales $\{2^j\}_{j\in\mathbb{Z}}$. Singularities create sequences of maxima that converge toward the corresponding location at fine scales, and the Lipschitz regularity is calculated from the decay of the maxima amplitude.

Scale–Space Maxima Support

Mallat and Zhong [367] introduced a dyadic wavelet maxima representation with a scale–space approximation support Λ of modulus maxima $(u, 2^j)$ of Wf.

Wavelet maxima can be interpreted as points of 0 or π phase for an appropriate complex wavelet transform. Let ψ' be the derivative of ψ and $\psi'_{u,2^j}(t) = 2^{-j/2}\psi'(2^{-j}(t-u))$. If Wf has a local extremum at u_0, then

$$\frac{\partial Wf(u_0, 2^j)}{\partial u} = -2^{-j}\,\langle f, \psi'_{2^j,u_0}\rangle = 0.$$

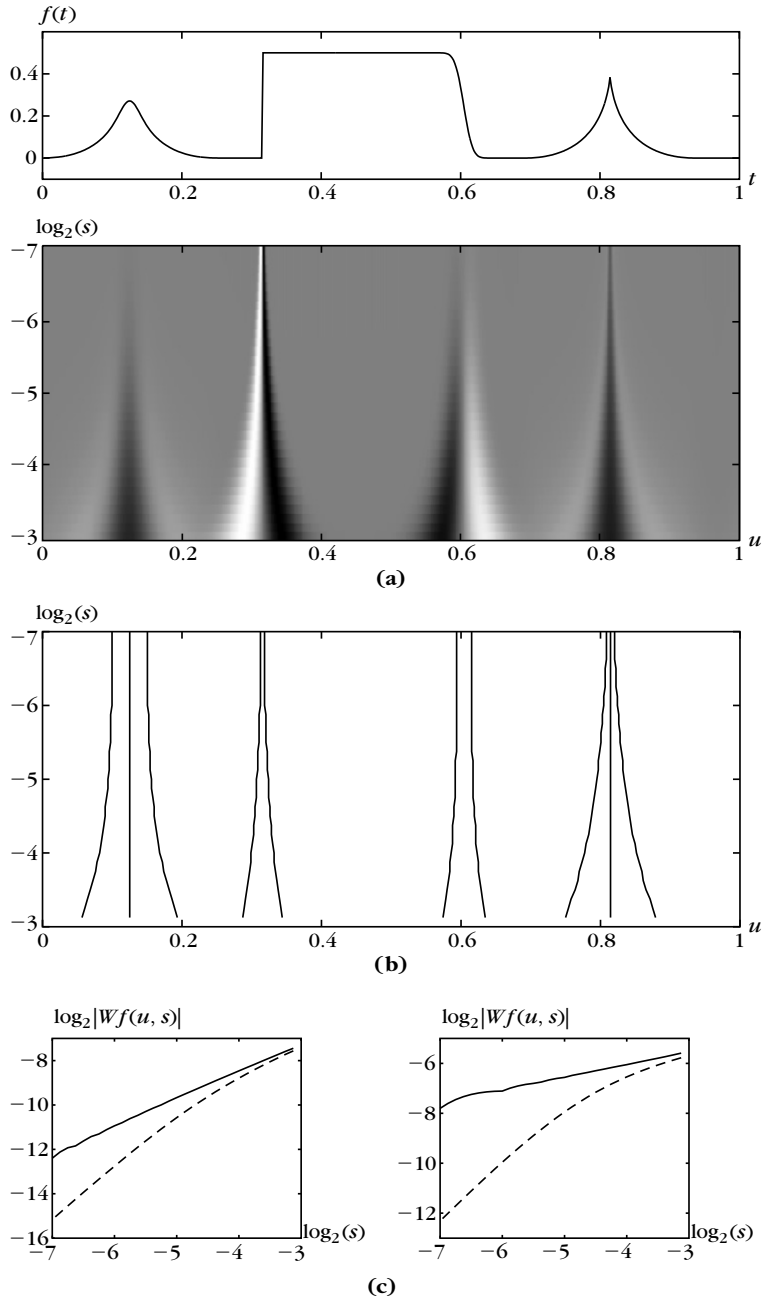

FIGURE 6.6

(a) Wavelet transform $Wf(u,s)$. **(b)** Modulus maxima of a wavelet transform computed $\psi = \theta''$, where θ is a Gaussian with variance $\beta = 1$. **(c)** Decay of $\log_2 |Wf(u,s)|$ along maxima curves. The solid and dotted lines (*left*) correspond to the maxima curves converging to $t = 0.81$ and $t = 0.12$, respectively. They correspond to the curves (*right*) converging to $t = 0.38$ and $t = 0.55$, respectively. The diffusion at $t = 0.12$ and $t = 0.55$ modifies the decay for $s \leqslant \sigma = 2^{-5}$.

Let us introduce a complex wavelet $\psi^c(t) = \psi(t) + i\psi'(t)$. If $(u, s) \in \Lambda$, then the resulting complex wavelet transform value is

$$W^c f(u, 2^j) = \langle f, \psi^c_{2^j,u} \rangle = \langle f, \psi_{2^j,u} \rangle + i \langle f, \psi'_{2^j,u} \rangle = W f(u, s), \tag{6.47}$$

because $\langle f, \psi'_{2^j,u} \rangle = 0$. The complex wavelet value $W^c f(u, s)$ has a phase equal to 0 or π depending on the sign of $W f(u, s)$, and a modulus $|W^c f(u, s)| = |W f(u, s)|$.

Figure 6.7(c) gives an example computed with the quadratic spline dyadic wavelet in Figure 5.3. This adaptive sampling of u produces a translation-invariant representation, which is important for pattern recognition. When f is translated by τ each $W f(2^j, u)$ is translated by τ, so the maxima support is translated by τ, as illustrated by Figure 6.8. This is not the case for wavelet frame coefficients, where the translation parameter u is sampled with an interval proportional to the scale a^j, as explained in Section 5.3.

Approximations from Wavelet Maxima

Mallat and Zhong [367] recover signal approximations from their wavelet maxima with an alternate projection algorithm, but several other algorithms have been proposed [150, 190, 286]. In the following we concentrate on orthogonal projection approximations on the space generated by wavelets in the scale–space maxima support. Numerical experiments show that dyadic wavelets of compact support recover signal approximations with a relative mean-square error that is typically of the order of 10^{-2}.

For general dyadic wavelets, Meyer [45] and Berman and Baras [107] proved that exact reconstruction is not possible. They found families of continuous or discrete signals having the same dyadic wavelet transforms and modulus maxima. However, signals with the same wavelet maxima differ from each other by small amplitude errors introducing no oscillation, which explains the success of numerical reconstructions [367]. If the signal has a band-limited Fourier transform and if $\hat{\psi}$ has a compact support, then Kicey and Lennard [328] proved that wavelet modulus maxima define a complete and stable signal representation.

As a result of (6.47), the wavelet modulus maxima specifies the complex wavelet inner products $\{\langle f, \psi^c_{u,2^j} \rangle\}_{(u,2^j)\in\Lambda}$. Thus, a modulus maxima approximation can be computed as an orthogonal projection of f on the space generated by the complex wavelets $\{\psi^c_{u,2^j}\}_{(u,2^j)\in\Lambda}$. To reduce computations, the explicit extrema condition $\langle \tilde{f}, \psi'_{u,2^j} \rangle = 0$ is often removed, because it is indirectly almost obtained by calculating the orthogonal projection over the space $\mathbf{V}_\Lambda$ generated by the real maxima wavelets $\{\psi_{u,2^j}\}_{(u,2^j)\in\Lambda}$. Section 5.1.3 shows that this orthogonal projection is obtained from the dual frame $\{\tilde{\psi}_{u,2^j}\}_{(u,2^j)\in\Lambda}$ of $\{\psi_{u,2^j}\}_{(u,2^j)\in\Lambda}$ in $\mathbf{V}_\Lambda$:

$$f_\Lambda = \sum_{(u,2^j)\in\Lambda} \langle f, \psi_{u,2^j} \rangle \, \tilde{\psi}_{u,2^j}. \tag{6.48}$$

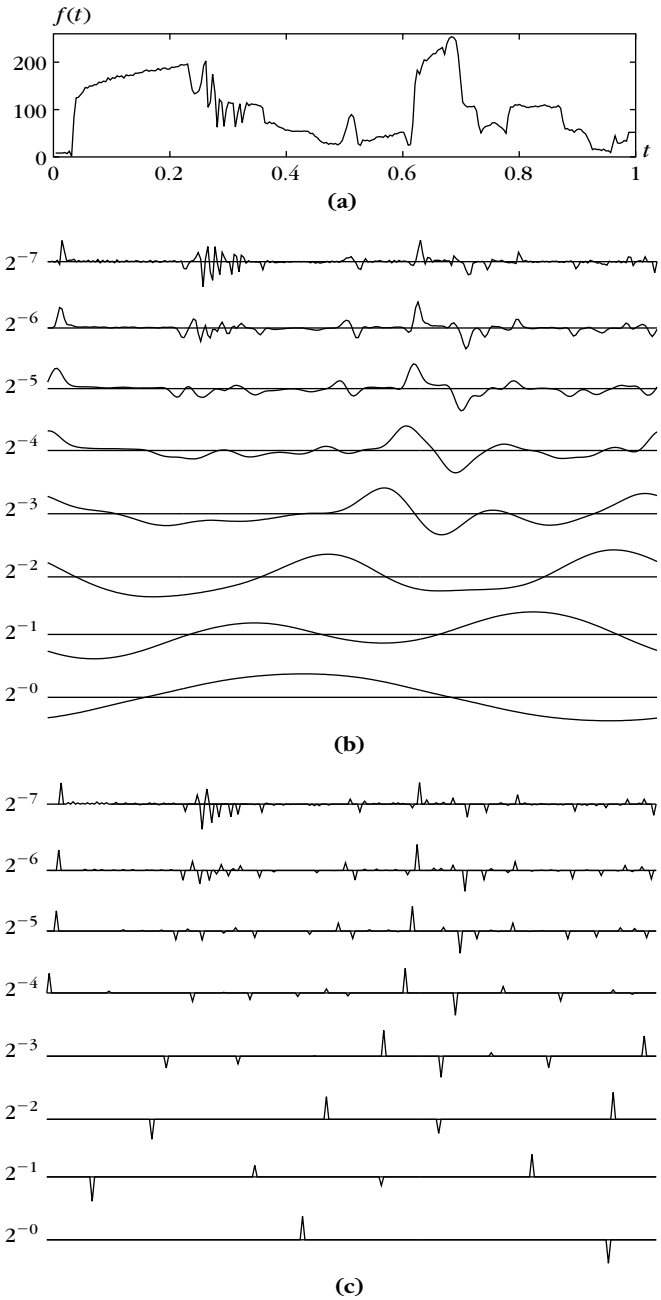

FIGURE 6.7

(a) Intensity variation along one row of the Lena image. **(b)** Dyadic wavelet transform computed at all scales $2N^{-1} \leq 2^j \leq 1$, with the quadratic spline wavelet $\psi = -\theta'$ shown in Figure 5.3. **(c)** Modulus maxima of the dyadic wavelet transform.

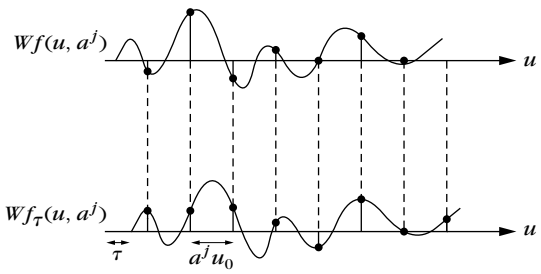

FIGURE 6.8

If $f_\tau(t) = f(t-\tau)$, then $Wf_\tau(u, a^j) = Wf(u-\tau, a^j)$. Uniformly sampling $Wf_\tau(u, a^j)$ and $Wf(u, a^j)$ at $u = na^j u_0$ may yield very different values if $\tau \neq ku_0 a^j$.

The dual-synthesis algorithm from Section 5.1.3 computes this orthogonal projection by inverting a symmetric operator L in $\mathbf{V}_\Lambda$:

$$Ly = \sum_{(u,2^j)\in\Lambda} \langle y, \psi_{u,2^j} \rangle \, \psi_{u,2^j}, \tag{6.49}$$

with a conjugate gradient algorithm Indeed $f_\Lambda = L^{-1}(Lf)$.

EXAMPLE 6.1

Figure 6.9(b) shows the approximation f_Λ, recovered with 10 conjugate gradient iterations, from the wavelet maxima in Figure 6.7(c). This reconstruction is calculated with real instead of complex wavelets. After 20 iterations, the reconstruction error is $\|f - \tilde{f}\|/\|f\| = 2.5\,10^{-2}$. Figure 6.9(c) shows the signal reconstructed from 50% of the wavelet maxima that have the largest amplitude. Sharp signal transitions corresponding to large wavelet maxima have not been affected, but small texture variations disappear because the corresponding maxima are removed. The resulting signal is piecewise regular.

Fast Discrete Calculations

The conjugate-gradient inversion of the operator (6.49) iterates on this operator many times. If there are many local maxima, it is more efficient to compute $Wy(u, 2^j) = \langle y, \psi_{u,2^j} \rangle$ for all $(u, 2^j)$, with the "algorithme à trous" (see Section 5.2.2). For a signal of size N, it cascades convolutions with two filters $h[n]$ and $g[n]$, up to a maximum scale $J = \log_2 N$, with $O(N \log_2 N)$ operations. All nonmaxima coefficients for $(u, 2^j) \notin \Lambda$ are then set to zero. The reconstruction of Ly is computed by modifying the filter bank reconstruction given by Theorem 5.14, which also requires $O(N \log_2 N)$ operations. The decomposition and reconstruction wavelets are the same in (6.49), so the reconstruction filters are $\tilde{h}[n] = h[n]$ and $\tilde{g}[n] = g[n]$. The factor $1/2$ in (5.72) is also removed because the reconstruction

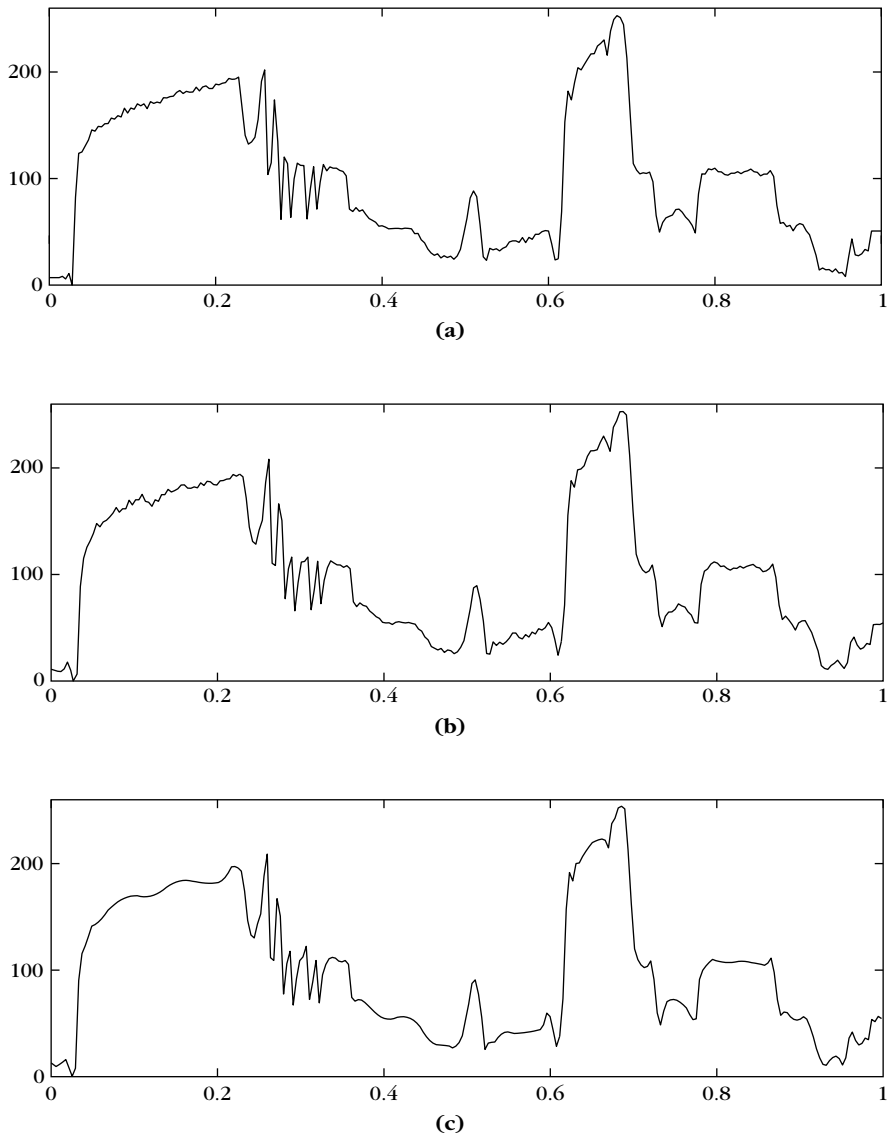

FIGURE 6.9

(a) Original signal f. **(b)** Signal approximation f_Λ recovered from the dyadic wavelet maxima shown in Figure 6.7(c). **(c)** Approximation recovered from 50% largest maxima.

wavelets in (6.49) are not attenuated by 2^{-j} as in a nonsampled wavelet reconstruction (5.50). For $J = \log_2 N$, we initialize $\tilde{a}_J[n] = C/\sqrt{N}$ where C is the average signal value, and for $\log_2 N > j \geq 0$ we compute

$$\tilde{a}_j[n] = \tilde{a}_{j+1} \star h_j[n] + \tilde{d}_{j+1} \star g_j[n]. \tag{6.50}$$

One can verify that $Ly[n] = \tilde{a}_0[n]$ with the same derivations as in the proof of Theorem 5.14.

The signal approximations shown in Figure 6.9 are computed with the filters of Table 5.1. About 10 iterations of conjugate gradient are usually sufficient to recover an approximation with $\|f_\Lambda - f\|/\|f\|$ of the order of 10^{-2}, if all wavelet maxima are kept.

6.3 MULTISCALE EDGE DETECTION

Image edges are often important for pattern recognition. This is clearly illustrated by our visual ability to recognize an object from a drawing that gives a rough outline of contours. But, what is an edge? It could be defined as points where the image intensity has sharp transitions. A closer look shows that this definition is often not satisfactory. Image textures do have sharp intensity variations that are often not considered as edges. When looking at a brick wall, we may decide that the edges are the contours of the wall whereas the bricks define a texture. Alternatively, we may include the contours of each brick in the set of edges and consider the irregular surface of each brick as a texture. The discrimination of edges versus textures depends on the scale of analysis.

This has motivated computer vision researchers to detect sharp image variations at different scales [42, 416]. Section 6.3.1 describes the multiscale Canny edge detector [146]. It is equivalent to detecting modulus maxima in a two-dimensional dyadic wavelet transform [367]. Thus, the scale–space support of these modulus maxima correspond to multiscale edges. The Lipschitz regularity of edge points is derived from the decay of wavelet modulus maxima across scales. Image approximations are recovered with an orthogonal projection on the wavelets of the modulus maxima support with no visual degradation. Thus, image-processing algorithms can be implemented on multiscale edges.

6.3.1 Wavelet Maxima for Images

Canny Edge Detection

The Canny algorithm detects points of sharp variation in an image $f(x_1, x_2)$ by calculating the modulus of its gradient vector

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x_1} , \frac{\partial f}{\partial x_2} \right). \tag{6.51}$$

The partial derivative of f in the direction of a unit vector $\vec{n} = (\cos\alpha, \sin\alpha)$ in the $x = (x_1, x_2)$ plane is calculated as an inner product with the gradient vector

$$\frac{\partial f}{\partial \vec{n}} = \vec{\nabla} f \,.\, \vec{n} = \frac{\partial f}{\partial x_1} \cos\alpha + \frac{\partial f}{\partial x_2} \sin\alpha.$$

The absolute value of this partial derivative is maximum if $\vec{n}$ is colinear to $\vec{\nabla} f$. This shows that $\vec{\nabla} f(x)$ is parallel to the direction of maximum change of the surface

$f(x)$. A point $y \in \mathbb{R}^2$ is defined as an edge if $|\vec{\nabla} f(x)|$ is locally maximum at $x = y$ when $x = y + \lambda\, \vec{\nabla} f(y)$ and $|\lambda|$ is small enough. This means that the partial derivatives of f reach a local maximum at $x = y$, when x varies in a one-dimensional neighborhood of y along the direction of maximum change of f at y. These edge points are inflection points of f.

Multiscale Edge Detection

A multiscale version of this edge detector is implemented by smoothing the surface with a convolution kernel $\theta(x)$ that is dilated. This is computed with two wavelets that are the partial derivatives of θ:

$$\psi^1 = -\frac{\partial \theta}{\partial x_1} \quad \text{and} \quad \psi^2 = -\frac{\partial \theta}{\partial x_2}. \tag{6.52}$$

The scale varies along the dyadic sequence $\{2^j\}_{j \in \mathbb{Z}}$ to limit computations and storage. For $1 \leqslant k \leqslant 2$, we denote for $x = (x_1, x_2)$,

$$\psi^k_{2^j}(x_1, x_2) = \frac{1}{2^j}\, \psi^k\left(\frac{x_1}{2^j}, \frac{x_2}{2^j}\right) \quad \text{and} \quad \bar{\psi}^k_{2^j}(x) = \psi^k_{2^j}(-x).$$

In the two directions indexed by $1 \leqslant k \leqslant 2$, the dyadic wavelet transform of $f \in \mathbf{L^2}(\mathbb{R}^2)$ at $u = (u_1, u_2)$ is

$$W^k f(u, 2^j) = \langle f(x), \psi^k_{2^j}(x - u) \rangle = f \star \bar{\psi}^k_{2^j}(u)\ . \tag{6.53}$$

Section 5.5 gives necessary and sufficient conditions for obtaining a complete and stable representation.

Let us denote $\theta_{2^j}(x) = 2^{-j}\, \theta(2^{-j} x)$ and $\bar{\theta}_{2^j}(x) = \theta_{2^j}(-x)$. The two scaled wavelets can be rewritten as

$$\bar{\psi}^1_{2^j} = 2^j\, \frac{\partial \bar{\theta}_{2^j}}{\partial x_1} \quad \text{and} \quad \bar{\psi}^2_{2^j} = 2^j\, \frac{\partial \bar{\theta}_{2^j}}{\partial x_2}.$$

Thus, let us derive from (6.53) that the wavelet transform components are proportional to the coordinates of the gradient vector of f smoothed by $\bar{\theta}_{2^j}$:

$$\begin{pmatrix} W^1 f(u, 2^j) \\ W^2 f(u, 2^j) \end{pmatrix} = 2^j \begin{pmatrix} \frac{\partial}{\partial u_1}(f \star \bar{\theta}_{2^j})(u) \\ \frac{\partial}{\partial u_2}(f \star \bar{\theta}_{2^j})(u) \end{pmatrix} = 2^j\, \vec{\nabla}(f \star \bar{\theta}_{2^j})(u). \tag{6.54}$$

The modulus of this gradient vector is proportional to the wavelet transform modulus

$$Mf(u, 2^j) = \sqrt{|W^1 f(u, 2^j)|^2 + |W^2 f(u, 2^j)|^2}. \tag{6.55}$$

Let $Af(u, 2^j)$ be the angle of the wavelet transform vector (6.54) in the plane (x_1, x_2):

$$Af(u, 2^j) = \begin{cases} \alpha(u) & \text{if } W^1 f(u, 2^j) \geqslant 0 \\ \pi + \alpha(u) & \text{if } W^1 f(u, 2^j) < 0 \end{cases} \tag{6.56}$$

with

$$\alpha(u) = \tan^{-1}\left(\frac{W^2 f(u, 2^j)}{W^1 f(u, 2^j)}\right).$$

The unit vector $\vec{n}_j(u) = (\cos Af(u, 2^j), \sin Af(u, 2^j))$ is colinear to $\vec{\nabla}(f \star \bar{\theta}_{2^j})(u)$. An edge point at the scale 2^j is a point v such that $Mf(u, 2^j)$ is locally maximum at $u = v$ when $u = v + \lambda \vec{n}_j(v)$ and $|\lambda|$ is small enough. These points are also called wavelet transform *modulus maxima*. The smoothed image $f \star \bar{\theta}_{2^j}$ has an inflection point at a modulus maximum location. Figure 6.10 gives an example where the wavelet modulus maxima are located along the contour of a circle.

Maxima Curves

Edge points are distributed along curves that often correspond to the boundary of important structures. Individual wavelet modulus maxima are chained together to form a maxima curve that follows an edge. At any location, the tangent of the edge curve is approximated by computing the tangent of a level set. This tangent direction is used to chain wavelet maxima that are along the same edge curve.

The level sets of $g(x)$ are the curves $x(s)$ in the (x_1, x_2) plane where $g(x(s))$ is constant. The parameter s is the arc-length of the level set. Let $\vec{\tau} = (\tau_1, \tau_2)$ be the direction of the tangent of $x(s)$. Since $g(x(s))$ is constant when s varies,

$$\frac{\partial g(x(s))}{\partial s} = \frac{\partial g}{\partial x_1}\tau_1 + \frac{\partial g}{\partial x_2}\tau_2 = \vec{\nabla} g \cdot \vec{\tau} = 0.$$

So, $\vec{\nabla} g(x)$ is perpendicular to the direction $\vec{\tau}$ of the tangent of the level set that goes through x.

This level set property applied to $g = f \star \bar{\theta}_{2^j}$ proves that at a maximum point v the vector $\vec{n}_j(v)$ of angle $Af(v, 2^j)$ is perpendicular to the level set of $f \star \bar{\theta}_{2^j}$ going through v. If the intensity profile remains constant along an edge, then the inflection points (maxima points) are along a level set. The tangent of the maxima curve is therefore perpendicular to $\vec{n}_j(v)$. The intensity profile of an edge may not be constant but its variations are often negligible over a neighborhood of size 2^j for a sufficiently small scale 2^j, unless we are near a corner. The tangent of the maxima curve is then nearly perpendicular to $\vec{n}_j(v)$. In discrete calculations, maxima curves are recovered by chaining together any two wavelet maxima at v and $v + \vec{n}$, which are neighbors over the image sampling grid and such that $\vec{n}$ is nearly perpendicular to $\vec{n}_j(v)$.

EXAMPLE 6.2

The dyadic wavelet transform of the image in Figure 6.10 yields modulus images $Mf(v, 2^j)$ with maxima along the boundary of a disk. This circular edge is also a level set of the image. Thus, the vector $\vec{n}_j(v)$ of angle $Af(v, 2^j)$ is perpendicular to the edge at the maxima locations.

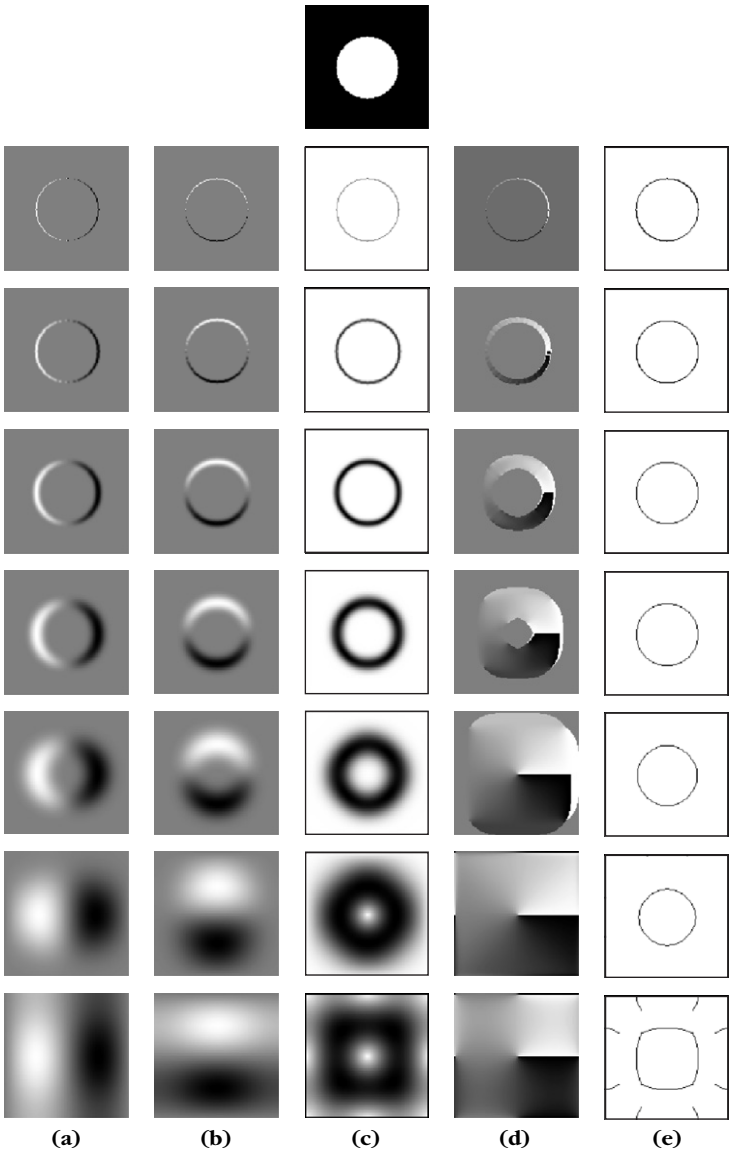

FIGURE 6.10

The very top image has $N = 128^2$ pixels. **(a)** Wavelet transform in the horizontal direction with a scale 2^j that increases from top to bottom: $\{W^1 f(u, 2^j)\}_{-6 \leqslant j \leqslant 0}$; black, gray, and white pixels correspond to negative, zero, and positive values, respectively. **(b)** Vertical direction: $\{W^2 f(u, 2^j)\}_{-6 \leqslant j \leqslant 0}$. **(c)** Wavelet transform modulus $\{Mf(u, 2^j)\}_{-6 \leqslant j \leqslant 0}$; white and black pixels correspond to zero and large-amplitude coefficients, respectively. **(d)** Angles $\{Af(u, 2^j)\}_{-6 \leqslant j \leqslant 0}$ at points where the modulus is nonzero. **(e)** The wavelet modulus maxima support is shown in black.

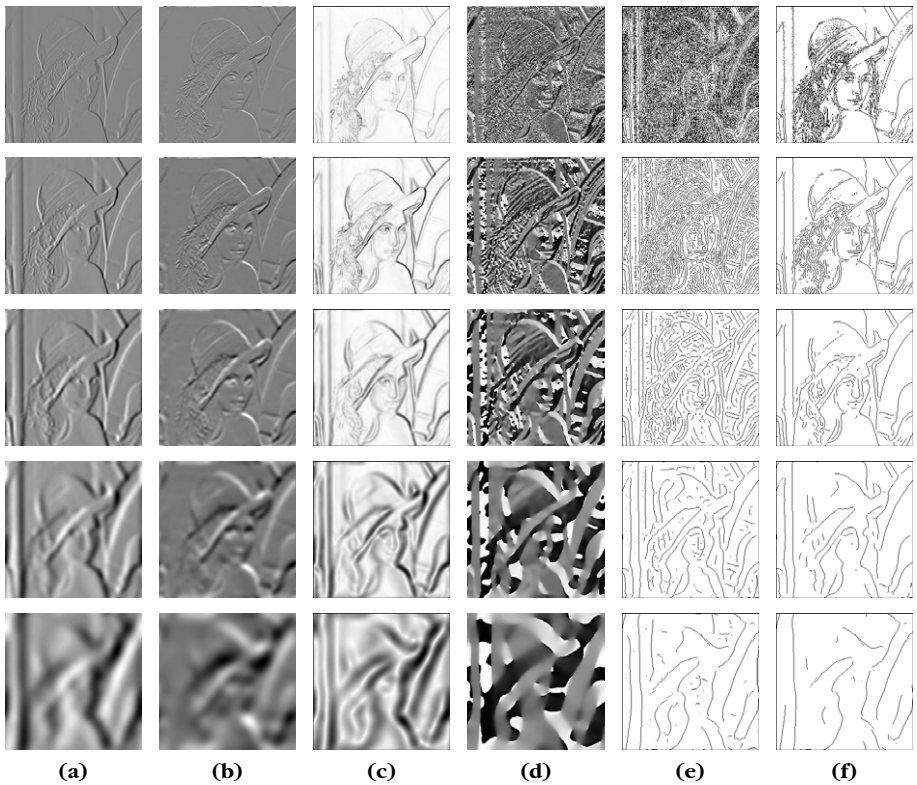

FIGURE 6.11

Multiscale edges of the Lena image shown in Figure 6.12. **(a)** $\{W^1 f(u, 2^j)\}_{-7 \leqslant j \leqslant -3}$. **(b)** $\{W^2 f(u, 2^j)\}_{-7 \leqslant j \leqslant -3}$. **(c)** $\{Mf(u, 2^j)\}_{-7 \leqslant j \leqslant -3}$. **(d)** $\{Af(u, 2^j)\}_{-7 \leqslant j \leqslant -3}$. **(e)** Modulus maxima support. **(f)** Support of maxima with modulus values above a threshold.

EXAMPLE 6.3

In the Lena image shown in Figure 6.11, some edges disappear when the scale increases. These correspond to fine-scale intensity variations that are removed by the averaging with $\bar{\theta}_{2^j}$ when 2^j is large. This averaging also modifies the position of the remaining edges. Figure 6.11(f) displays the wavelet maxima such that $Mf(v, 2^j) \geqslant T$ for a given threshold T. They indicate the location of edges where the image has large amplitude variations.

Lipschitz Regularity

The decay of the two-dimensional wavelet transform depends on the regularity of f. We restrict the analysis to Lipschitz exponents $0 \leqslant \alpha \leqslant 1$. A function f is said to

be Lipschitz α at $v = (v_1, v_2)$ if there exists $K > 0$ such that for all $(x_1, x_2) \in \mathbb{R}^2$,

$$|f(x_1, x_2) - f(v_1, v_2)| \leq K\,(|x_1 - v_1|^2 + |x_2 - v_2|^2)^{\alpha/2}. \tag{6.57}$$

If there exists $K > 0$ such that (6.57) is satisfied for any $v \in \Omega$, then f is uniformly Lipschitz α over Ω. As in one dimension, the Lipschitz regularity of a function f is related to the asymptotic decay $|W^1 f(u, 2^j)|$ and $|W^2 f(u, 2^j)|$ in the corresponding neighborhood. This decay is controlled by $Mf(u, 2^j)$. Like in Theorem 6.3, one can prove that f is uniformly Lipschitz α inside a bounded domain of $\mathbb{R}^2$ if and only if there exists $A > 0$ such that for all u inside this domain and all scales 2^j,

$$|Mf(u, 2^j)| \leq A\, 2^{j(\alpha+1)}. \tag{6.58}$$

Suppose that the image has an isolated edge curve along which f has Lipschitz regularity α. The value of $|Mf(u, 2^j)|$ in a two-dimensional neighborhood of the edge curve can be bounded by the wavelet modulus values along the edge curve. The Lipschitz regularity α of the edge is estimated with (6.58) by measuring the slope of $\log_2 |Mf(u, 2^j)|$ as a function of j. If f is not singular but has a smooth transition along the edge, the smoothness can be quantified by the variance σ^2 of a two-dimensional Gaussian blur. The value of σ^2 is estimated by generalizing Theorem 6.7.

Reconstruction from Edges

In his book about vision, Marr [42] conjectured that images can be reconstructed from multiscale edges. For a Canny edge detector, this is equivalent to recovering images from wavelet modulus maxima. Whether dyadic wavelet maxima define a complete and stable representation in two dimensions is still an open mathematical problem. However, the algorithm of Mallat and Zhong [367] recovers an image approximation that is visually identical to the original one. In the following, image approximations are computed by projecting the image on the space generated by wavelets on the modulus maxima support.

Let Λ be the set of all modulus maxima points $(u, 2^j)$. Let $\vec{n}$ be the unit vector in the direction of $Af(u, 2^j)$ and

$$\psi^3_{u,2^j}(x) = 2^{2j}\, \frac{\partial^2 \theta_{2^j}(x - u)}{\partial \vec{n}^2}.$$

Since the wavelet transform modulus $Mf(u, 2^j)$ has a local extremum at u in the direction of $\vec{n}$, it results that

$$\langle f, \psi^3_{u,2^j} \rangle = 0. \tag{6.59}$$

A modulus maxima representation provides the set of inner products $\{\langle f, \psi^k_{u,2^j} \rangle\}_{(u,2^j) \in \Lambda, 1 \leq k \leq 3}$. A modulus maxima approximation f_Λ can be computed as an orthogonal projection of f on the space generated by the family of maxima wavelets $\{\psi^k_{u,2^j}\}_{(u,2^j) \in \Lambda, 1 \leq k \leq 3}$.

To reduce computations, the condition on the third wavelets $\langle f, \psi^3_{u,2^j} \rangle = 0$ is removed, because it is indirectly almost imposed by the orthogonal projection over the space $\mathbf{V}_\Lambda$ generated by the other two wavelets for $k = 1, 2$. The dual-synthesis algorithm from Section 5.1.3 computes this orthogonal projection f_Λ by inverting a symmetric operator L in $\mathbf{V}_\Lambda$:

$$Ly = \sum_{(u,2^j)\in\Lambda} \sum_{k=1}^{2} \langle y, \psi^k_{u,2^j} \rangle \, \psi^k_{u,2^j}, \tag{6.60}$$

with a conjugate-gradient algorithm. Indeed $f_\Lambda = L^{-1}(Lf)$. When keeping all modulus maxima, the resulting modulus maxima approximation f_Λ satisfies $\|f_\Lambda - f\|/\|f\| \leqslant 10^{-2}$. Singularities and edges are nearly perfectly recovered and no spurious oscillations are introduced. The images differ slightly in smooth regions, but visually this is not noticeable.

EXAMPLE 6.4

The image reconstructed in Figure 6.12(b) is visually identical to the original image. It is recovered with 10 conjugate-gradient iterations. After 20 iterations, the relative mean-square reconstruction error is $\|\tilde{f} - f\|/\|f\| = 4\,10^{-3}$. The thresholding of maxima accounts for the disappearance of image structures from the reconstruction shown in Figure 6.12(c). Sharp image variations are recovered.

Denoising by Multiscale Edge Thresholding

Multiscale edge representations can be used to reduce additive noise. Denoising algorithms by thresholding wavelet coefficients are presented in Section 11.3.1. Block thresholding (see Section 11.4.2) regularizes this coefficient selection by regrouping them in square blocks. Similarly, a noise-removal algorithm can be implemented by thresholding multiscale wavelet maxima, while taking into account their geometric properties.

A simple approach implemented by Hwang and Mallat [364] chains the maxima into curves that are thresholded as a block. In Figure 6.13 noisy modulus maxima are shown on the second row and the third row displays the thresholded modulus maxima chains. At the finest scale shown on the left, the noise is masking the image structures. Maxima chains are selected by using the position of the selected maxima at the previous scale. An image approximation is recovered from the selected wavelet maxima. Edges are well-recovered visually but textures and fine structures are removed. This produces a cartoonlike image.

Illusory Contours

A multiscale wavelet edge detector defines edges as points where the image intensity varies sharply. However, this definition is too restrictive when edges are used to find the contours of objects. For image segmentation, edges must define closed curves

(a) (b)

(c)

FIGURE 6.12

(a) Original Lena image. **(b)** Image reconstructed from the wavelet maxima displayed in Figure 6.11(e) and larger-scale maxima. **(c)** Image reconstructed from the thresholded wavelet maxima displayed in Figure 6.11(f) and larger-scale maxima.

that outline the boundaries of each region. Because of noise or light variations, local edge detectors produce contours with holes. Filling these holes requires some prior knowledge about the behavior of edges in the image. The illusion of the Kanizsa triangle [37] shows that such an edge filling is performed by the human visual system.

In Figure 6.14 one can "see" the edges of a straight and a curved triangle although the image gray level remains uniformly white between the black discs. Closing edge curves and understanding illusory contours requires computational models that are

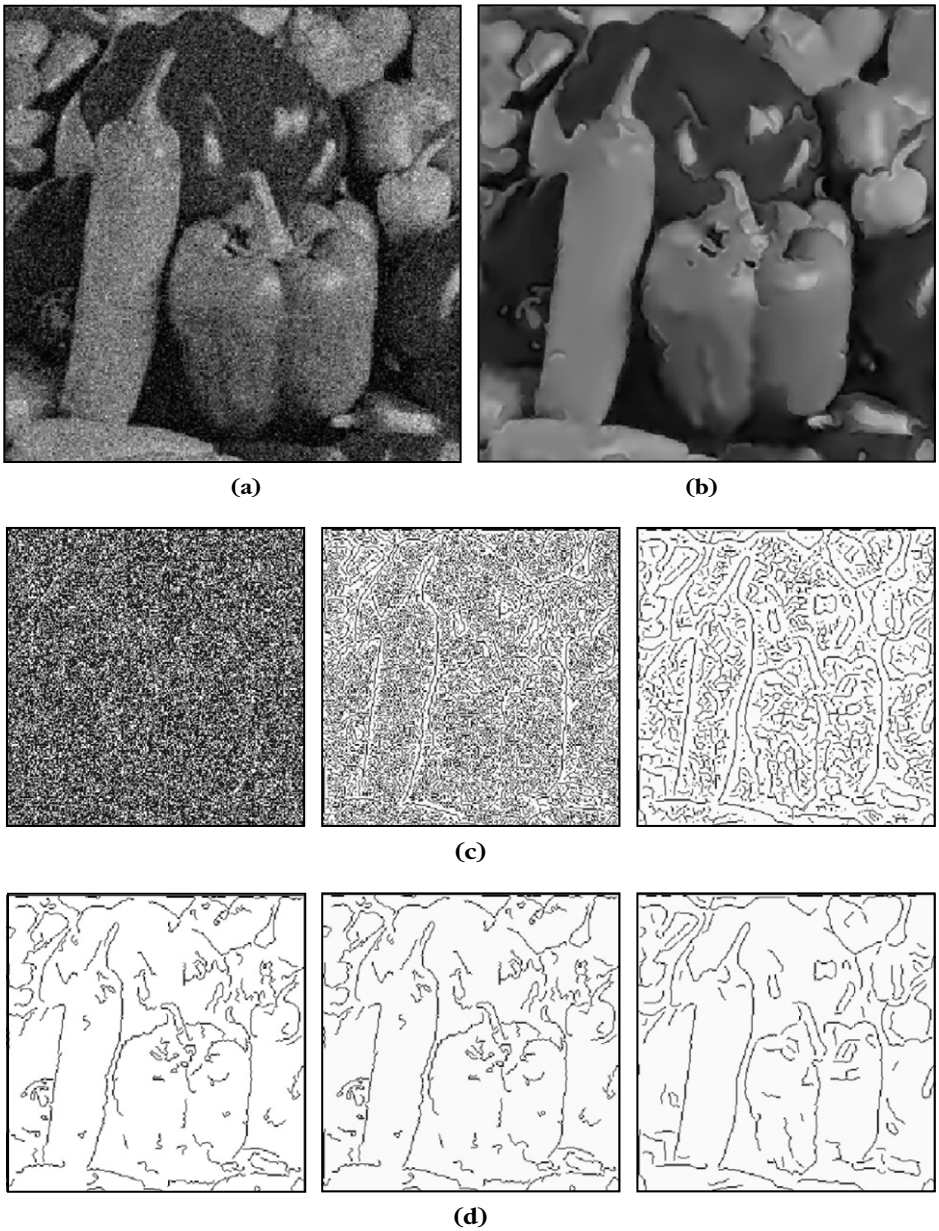

FIGURE 6.13

(a) Noisy peppers image. **(b)** Peppers image restored from the thresholding maxima chains shown in **(d)**. The images in row **(c)** show the wavelet maxima support of the noisy image—the scale increases from left to right, from 2^{-7} to 2^{-5}. The images in row (d) give the maxima support computed with a thresholding selection of the maxima chains.

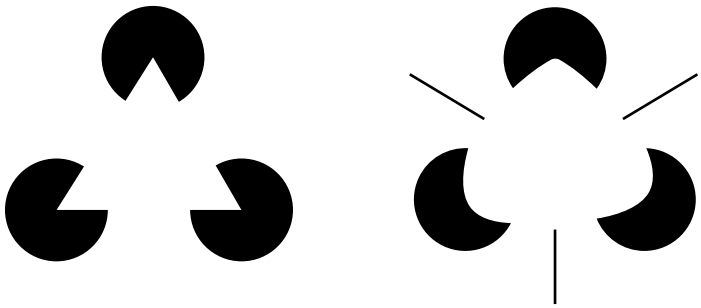

FIGURE 6.14

The illusory edges of a **(a)** straight and **(b)** curved triangle are perceived in domains where the images are uniformly white.

not as local as multiscale differential operators. Such contours can be obtained as the solution of a global optimization that incorporates constraints on the regularity of contours and takes into account the existence of occlusions [269].

6.3.2 Fast Multiscale Edge Computations

The dyadic wavelet transform of an image of N pixels is computed with a separable extension of the filter bank algorithm described in Section 5.2.2. A fast multiscale edge detection is derived [367].

Wavelet Design

Edge-detection wavelets (6.52) are designed as separable products of the one-dimensional dyadic wavelets constructed in Section 5.2.1. Their Fourier transform is

$$\hat{\psi}^1(\omega_1, \omega_2) = \hat{g}\left(\frac{\omega_1}{2}\right)\hat{\phi}\left(\frac{\omega_1}{2}\right)\hat{\phi}\left(\frac{\omega_2}{2}\right), \tag{6.61}$$

and

$$\hat{\psi}^2(\omega_1, \omega_2) = \hat{g}\left(\frac{\omega_2}{2}\right)\hat{\phi}\left(\frac{\omega_1}{2}\right)\hat{\phi}\left(\frac{\omega_2}{2}\right), \tag{6.62}$$

where $\hat{\phi}(\omega)$ is a scaling function that has energy concentrated at low frequencies and

$$\hat{g}(\omega) = -i\sqrt{2}\,\sin\left(\frac{\omega}{2}\right)\exp\left(\frac{-i\omega}{2}\right). \tag{6.63}$$

This transfer function is the Fourier transform of a finite difference filter, which is a discrete approximation of a derivative

$$\frac{g[p]}{\sqrt{2}} = \begin{cases} -0.5 & \text{if } p = 0 \\ 0.5 & \text{if } p = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{6.64}$$

The resulting wavelets ψ^1 and ψ^2 are finite difference approximations of partial derivatives along x_1 and x_2 of $\theta(x_1, x_2) = 4\,\phi(2x_1)\phi(2x_2)$.

To implement the dyadic wavelet transform with a filter bank algorithm, the scaling function $\hat{\phi}$ is calculated, as in (5.60), with an infinite product:

$$\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} \frac{\hat{h}(2^{-p}\omega)}{\sqrt{2}} = \frac{1}{\sqrt{2}}\,\hat{h}\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right). \tag{6.65}$$

The 2π periodic function $\hat{h}$ is the transfer function of a finite impulse-response low-pass filter $h[p]$. We showed in (5.61) that the Fourier transform of a box spline of degree m,

$$\hat{\phi}(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^{m+1} \exp\left(\frac{-i\varepsilon\omega}{2}\right) \quad \text{with } \varepsilon = \begin{cases} 1 & \text{if } m \text{ is even} \\ 0 & \text{if } m \text{ is odd} \end{cases}$$

is obtained with

$$\hat{h}(\omega) = \sqrt{2}\,\frac{\hat{\phi}(2\omega)}{\hat{\phi}(\omega)} = \sqrt{2}\left(\cos\frac{\omega}{2}\right)^{m+1} \exp\left(\frac{-i\varepsilon\omega}{2}\right).$$

Table 5.1 gives $h[p]$ for $m = 2$.

Algorithme à Trous

The one-dimensional *algorithme à trous* (see Section 5.2.2) is extended in two dimensions with convolutions along the image rows and columns.

Each sample $a_0[n]$ of the normalized discrete image is considered to be an average of the input analog image f calculated with the kernel $\phi(x_1)\,\phi(x_2)$ translated at $n = (n_1, n_2)$:

$$a_0[n_1, n_2] = \langle f(x_1, x_2),\ \phi(x_1 - n_1)\,\phi(x_2 - n_2)\rangle.$$

This is further justified in Section 7.7.3. For any $j \geqslant 0$, we denote

$$a_j[n_1, n_2] = \langle f(x_1, x_2),\ \phi_{2^j}(x_1 - n_1)\,\phi_{2^j}(x_2 - n_2)\rangle.$$

The discrete wavelet coefficients at $n = (n_1, n_2)$ are

$$d_j^1[n] = W^1 f(n, 2^j) \quad \text{and} \quad d_j^2[n] = W^2 f(n, 2^j).$$

They are calculated with separable convolutions.

For any $j \geqslant 0$, the filter $h[p]$ "dilated" by 2^j is defined by

$$\bar{h}_j[p] = \begin{cases} h[-p/2^j] & \text{if } p/2^j \in \mathbb{Z} \\ 0 & \text{otherwise;} \end{cases} \tag{6.66}$$

and for $j > 0$, a centered finite difference filter is defined by

$$\frac{\bar{g}_j[p]}{\sqrt{2}} = \begin{cases} 0.5 & \text{if } p = -2^{j-1} \\ -0.5 & \text{if } p = 2^{j-1} \\ 0 & \text{otherwise.} \end{cases} \tag{6.67}$$

For $j=0$, we define $\bar{g}_0[0]/\sqrt{2}=-0.5$, $\bar{g}_0[-1]/\sqrt{2}=-0.5$ and $\bar{g}_0[p]=0$ for $p\neq 0,-1$. A separable two-dimensional filter is written as

$$\alpha\beta[n_1,n_2]=\alpha[n_1]\,\beta[n_2],$$

and $\delta[n]$ is a discrete Dirac. Similar to Theorem 5.14, one can prove that for any $j\geqslant 0$ and any $n=(n_1,n_2)$,

$$a_{j+1}[n]=a_j\star\bar{h}_j\bar{h}_j[n], \tag{6.68}$$

$$d^1_{j+1}[n]=a_j\star\bar{g}_j\delta[n], \tag{6.69}$$

$$d^2_{j+1}[n]=a_j\star\delta\bar{g}_j[n]. \tag{6.70}$$

Dyadic wavelet coefficients up to the scale 2^J are therefore calculated by cascading the convolutions (6.68–6.70) for $0<j\leqslant J$. To take into account border problems, all convolutions are replaced by circular convolutions, which means that the input image $a_0[n]$ is considered to be periodic along its rows and columns. For an image of N pixels, this algorithm requires $O(N\log_2 N)$ operations. For a square image with a maximum scale $J=\log_2 N^{1/2}$, one can verify that the larger-scale approximation is a constant proportional to the gray-level average C:

$$a_J[n_1,n_2]=N^{-1/2}\sum_{n_1,n_2=0}^{N^{1/2}-1}a_0[n_1,n_2]=N^{1/2}\,C.$$

The wavelet transform modulus is $Mf(n,2^j)=|d^1_j[n]|^2+|d^2_j[n]|^2$, whereas $Af(n,2^j)$ is the angle of the vector $(d^1_j[n],d^2_j[n])$.

The support Λ of wavelet modulus maxima $(u,2^j)$ is the set of points $Mf(u,2^j)$, which is larger than its two neighbors $Mf(u\pm\vec{\varepsilon},2^j)$, where $\vec{\varepsilon}=(\varepsilon_1,\varepsilon_2)$ is the vector with coordinates ε_1 and ε_2 that are either 0 or 1 and have an angle that is the closest to $Af(u,2^j)$.

Reconstruction from Maxima

The orthogonal projection from wavelet maxima is computed with the dual-synthesis algorithm from Section 5.1.3, which inverts the symmetric operator (6.60) with conjugate-gradient iterations. This requires computing Ly efficiently for any image $y[n]$. For this purpose, the wavelet coefficients of y are first calculated with the *algorithme à trous*, and all coefficients for $(u,2^j)\notin\Lambda$ are set to 0. The signal $Ly[n]$ is recovered from these nonzero wavelet coefficients. Let $h_j[n]=\bar{h}_j[-n]$ and $g_j[n]=\bar{g}_j[-n]$ be the two filters defined with (6.66) and (6.67). The calculation is initialized for $J=\log_2 N^{1/2}$ by setting $\tilde{a}_J[n]=C\,N^{-1/2}$, where C is the average image intensity. For $\log_2 N>j\geqslant 0$, we compute

$$\tilde{a}_j[n]=\tilde{a}_{j+1}\star h_j h_j[n]+d^1_{j+1}\star g_j\delta[n]+d^2_{j+1}[n]\star\delta g_j[n],$$

and one can verify that $Ly[n]=\tilde{a}_0[n]$ is recovered with $O(N\log_2 N)$ operations. The reconstructed images that were shown in Figure 6.12 are obtained with 10 conjugate-gradient iterations implemented with this filter bank algorithm.

6.4 MULTIFRACTALS

Signals that are singular at almost every point were originally studied as pathological objects of pure mathematical interest. Mandelbrot [41] was the first to recognize that such phenomena are encountered everywhere. Among the many examples [25] are economic records such as the Dow Jones industrial average, physiological data including heart records, electromagnetic fluctuations in galactic radiation noise, textures in images of natural terrains, variations of traffic flow, and so on.

The singularities of multifractals often vary from point to point, and knowing the distribution of these singularities is important in analyzing their properties. Pointwise measurements of Lipschitz exponents are not possible because of the finite numerical resolution. After discretization, each sample corresponds to a time interval where the signal has an infinite number of singularities that may all be different. The singularity distribution must therefore be estimated from global measurements that take advantage of multifractal self-similarities. Section 6.4.2 computes the fractal dimension of sets of points having the same Lipschitz regularity, with a global partition function calculated from wavelet transform modulus maxima. Applications to fractal noises, such as fractional Brownian motions and hydrodynamic turbulence, are studied in Section 6.4.3.

6.4.1 Fractal Sets and Self-Similar Functions

A set $\mathcal{S} \subset \mathbb{R}^n$ is said to be self-similar if it is the union of disjoint subsets $\mathcal{S}_1, \ldots, \mathcal{S}_k$ that can be obtained from $\mathcal{S}$ with a scaling, translation, and rotation. This self-similarity often implies an infinite multiplication of details, which creates irregular structures. The triadic Cantor set and the Von Koch curve are simple examples.

EXAMPLE 6.5

The Von Koch curve is a fractal set obtained by recursively dividing each segment of length l in four segments of length $l/3$, as illustrated in Figure 6.15. Each subdivision multiplies the length by $4/3$; therefore, the limit of these subdivisions is a curve of infinite length.

EXAMPLE 6.6

The triadic Cantor set is constructed by recursively dividing intervals of size l in two subintervals of size $l/3$ and a central hole, illustrated in Figure 6.16. The iteration begins from $[0, 1]$. The Cantor set obtained as a limit of these subdivisions is a dust of points in $[0, 1]$.

Fractal Dimension

The Von Koch curve has infinite length in a finite square of $\mathbb{R}^2$; therefore, the usual length measurement is not well adapted to characterize the topological properties of such fractal curves. This motivated Hausdorff in 1919 to introduce a new definition

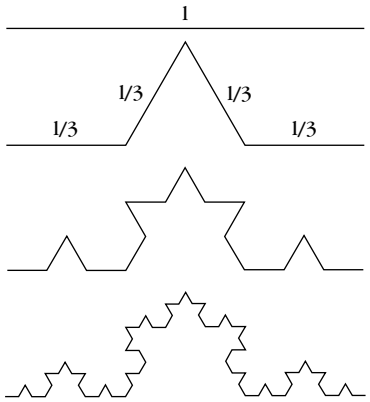

FIGURE 6.15

Three iterations of the Von Koch subdivision. The Von Koch curve is the fractal obtained as a limit of an infinite number of subdivisions.

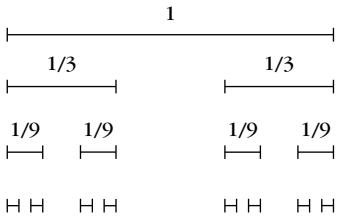

FIGURE 6.16

Three iterations of the Cantor subdivision of $[0, 1]$. The limit of an infinite number of subdivisions is a closed set in $[0, 1]$.

of dimension—the *capacity dimension*—based on the size variations of sets when measured at different scales.

The capacity dimension is a simplification of the Hausdorff dimension that is easier to compute numerically. Let $\mathcal{S}$ be a bounded set in $\mathbb{R}^n$. We count the minimum number $N(s)$ of balls of radius s needed to cover $\mathcal{S}$. If $\mathcal{S}$ is a set of dimension D with a finite length ($D = 1$), surface ($D = 2$), or volume ($D = 3$), then

$$N(s) \sim s^{-D},$$

so

$$D = -\lim_{s \to 0} \frac{\log N(s)}{\log s}. \tag{6.71}$$

The capacity dimension D of $\mathcal{S}$ generalizes this result and is defined by

$$D = -\liminf_{s \to 0} \frac{\log N(s)}{\log s}. \tag{6.72}$$

The measure of $\mathcal{S}$ is then

$$M = \limsup_{s \to 0} N(s)\, s^D.$$

It may be finite or infinite.

The Hausdorff dimension is a refined fractal measure that considers all covers of $\mathcal{S}$ with balls of radius smaller than s. It is most often, but not always, equal to the capacity dimension. In the following examples, the capacity dimension is called *fractal dimension*.

EXAMPLE 6.7

The Von Koch curve has infinite length because its fractal dimension is $D > 1$. We need $N(s) = 4^n$ balls of size $s = 3^{-n}$ to cover the whole curve, thus,

$$N(3^{-n}) = (3^{-n})^{-\log 4/\log 3}.$$

One can verify that at any other scale s, the minimum number of balls $N(s)$ to cover this curve satisfies

$$D = -\liminf_{s \to 0} \frac{\log N(s)}{\log s} = \frac{\log 4}{\log 3}.$$

As expected, it has a fractal dimension between 1 and 2.

EXAMPLE 6.8

The triadic Cantor set is covered by $N(s) = 2^n$ intervals of size $s = 3^{-n}$, so

$$N(3^{-n}) = (3^{-n})^{-\log 2/\log 3}.$$

One can also verify that

$$D = -\liminf_{s \to 0} \frac{\log N(s)}{\log s} = \frac{\log 2}{\log 3}.$$

Self-Similar Functions

Let f be a continuous function with a compact support $\mathcal{S}$. We say that f is *self-similar* if there exist disjoint subsets $\mathcal{S}_1, \ldots, \mathcal{S}_k$ such that the graph of f restricted to each $\mathcal{S}_i$ is an affine transformation of f. This means that there exist a scale $l_i > 1$, a translation r_i, a weight p_i, and a constant c_i such that

$$\forall t \in \mathcal{S}_i, \quad f(t) = c_i + p_i f\Big(l_i(t - r_i)\Big). \tag{6.73}$$

Outside these subsets, we suppose that f is constant. Generalizations of this definition can also be used [128].

If a function is self-similar then its wavelet transform is also self-similar. Let g be an affine transformation of f:

$$g(t) = p f\Big(l(t - r)\Big) + c. \tag{6.74}$$

Its wavelet transform is

$$Wg(u,s) = \int_{-\infty}^{+\infty} g(t)\,\frac{1}{\sqrt{s}}\,\psi\left(\frac{t-u}{s}\right)dt.$$

With the change of variable $t' = l(t-r)$, since ψ has a zero average, the affine relation (6.74) implies

$$Wg(u,s) = \frac{p}{\sqrt{l}}\,Wf\Big(l(u-r), sl\Big).$$

Suppose that ψ has a compact support included in $[-K, K]$. The affine invariance (6.73) of f over $\mathcal{S}_i = [a_i, b_i]$ produces an affine invariance for all wavelets having a support included in $\mathcal{S}_i$. For any $s < (b_i - a_i)/K$ and any $u \in [a_i + Ks, b_i - Ks]$,

$$Wf(u,s) = \frac{p_i}{\sqrt{l_i}}\,Wf\Big(l_i(u-r_i), sl_i\Big).$$

The wavelet transform's self-similarity implies that the positions and values of its modulus maxima are also self-similar. This can be used to recover unknown affine-invariance properties with a voting procedure based on wavelet modulus maxima [310].

EXAMPLE 6.9

A Cantor measure is constructed over a Cantor set. Let $d\mu_0(x) = dx$ be the uniform Lebesgue measure on $[0, 1]$. As in the Cantor set construction, this measure is subdivided into three uniform measures over $[0, 1/3]$, $[1/3, 2/3]$, and $[2/3, 1]$ with integrals equal to p_1, 0, and p_2, respectively. We impose $p_1 + p_2 = 1$ to obtain a total measure $d\mu_1$ on $[0, 1]$ with an integral equal to 1. This operation is iteratively repeated by dividing each uniform measure of integral p over $[a, a+l]$ into three equal parts where the integrals are $p_1 p$, 0, and $p_2 p$, respectively, over $[a, a+l/3]$, $[a+l/3, a+2l/3]$, and $[a+2l/3, a+l]$. This is illustrated in Figure 6.17. After each subdivision, the resulting measure $d\mu_n$ has a unit integral. In the limit, we obtain a Cantor measure $d\mu_\infty$ of unit integral with a support that is the triadic Cantor set.

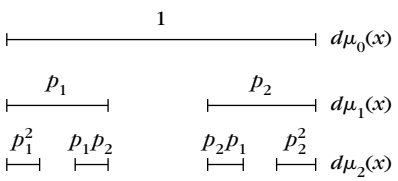

FIGURE 6.17

Two subdivisions of the uniform measure on $[0, 1]$ with left and right weights p_1 and p_2. The Cantor measure $d\mu_\infty$ is the limit of an infinite number of these subdivisions.

EXAMPLE 6.10

A devil's staircase is the integral of a Cantor measure:

$$f(t) = \int_0^t d\mu_\infty(x). \tag{6.75}$$

It is a continuous function that increases from 0 to 1 on $[0, 1]$. The recursive construction of the Cantor measure implies that f is self-similar:

$$f(t) = \begin{cases} p_1 f(3t) & \text{if } t \in [0, 1/3] \\ p_1 & \text{if } t \in [1/3, 2/3] \\ p_1 + p_2 f(3t-2) & \text{if } t \in [2/3, 0]. \end{cases}$$

Figure 6.18 displays the devil's staircase obtained with $p_1 = p_2 = 0.5$. The wavelet transform in (b) is calculated with a wavelet that is the first derivative of a Gaussian. The self-similarity of f yields a wavelet transform and modulus maxima that are self-similar. The subdivision of each interval in three parts appears through the multiplication by 2 maxima lines when the scale is multiplied by 3. This Cantor construction is generalized with different interval subdivisions and weight allocations beginning from the same Lebesgue measure $d\mu_0$ on $[0, 1]$ [5].

6.4.2 Singularity Spectrum

Finding the distribution of singularities in a multifractal signal f is particularly important for analyzing its properties. The spectrum of singularity measures the global repartition of singularities having different Lipschitz regularity. The pointwise Lipschitz regularity of f is given by Definition 6.1.

Definition 6.1: *Spectrum.* Let $\mathcal{S}_\alpha$ be the set of all points $t \in \mathbb{R}$ where the pointwise Lipschitz regularity of f is equal to α. The spectrum of singularity $D(\alpha)$ of f is the fractal dimension of $\mathcal{S}_\alpha$. The support of $D(\alpha)$ is the set of α such that $\mathcal{S}_\alpha$ is not empty.

This spectrum was originally introduced by Frisch and Parisi [264] to analyze the homogeneity of multifractal measures that model the energy dissipation of turbulent fluids. It was then extended by Arneodo, Bacry, and Muzy [381] to multifractal signals. The fractal dimension definition (6.72) shows that if we make a disjoint cover of the support of f with intervals of size s, then the number of intervals that intersect $\mathcal{S}_\alpha$ is

$$N_\alpha(s) \sim s^{-D(\alpha)}. \tag{6.76}$$

The singularity spectrum gives the proportion of Lipschitz α singularities that appear at any scale s. A multifractal f is said to be homogeneous if all singularities have the same Lipschitz exponent α_0, which means the support of $D(\alpha)$ is restricted to $\{\alpha_0\}$. Fractional Brownian motions are examples of homogeneous multifractals.

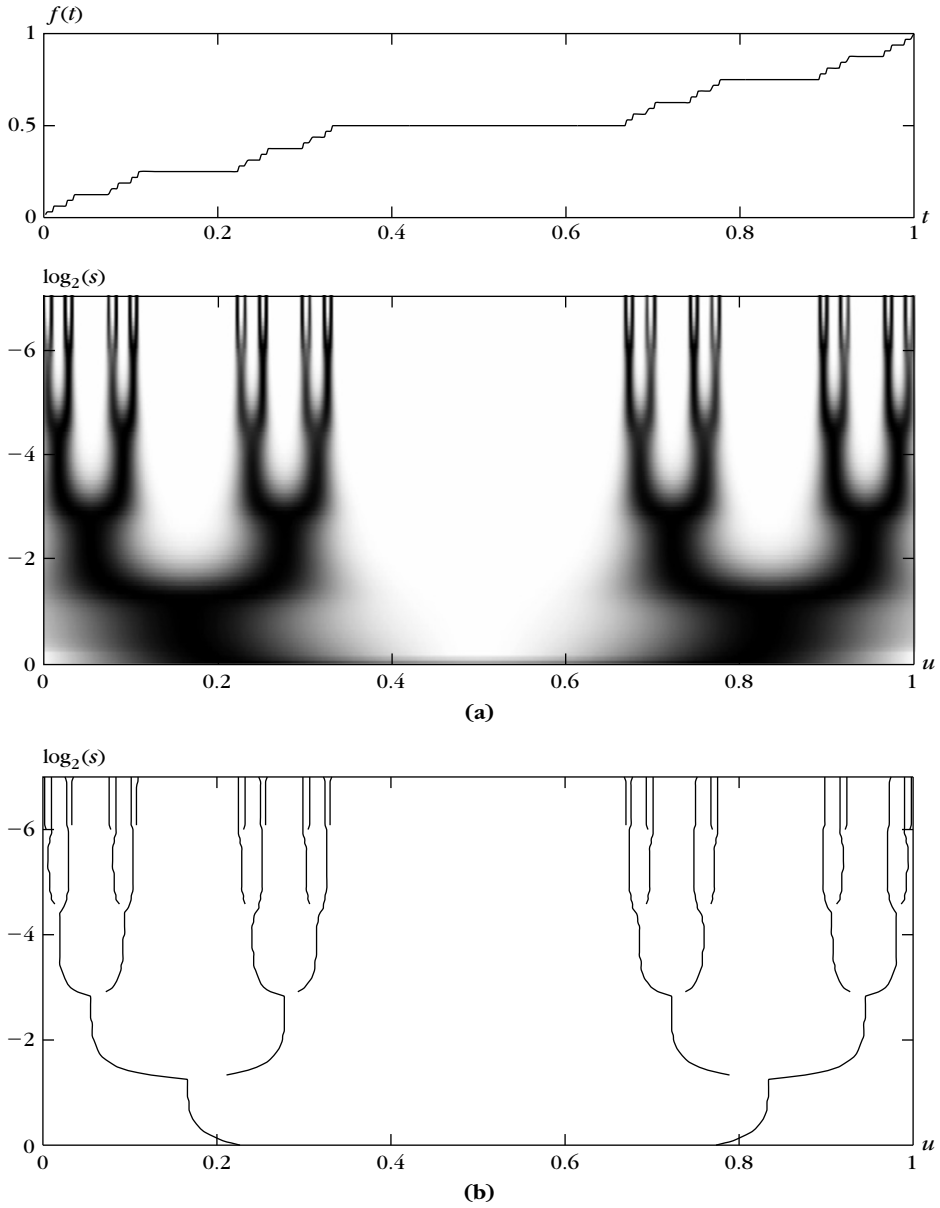

FIGURE 6.18

Devil's staircase calculated from a Cantor measure with equal weights $p_1 = p_2 = 0.5$. **(a)** Wavelet transform $Wf(u, s)$ computed with $\psi = -\theta'$ where θ is Gaussian. **(b)** Wavelet transform modulus maxima.

Partition Function

One cannot compute the pointwise Lipschitz regularity of a multifractal because its singularities are not isolated, and the finite numerical resolution is not sufficient to discriminate them. It is possible, however, to measure the singularity spectrum of multifractals from the wavelet transform local maxima using a global partition function introduced by Arneodo, Bacry, and Muzy [381].

Let ψ be a wavelet with n vanishing moments. Theorem 6.5 proves that if f has pointwise Lipschitz regularity $\alpha_0 < n$ at v, then the wavelet transform $Wf(u, s)$ has a sequence of modulus maxima that converges toward v at fine scales. Thus, the set of maxima at the scale s can be interpreted as a covering of the singular support of f with wavelets of scale s. At these maxima locations,

$$|Wf(u, s)| \sim s^{\alpha_0 + 1/2}.$$

Let $\{u_p(s)\}_{p \in \mathbb{Z}}$ be the position of all local maxima of $|Wg(u, s)|$ at a fixed scale s. The partition function $\mathcal{Z}$ measures the sum at a power q of all these wavelet modulus maxima:

$$\mathcal{Z}(q, s) = \sum_p |Wf(u_p, s)|^q. \tag{6.77}$$

At each scale s, any two consecutive maxima u_p and u_{p+1} are supposed to have a distance $|u_{p+1} - u_p| > \varepsilon s$, for some $\varepsilon > 0$. If not, over intervals of size εs, the sum (6.77) includes only the maxima of largest amplitude. This protects the partition function from the multiplication of very close maxima created by fast oscillations.

For each $q \in \mathbb{R}$, the scaling exponent $\tau(q)$ measures the asymptotic decay of $\mathcal{Z}(q, s)$ at fine scales s:

$$\tau(q) = \liminf_{s \to 0} \frac{\log \mathcal{Z}(q, s)}{\log s}.$$

This typically means that

$$\mathcal{Z}(q, s) \sim s^{\tau(q)}.$$

Legendre Transform

Theorem 6.8 relates $\tau(q)$ to the Legendre transform of $D(\alpha)$ for self-similar signals. This result was established in [91] for a particular class of fractal signals and generalized by Jaffard [313].

Theorem 6.8: *Arneodo, Bacry, Jaffard, Muzy.* Let $\Lambda = [\alpha_{\min}, \alpha_{\max}]$ be the support of $D(\alpha)$. Let ψ be a wavelet with $n > \alpha_{\max}$ vanishing moments. If f is a self-similar signal, then

$$\tau(q) = \min_{\alpha \in \Lambda} \Big(q\,(\alpha + 1/2) - D(\alpha) \Big). \tag{6.78}$$

Proof. The detailed proof is long; we only give an intuitive justification. The sum (6.77) over all maxima positions is replaced by an integral over the Lipschitz parameter. At the scale s,

(6.76) indicates that the density of modulus maxima that cover a singularity with Lipschitz exponent α is proportional to $s^{-D(\alpha)}$. At locations where f has Lipschitz regularity α, the wavelet transform decay is approximated by

$$|Wf(u,s)| \sim s^{\alpha+1/2}.$$

It follows that

$$\mathcal{Z}(q,s) \sim \int_{\Lambda} s^{q(\alpha+1/2)}\, s^{-D(\alpha)}\, d\alpha.$$

When s goes to 0 we derive that $\mathcal{Z}(q,s) \sim s^{\tau(q)}$ for $\tau(q) = \min_{\alpha\in\Lambda}(q(\alpha+1/2) - D(\alpha))$. ■

This theorem proves that the scaling exponent $\tau(q)$ is the Legendre transform of $D(\alpha)$. It is necessary to use a wavelet with enough vanishing moments to measure all Lipschitz exponents up to $\alpha_{\max}$. In numerical calculations $\tau(q)$ is computed by evaluating the sum $\mathcal{Z}(q,s)$. Thus, we need to invert the Legendre transform (6.78) to recover the spectrum of singularity $D(\alpha)$.

Theorem 6.9.

- The scaling exponent $\tau(q)$ is a concave and increasing function of q.
- The Legendre transform (6.78) is invertible if and only if $D(\alpha)$ is concave, in which case

$$D(\alpha) = \min_{q\in\mathbb{R}} \Big(q\,(\alpha+1/2) - \tau(q) \Big). \tag{6.79}$$

- The spectrum $D(\alpha)$ of self-similar signals is concave.

Proof. The proof that $D(\alpha)$ is concave for self-similar signals can be found in [313]. We concentrate on the properties of the Legendre transform that are important in numerical calculations. To simplify the proof, let us suppose that $D(q)$ is twice differentiable. The minimum of the Legendre transform (6.78) is reached at a critical point $q(\alpha)$. Computing the derivative of $q(\alpha+1/2) - D(\alpha)$ with respect to α gives

$$q(\alpha) = \frac{dD}{d\alpha}, \tag{6.80}$$

with

$$\tau(q) = q\left(\alpha + \frac{1}{2}\right) - D(\alpha). \tag{6.81}$$

Since it is a minimum, the second derivative of $\tau(q(\alpha))$ with respect to α is positive, from which we derive that

$$\frac{d^2 D(\alpha(q))}{d\alpha^2} \leqslant 0.$$

This proves that $\tau(q)$ depends only on the values where $D(\alpha)$ has a negative second derivative. Thus, we can recover $D(\alpha)$ from $\tau(q)$ only if it is concave.

The derivative of $\tau(q)$ is

$$\frac{d\tau(q)}{dq} = \alpha + \frac{1}{2} + q\,\frac{d\alpha}{dq} - \frac{d\alpha}{dq}\,\frac{dD(\alpha)}{d\alpha} = \alpha + \frac{1}{2} \geqslant 0. \tag{6.82}$$

Therefore, it is increasing. Its second derivative is

$$\frac{d^2\tau(q)}{dq^2} = \frac{d\alpha}{dq}.$$

Taking the derivative of (6.80) with respect to q proves that

$$\frac{d\alpha}{dq}\,\frac{d^2D(\alpha)}{d\alpha^2} = 1.$$

Since $\frac{d^2D(\alpha)}{d\alpha^2} \leqslant 0$, we derive that $\frac{d^2\tau(q)}{dq^2} \leqslant 0$. Thus, $\tau(q)$ is concave. By using (6.81), (6.82), and the fact that $\tau(q)$ is concave, we verify that

$$D(\alpha) = \min_{q\in\mathbb{R}} \Big(q\,(\alpha + 1/2) - \tau(q) \Big).$$

■

The spectrum $D(\alpha)$ of self-similar signals is concave and therefore can be calculated from $\tau(q)$ with the inverse Legendre formula (6.79). This formula is also valid for a much larger class of multifractals; for example, it is verified for statistical self-similar signals such as realizations of fractional Brownian motions. Multifractals having some stochastic self-similarity have a spectrum that can often be calculated as an inverse Legendre transform (6.79). However, let us emphasize that this formula is not exact for any function f because its spectrum of singularity $D(\alpha)$ is not necessarily concave. In general, Jaffard proved [313] that the Legendre transform (6.79) gives only an upper bound of $D(\alpha)$. These singularity spectrum properties are studied in detail in [46].

Figure 6.19 illustrates the properties of a concave spectrum $D(\alpha)$. The Legendre transform (6.78) proves that its maximum is reached at

$$D(\alpha_0) = \max_{\alpha\in\Lambda} D(\alpha) = -\tau(0).$$

It is the fractal dimension of the Lipschitz exponent α_0 most frequently encountered in f. Since all other Lipschitz α singularities appear over sets of lower dimension, if $\alpha_0 < 1$, then $D(\alpha_0)$ is also the fractal dimension of the singular support of f. The spectrum $D(\alpha)$ for $\alpha < \alpha_0$ depends on $\tau(q)$ for $q > 0$, and for $\alpha > \alpha_0$ it depends on $\tau(q)$ for $q < 0$.

Numerical Calculations

To compute $D(\alpha)$, we assume that the Legendre transform formula (6.79) is valid. We first calculate $\mathcal{Z}(q, s) = \sum_p |Wf(u_p, s)|^q$, then derive the decay scaling exponent $\tau(q)$, and finally compute $D(\alpha)$ with a Legendre transform. If $q < 0$, then the value of $\mathcal{Z}(q, s)$ depends mostly on the small-amplitude maxima $|Wf(u_p, s)|$. Numerical calculations may then become unstable. To avoid introducing spurious modulus

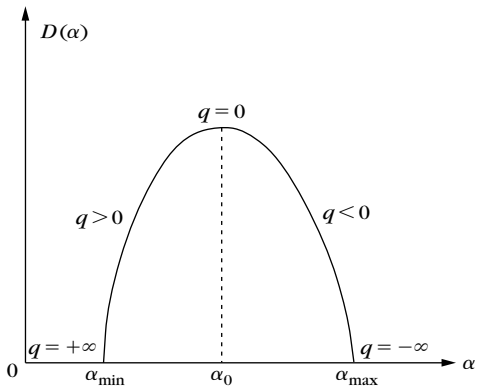

FIGURE 6.19

Concave spectrum $D(\alpha)$.

maxima created by numerical errors in regions where f is nearly constant, wavelet maxima are chained to produce maxima curve across scales. If $\psi = (-1)^p \theta^{(p)}$ where θ is a Gaussian, Theorem 6.6 proves that all maxima lines $u_p(s)$ define curves that propagate up to the limit $s = 0$. Thus, all maxima lines that do not propagate up to the finest scale are removed in the calculation of $\mathcal{Z}(q, s)$. The calculation of the spectrum $D(\alpha)$ proceeds as follows:

1. *Maxima*. Compute $Wf(u, s)$ and the modulus maxima at each scale s. Chain the wavelet maxima across scales.
2. *Partition function*. Compute
$$\mathcal{Z}(q, s) = \sum_p |Wf(u_p, s)|^q.$$
3. *Scaling*. Compute $\tau(q)$ with a linear regression of $\log_2 \mathcal{Z}(s, q)$ as a function of $\log_2 s$:
$$\log_2 \mathcal{Z}(q, s) \approx \tau(q) \log_2 s + C(q).$$
4. *Spectrum*. Compute
$$D(\alpha) = \min_{q \in \mathbb{R}} \Big(q(\alpha + 1/2) - \tau(q) \Big).$$

EXAMPLE 6.11

The spectrum of singularity $D(\alpha)$ of the devil's staircase (6.75) is a concave function that can be calculated analytically [292]. Suppose that $p_1 < p_2$. The support of $D(\alpha)$ is $[\alpha_{\min}, \alpha_{\max}]$ with
$$\alpha_{\min} = \frac{-\log p_2}{\log 3} \quad \text{and} \quad \alpha_{\max} = \frac{-\log p_1}{\log 3}.$$

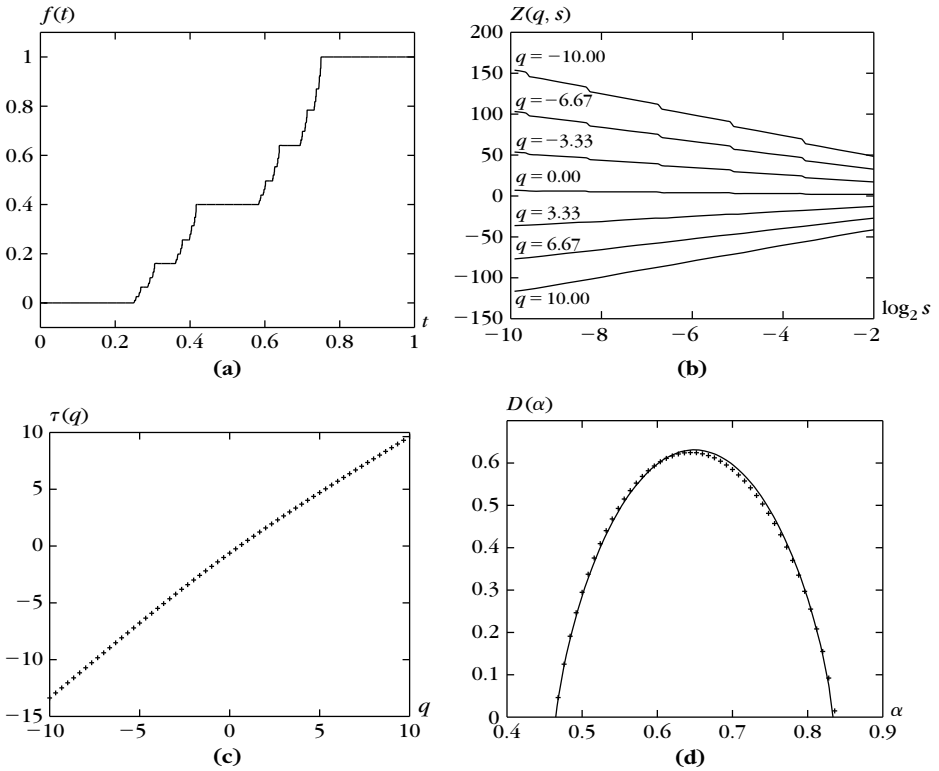

FIGURE 6.20

(a) Devil's staircase with $p_1 = 0.4$ and $p_2 = 0.6$. **(b)** Partition function $\mathcal{Z}(q, s)$ for several values of q. **(c)** Scaling exponent $\tau(q)$. **(d)** The theoretical spectrum $D(\alpha)$ is shown with a solid line. The spectrum values are calculated numerically with a Legendre transform of $\tau(q)$.

If $p_1 = p_2 = 1/2$, then the support of $D(\alpha)$ is reduced to a point, which means that all the singularities of f have the same Lipschitz $\log 2/\log 3$ regularity. The value $D(\log 2/\log 3)$ is then the fractal dimension of the triadic Cantor set and is equal to $\log 2/\log 3$.

Figure 6.20(a) shows a devil's staircase calculated with $p_1 = 0.4$ and $p_2 = 0.6$. Its wavelet transform is computed with $\psi = -\theta'$ where θ is a Gaussian. The decay of $\log_2 \mathcal{Z}(q, s)$ as a function of $\log_2 s$ is shown in Figure 6.20(b) for several values of q. The resulting $\tau(q)$ and $D(\alpha)$ are given by Figures 6.20(c, d). There is no numerical instability for $q < 0$, because there is no modulus maximum that has an amplitude close to zero. This is not the case if the wavelet transform is calculated with a wavelet that has more vanishing moments.

Smooth Perturbations

Let f be a multifractal with a spectrum of singularity $D(\alpha)$ calculated from $\tau(q)$. If a $\mathbf{C}^\infty$ signal g is added to f then the singularities are not modified and the singularity

spectrum of $\tilde{f} = f + g$ remains $D(\alpha)$. We study the effect of this smooth perturbation on the spectrum calculation.

The wavelet transform of $\tilde{f}$ is

$$W\tilde{f}(u,s) = Wf(u,s) + Wg(u,s).$$

Let $\tau(q)$ and $\tilde{\tau}(q)$ be the scaling exponent of the partition functions $\mathcal{Z}(q,s)$ and $\tilde{\mathcal{Z}}(q,s)$ calculated from the modulus maxima of $Wf(u,s)$ and $W\tilde{f}(u,s)$, respectively. We denote by $D(\alpha)$ and $\tilde{D}(\alpha)$ the Legendre transforms of $\tau(q)$ and $\tilde{\tau}(q)$, respectively. Theorem 6.10 relates $\tau(q)$ and $\tilde{\tau}(q)$.

Theorem 6.10: *Arneodo, Bacry, Muzy.* Let ψ be a wavelet with exactly n vanishing moments. Suppose that f is a self-similar function.

- If g is a polynomial of degree $p < n$, then $\tau(q) = \tilde{\tau}(q)$ for all $q \in \mathbb{R}$.
- If $g^{(n)}$ is almost everywhere nonzero, then

$$\tilde{\tau}(q) = \begin{cases} \tau(q) & \text{if } q > q_c \\ (n+1/2)\,q & \text{if } q \leqslant q_c \end{cases} \tag{6.83}$$

  where q_c is defined by $\tau(q_c) = (n+1/2)q_c$.

Proof. If g is a polynomial of degree $p < n$, then $Wg(u,s) = 0$. The addition of g does not modify the calculation of the singularity spectrum based on wavelet maxima, so $\tau(q) = \tilde{\tau}(q)$ for all $q \in \mathbb{R}$.

If g is a $\mathbf{C}^\infty$ function that is not a polynomial then its wavelet transform is generally nonzero. We justify (6.83) with an intuitive argument that is not a proof. A rigorous proof can be found in [91]. Since ψ has exactly n vanishing moments, (6.15) proves that

$$|Wg(u,s)| \sim K\, s^{n+1/2}\, g^{(n)}(u).$$

We suppose that $g^{(n)}(u) \neq 0$. For $\tau(q) \leqslant (n+1/2)q$, since $|Wg(u,s)|^q \sim s^{q(n+1/2)}$ has a faster asymptotic decay than $s^{\tau(q)}$ when s goes to zero, one can verify that $\tilde{\mathcal{Z}}(q,s)$ and $\mathcal{Z}(q,s)$ have the same scaling exponent, $\tilde{\tau}(q) = \tau(q)$. If $\tau(q) > (n+1/2)q$, which means that $q \leqslant q_c$, then the decay of $|W\tilde{f}(u,s)|^q$ is controlled by the decay of $|Wg(u,s)|^q$, so $\tilde{\tau}(q) = (n+1/2)q$. ■

This theorem proves that the addition of a nonpolynomial smooth function introduces a bias in the calculation of the singularity spectrum. Let α_c be the critical Lipschitz exponent corresponding to q_c:

$$D(\alpha_c) = q_c\,(\alpha_c + 1/2) - \tau(q_c).$$

The Legendre transform of $\tilde{\tau}(q)$ in (6.83) yields

$$\tilde{D}(\alpha) = \begin{cases} D(\alpha) & \text{if } \alpha \leqslant \alpha_c \\ 0 & \text{if } \alpha = n \\ -\infty & \text{if } \alpha > \alpha_c \text{ and } \alpha \neq n. \end{cases} \tag{6.84}$$

This modification is illustrated by Figure 6.21.

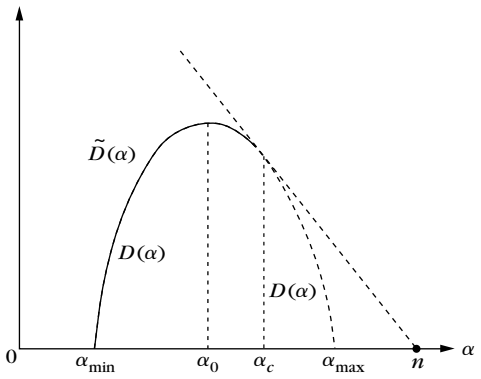

FIGURE 6.21

If ψ has n vanishing moments, in the presence of a $\mathbf{C}^\infty$ perturbation the computed spectrum $\tilde{D}(\alpha)$ is identical to the true spectrum $D(\alpha)$ for $\alpha \leqslant \alpha_c$. Its support is reduced to $\{n\}$ for $\alpha > \alpha_c$.

The bias introduced by the addition of smooth components can be detected experimentally by modifying the number n of vanishing moments of ψ. Indeed the value of q_c depends on n. If the singularity spectrum varies when changing the number of vanishing moments of the wavelet then it indicates the presence of a bias.

6.4.3 Fractal Noises

Fractional Brownian motions are statistically self-similar Gaussian processes that give interesting models for a wide class of natural phenomena [371]. Despite their nonstationarity, one can define a power spectrum that has a power decay. Realizations of fractional Brownian motions are almost everywhere singular, with the same Lipschitz regularity at all points.

We often encounter fractal noise processes that are not Gaussian although their power spectrum has a power decay. Realizations of these processes may include singularities of various types. The spectrum of singularity is then important in analyzing their properties. This is illustrated by an application to hydrodynamic turbulence.

Definition 6.2: *Fractional Brownian Motion.* A fractional Brownian motion of Hurst exponent $0 < H < 1$ is a zero-mean Gaussian process B_H such that

$$B_H(0) = 0,$$

and

$$E\{|B_H(t) - B_H(t-\Delta)|^2\} = \sigma^2 |\Delta|^{2H}. \tag{6.85}$$

Property (6.85) imposes that the deviation of $|B_H(t) - B_H(t-\Delta)|$ be proportional to $|\Delta|^H$. As a consequence, one can prove that any realization f of B_H is almost everywhere singular with a pointwise Lipschitz regularity $\alpha = H$. The smaller

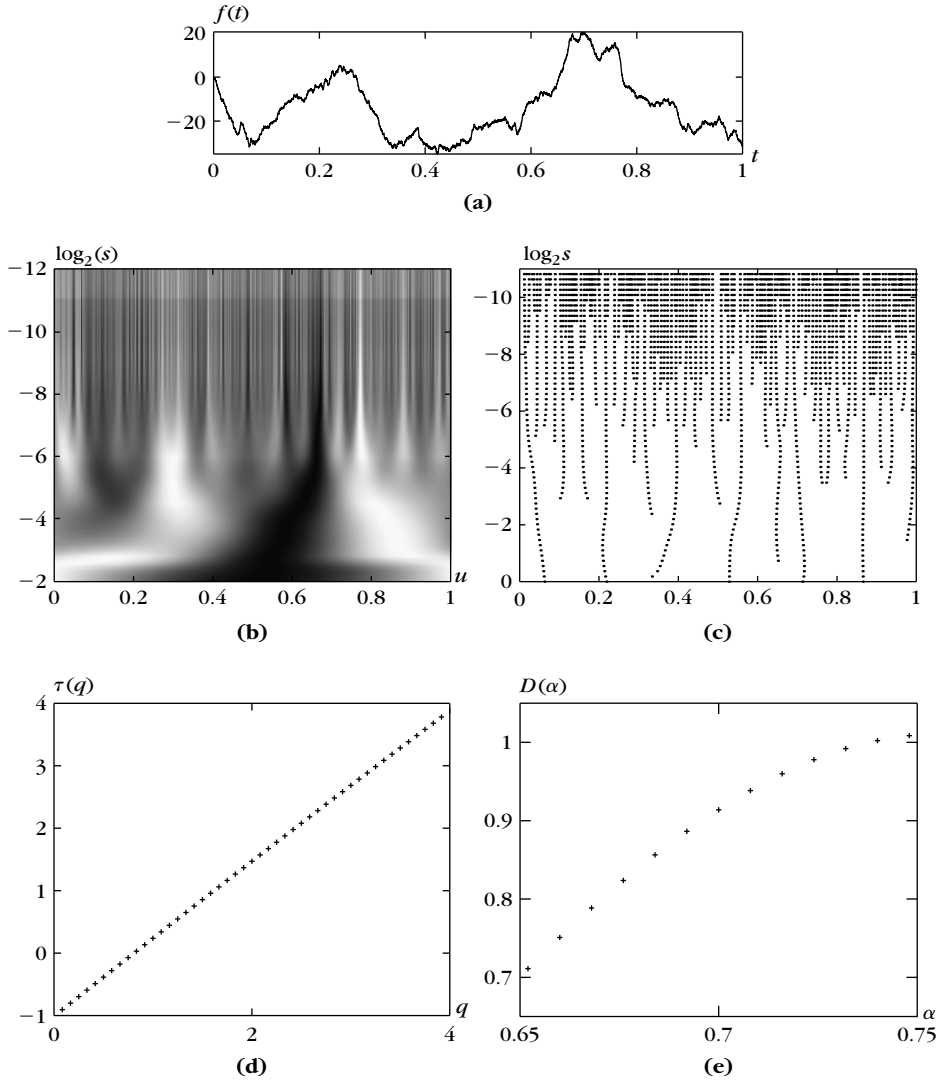

FIGURE 6.22

(a) One realization of a fractional Brownian motion for a Hurst exponent $H = 0.7$. **(b)** Wavelet transform. **(c)** Modulus maxima of its wavelet transform. **(d)** Scaling exponent $\tau(q)$. **(e)** Resulting $D(\alpha)$ over its support.

H is, the more singular f is. Figure 6.22(a) shows the graph of one realization for $H = 0.7$.

Setting $\Delta = t$ in (6.85) yields

$$E\{|B_H(t)|^2\} = \sigma^2 |t|^{2H}.$$

Developing (6.85) for $\Delta = t - u$ also gives

$$E\{B_H(t)\, B_H(u)\} = \frac{\sigma^2}{2}\left(|t|^{2H} + |u|^{2H} - |t-u|^{2H}\right). \tag{6.86}$$

The covariance does not depend only on $t-u$, which proves that a fractional Brownian motion is nonstationary.

The statistical self-similarity appears when scaling this process. One can derive from (6.86) that for any $s > 0$,

$$E\{B_H(st)\, B_H(su)\} = E\{s^H\, B_H(t)\, s^H\, B_H(u)\}.$$

Since $B_H(st)$ and $s^H\, B_H(t)$ are two Gaussian processes with the same mean and covariance, they have the same probability distribution,

$$B_H(st) \equiv s^H\, B_H(t),$$

where $\equiv$ denotes an equality of finite-dimensional distributions.

Power Spectrum

Although B_H is not stationary, one can define a generalized power spectrum. This power spectrum is introduced by proving that the increments of a fractional Brownian motion are stationary and by computing their power spectrum [73].

Theorem 6.11. Let $g_\Delta(t) = \delta(t) - \delta(t-\Delta)$. The increment

$$I_{H,\Delta}(t) = B_H \star g_\Delta(t) = B_H(t) - B_H(t-\Delta) \tag{6.87}$$

is a stationary process with power spectrum

$$\hat{R}_{I_{H,\Delta}}(\omega) = \frac{\sigma_H^2}{|\omega|^{2H+1}}\, |\hat{g}_\Delta(\omega)|^2. \tag{6.88}$$

Proof. The covariance of $I_{H,\Delta}$ is computed with (6.86):

$$E\{I_{H,\Delta}(t)\, I_{H,\Delta}(t-\tau)\} = \frac{\sigma^2}{2}(|\tau-\Delta|^{2H} + |\tau+\Delta|^{2H} - 2|\tau|^{2H}) = R_{I_{H,\Delta}}(\tau). \tag{6.89}$$

The power spectrum $\hat{R}_{I_{H,\Delta}}(\omega)$ is the Fourier transform of $R_{I_{H,\Delta}}(\tau)$. One can verify that the Fourier transform of the distribution $f(\tau) = |\tau|^{2H}$ is $\hat{f}(\omega) = -\lambda_H\, |\omega|^{-(2H+1)}$, with $\lambda_H > 0$. Thus, we derive that the Fourier transform of (6.89) can be written as

$$\hat{R}_{I_{H,\Delta}}(\omega) = 2\,\sigma^2\,\lambda_H\, |\omega|^{-(2H+1)}\, \sin^2\frac{\Delta\omega}{2},$$

which proves (6.88) for $\sigma_H^2 = \sigma^2 \lambda_H / 2$. ■

If $X(t)$ is a stationary process, then we know that $Y(t) = X \star g(t)$ is also stationary and the power spectrum of both processes is related by

$$\hat{R}_X(\omega) = \frac{\hat{R}_Y(\omega)}{|\hat{g}(\omega)|^2}. \tag{6.90}$$

Although $B_H(t)$ is not stationary, Theorem 6.11 proves that $I_{H,\Delta}(t) = B_H \star g_\Delta(t)$ is stationary. As in (6.90), it is tempting to define a "generalized" power spectrum calculated with (6.88):

$$\hat{R}_{B_H}(\omega) = \frac{\hat{R}_{I_{H,\Delta}}(\omega)}{|\hat{g}_\Delta(\omega)|^2} = \frac{\sigma_H^2}{|\omega|^{2H+1}}. \tag{6.91}$$

The nonstationarity of $B_H(t)$ appears in the energy blow-up at low frequencies. The increments $I_{H,\Delta}(t)$ are stationary because the multiplication by $|\hat{g}_\Delta(\omega)|^2 = O(\omega^2)$ removes the explosion of the low-frequency energy. One can generalize this result and verify that if g is an arbitrary stable filter with a transfer function that satisfies $|\hat{g}(\omega)| = O(\omega)$, then $Y(t) = B_H \star g(t)$ is a stationary Gaussian process with a power spectrum that is

$$\hat{R}_Y(\omega) = \frac{\sigma_H^2}{|\omega|^{2H+1}} |\hat{g}(\omega)|^2. \tag{6.92}$$

Wavelet Transform

The wavelet transform of a fractional Brownian motion is

$$WB_H(u, s) = B_H \star \bar{\psi}_s(u). \tag{6.93}$$

Since ψ has at least one vanishing moment, necessarily $|\hat{\psi}(\omega)| = O(\omega)$ in the neighborhood of $\omega = 0$. The wavelet filter $g = \bar{\psi}_s$ has a Fourier transform $\hat{g}(\omega) = \sqrt{s}\,\hat{\psi}^*(s\omega) = O(\omega)$ near $\omega = 0$. This proves that for a fixed s the process $Y_s(u) = WB_H(u, s)$ is a Gaussian stationary process [258] with a power spectrum that is calculated with (6.92):

$$\hat{R}_{Y_s}(\omega) = s\,|\hat{\psi}(s\omega)|^2 \frac{\sigma_H^2}{|\omega|^{2H+1}} = s^{2H+2}\,\hat{R}_{Y_1}(s\omega). \tag{6.94}$$

The self-similarity of the power spectrum and the fact that B_H is Gaussian are sufficient to prove that $WB_H(u, s)$ is self-similar across scales:

$$WB_H(u, s) \equiv s^{H+1/2}\, WB_H\left(\frac{u}{s}, 1\right),$$

where the equivalence means that they have the same finite distributions. Interesting characterizations of fractional Brownian motion properties are also obtained by decomposing these processes in wavelet bases [46, 73, 490].

EXAMPLE 6.12

Figure 6.22(a) on page 253 displays one realization of a fractional Brownian with $H = 0.7$. The wavelet transform and its modulus maxima are shown in Figures 6.22(b) and 6.22(c). The partition function (6.77) is computed from the wavelet modulus maxima. Figure 6.22(d) gives the scaling exponent $\tau(q)$, which is nearly a straight line.

Fractional Brownian motions are homogeneous fractals with Lipschitz exponents equal to H. In this example, the theoretical spectrum $D(\alpha)$ has a support reduced to $\{0.7\}$ with $D(0.7) = 1$. The estimated spectrum in Figure 6.22(e) is calculated with a Legendre transform of $\tau(q)$. Its support is $[0.65, 0.75]$. There is an estimation error because the calculations are performed on a signal of finite size.

Fractal Noises

Some physical phenomena produce more general fractal noises $X(t)$, which are not Gaussian processes, but they do have stationary increments. As for fractional Brownian motions, one can define a generalized power spectrum with a power decay

$$\hat{R}_X(\omega) = \frac{\sigma_H^2}{|\omega|^{2H+1}}.$$

These processes are transformed into a wide-sense stationary process by a convolution with a stable filter g that removes the lowest frequencies $|\hat{g}(\omega)| = O(\omega)$. Thus, one can determine that the wavelet transform $Y_s(u) = WX(u, s)$ is a stationary process at any fixed scale s. Its spectrum is the same as the spectrum (6.94) of fractional Brownian motions. If $H < 1$, the asymptotic decay of $\hat{R}_X(\omega)$ indicates that realizations of $X(t)$ are singular functions; however, it gives no information about the distribution of these singularities.

As opposed to fractional Brownian motions, general fractal noises have realizations that may include singularities of various types. Such multifractals are differentiated from realizations of fractional Brownian motions by computing their singularity spectrum $D(\alpha)$. For example, the velocity fields of fully developed turbulent flows have been modeled by fractal noises, but the calculation of the singularity spectrum clearly shows that these flows differ in important ways from fractional Brownian motions.

Hydrodynamic Turbulence

Fully developed turbulence appears in incompressible flows at high Reynolds numbers. Understanding the properties of hydrodynamic turbulence is a major problem of modern physics, which remains mostly open despite an intense research effort since the first theory of Kolmogorov in 1941 [331]. The number of degrees of liberty of three-dimensional turbulence is considerable, which produces extremely complex spatio-temporal behavior. No formalism is yet able to build a statistical physics framework based on the Navier-Stokes equations that would enable us to understand the global behavior of turbulent flows as it is done in thermodynamics.

In 1941, Kolmogorov [331] formulated a statistical theory of turbulence. The velocity field is modeled as a process $V(x)$ that has increments with variance

$$E\{|V(x+\Delta) - V(x)|^2\} \sim \varepsilon^{2/3}\, \Delta^{2/3}.$$

The constant ε is a rate of dissipation of energy per unit of mass and time, which is supposed to be independent of the location. This indicates that the velocity field is statistically homogeneous with Lipschitz regularity $\alpha = H = 1/3$. The theory predicts that a one-dimensional trace of a three-dimensional velocity field is a fractal noise process with stationary increments that have spectrum decays with a power exponent $2H + 1 = 5/3$:

$$\hat{R}_V(\omega) = \frac{\sigma_H^2}{|\omega|^{5/3}}.$$

The success of this theory comes from numerous experimental verifications of this power spectrum decay. However, the theory does not take into account the existence of coherent structures such as vortices. These phenomena contradict the hypothesis of homogeneity, which is at the root of Kolmogorov's 1941 theory.

Kolmogorov [332] modified the homogeneity assumption in 1962 by introducing an energy dissipation rate $\varepsilon(x)$ that varies with the spatial location x. This opens the door to "local stochastic self-similar" multifractal models, first developed by Mandelbrot [370] to explain energy exchanges between fine-scale structures and large-scale structures. The spectrum of singularity $D(\alpha)$ is playing an important role in testing these models [264]. Calculations with wavelet maxima on turbulent velocity fields [5] show that $D(\alpha)$ is maximum at $1/3$, as predicted by the Kolmogorov theory. However, $D(\alpha)$ does not have a support reduced to $\{1/3\}$, which verifies that a turbulent velocity field is not a homogeneous process. Models based on the wavelet transform have been introduced to explain the distribution of vortices in turbulent fluids [13, 251, 252].

6.5 EXERCISES

6.1 [2] *Lipschitz regularity*:

(a) Prove that if f is uniformly Lipschitz α on $[a, b]$, then it is pointwise Lipschitz α at all $t_0 \in [a, b]$.

(b) Show that $f(t) = t \sin t^{-1}$ is Lipschitz 1 at all $t_0 \in [-1, 1]$ and verify that it is uniformly Lipschitz α over $[-1, 1]$ only for $\alpha \leqslant 1/2$. *Hint:* Consider the points $t_n = (n + 1/2)^{-1} \pi^{-1}$.

6.2 [2] *Regularity of derivatives*:

(a) Prove that f is uniformly Lipschitz $\alpha > 1$ over $[a, b]$ if and only if f' is uniformly Lipschitz $\alpha - 1$ over $[a, b]$.

(b) Show that f may be pointwise Lipschitz $\alpha > 1$ at t_0 while f' is not pointwise Lipschitz $\alpha - 1$ at t_0. Consider $f(t) = t^2 \cos t^{-1}$ at $t = 0$.

6.3 [2] Find $f(t)$ that is uniformly Lipschitz 1, but does not satisfy the sufficient Fourier condition (6.1).

6.4 [1] Let $f(t) = \cos \omega_0 t$ and $\psi(t)$ be a wavelet that is symmetric about 0.

(a) Verify that

$$Wf(u,s) = \sqrt{s}\,\hat{\psi}(s\omega_0)\cos\omega_0 t.$$

(b) Find the equations of the curves of wavelet modulus maxima in the time-scale plane (u,s). Relate the decay of $|Wf(u,s)|$ along these curves to the number n of vanishing moments of ψ.

6.5 [1] Let $f(t) = |t|^\alpha$. Show that $Wf(u,s) = s^{\alpha+1/2}\,Wf(u/s, 1)$. Prove that it is not sufficient to measure the decay of $|Wf(u,s)|$ when s goes to zero at $u = 0$ in order to compute the Lipschitz regularity of f at $t = 0$.

6.6 [3] Let $f(t) = |t|^\alpha \sin|t|^{-\beta}$ with $\alpha > 0$ and $\beta > 0$. What is the pointwise Lipschitz regularity of f and f' at $t = 0$? Find the equation of the ridge curve in the (u,s) plane along which the high-amplitude wavelet coefficients $|Wf(u,s)|$ converge to $t = 0$ when s goes to zero. Compute the maximum values of α and α' such that $Wf(u,s)$ satisfies (6.21).

6.7 [2] For a complex wavelet, we call *lines of constant phase* the curves in the (u,s) plane along which the complex phase of $Wf(u,s)$ remains constant when s varies.

(a) If $f(t) = |t|^\alpha$, prove that the lines of constant phase converge toward the singularity at $t = 0$ when s goes to zero. Verify this numerically.

(b) Let ψ be a real wavelet and $Wf(u,s)$ be the real wavelet transform of f. Show that the modulus maxima of $Wf(u,s)$ correspond to lines of constant phase of an analytic wavelet transform, which is calculated with a particular analytic wavelet ψ^a that you will specify.

6.8 [3] Prove that if $f = \mathbf{1}_{[0,+\infty)}$, then the number of modulus maxima of $Wf(u,s)$ at each scale s is larger than or equal to the number of vanishing moments of ψ.

6.9 [2] The spectrum of singularity of the Riemann function

$$f(t) = \sum_{n=-\infty}^{+\infty} \frac{1}{n^2} \sin n^2 t$$

is defined on its support by $D(\alpha) = 4\alpha - 2$ if $\alpha \in [1/2, 3/4]$ and $D(3/2) = 0$ [304, 313]. Verify this result numerically by computing this spectrum from the partition function of a wavelet transform modulus maxima.

6.10 [3] Let $\psi = -\theta'$ where θ is a positive window of compact support. If f is a Cantor devil's staircase, prove that there exist lines of modulus maxima that converge toward each singularity.

6.11 [3] Implement an algorithm that detects oscillating singularities by following the ridges of an analytic wavelet transform when the scale s decreases. Test your algorithm on $f(t) = \sin t^{-1}$.

6.12 [2] Implement an algorithm that reconstructs a signal from the local maxima of its dyadic wavelet transform with a dual synthesis (6.48) using a conjugate-gradient algorithm.

6.13 [3] Let $X[n] = f[n] + W[n]$ be a signal of size N, where W is a Gaussian white noise of variance σ^2. Implement in WaveLab an estimator of f that thresholds at $T = \lambda\,\sigma$ the maxima of a dyadic wavelet transform of X. The estimation of f is reconstructed from the thresholded maxima representation with the dual synthesis (6.48) implemented with a conjugate-gradient algorithm. Compare numerically the risk of this estimator with the risk of a thresholding estimator over the translation-invariant dyadic wavelet transform of X.

6.14 [2] Let $\theta(t)$ be a Gaussian of variance 1.

(a) Prove that the Laplacian of a two-dimensional Gaussian

$$\psi(x_1, x_2) = \frac{\partial^2 \theta(x_1)}{\partial x^2}\,\theta(x_2) + \theta(x_1)\,\frac{\partial^2 \theta(x_2)}{\partial x_2^2}$$

satisfies the dyadic wavelet condition (5.101) (there is only one wavelet).

(b) Explain why the zero-crossings of this dyadic wavelet transform provide the locations of multiscale edges in images. Compare the position of these zero-crossings with the wavelet modulus maxima obtained with $\psi^1(x_1, x_2) = -\theta'(x_1)\,\theta(x_2)$ and $\psi^2(x_1, x_2) = -\theta(x_1)\,\theta'(x_2)$.

6.15 [2] The covariance of a fractional Brownian motion $B_H(t)$ is given by (6.86). Show that the wavelet transform at a scale s is stationary by verifying that

$$E\Big\{WB_H(u_1, s)\,WB_H(u_2, s)\Big\} = -\frac{\sigma^2}{2}\,s^{2H+1}\int_{-\infty}^{+\infty} |t|^{2H}\,\Psi\Big(\frac{u_1 - u_2}{s} - t\Big)\,dt,$$

with $\Psi(t) = \psi \star \bar{\psi}(t)$ and $\bar{\psi}(t) = \psi(-t)$.

6.16 [2] Let $X(t)$ be a stationary Gaussian process with a covariance $R_X(\tau) = E\{X(t)X(t-\tau)\}$ that is twice differentiable. One can prove that the average number of zero-crossings over an interval of size 1 is $-\pi R_X''(0)\left(\pi^2\,R_X(0)\right)^{-1}$ [53]. Let $B_H(t)$ be a fractional Brownian motion and ψ a wavelet that is $\mathbf{C}^2$. Prove that the average numbers, respectively, of zero-crossings and of modulus maxima of $WB_H(u, s)$ for $u \in [0, 1]$ are proportional to s. Verify this result numerically.

6.17 [2] Implement an algorithm that estimates the Lipschitz regularity α and the smoothing scale σ of sharp variation points in one-dimensional signals by applying the result of Theorem 6.7 on the dyadic wavelet transform maxima.