Recall that we are utilizing a real valued wavelet $\psi \in \mathbf{C}^n(\mathbb{R})$ with $n$ vanishing moments and with derivatives that have fast decay. We first remark that the proof of Theorem 31 can be adapted to give the following theorem, which measures the uniform Lipschitz regularity of $f$ over arbitrary intervals $[a, b]$.

**Theorem 32.** *If $f \in \mathbf{L}^2(\mathbb{R})$ is uniformly Lipschitz $\alpha \leq n$ over $[a, b]$, then there exists $A > 0$ such that*

$$|Wf(u, s)| \leq As^{\alpha+1/2}, \quad \forall\, (u, s) \in [a, b] \times (0, \infty) \tag{43}$$

*Conversely, suppose that $f$ is bounded and that $Wf(u, s)$ satisfies (43) for $\alpha < n$ that is not an integer. Then $f$ is uniformly Lipschitz $\alpha$ on $[a+\epsilon, b-\epsilon]$ for any $\epsilon > 0$.*

*Proof.* The proof relies on Theorem 31 and modifications of its proof. See pages 211–212 of *A Wavelet Tour of Signal Processing* for the details. □

We now make a few remarks. First, the condition (43) is only meaningful when $s \to 0$, since in general we have

$$|Wf(u, s)| = |\langle f, \psi_{u,s} \rangle| \leq \|f\|_2 \|\psi\|_2$$

which will supersede (43) for large $s$. Thus the localized regularity of $f$ is measured by zooming in on the points $u \in [a, b]$.

Second, if $\psi$ has exactly $n$ vanishing moments but $f$ is uniformly Lipschitz $\alpha > n$ on $[a, b]$, then $f \in \mathbf{C}^n(a, b)$ and we showed already in (34) that $\lim_{s\to 0} s^{-(n+1/2)} Wf(u, s) = Kf^{(n)}(u)$ with $K \neq 0$. Thus the wavelet coefficients will not decay as $O(s^{\alpha+1/2})$ despite the higher regularity of $f$.

Finally, for the converse of Theorems 31 and 32, there is the requirement that $\alpha \notin \mathbb{Z}$. Indeed, the wavelet decay conditions are not sufficient to conclude $\alpha$-Lipschitz regularity. In the case of $[a, b] = \mathbb{R}$, the decay (43) is only sufficient to conclude that $f$ is in the Zygmund class, which consists of all bounded, continuous functions for which there exists a constant $K$ such that

$$|f(t+v) + f(t-v) - 2f(t)| \leq K|v|, \quad \forall\, t, v \in \mathbb{R}$$

For more details, see [2, Chapter 6].

**Exercise 56.** Read Section 6.1.3 of *A Wavelet Tour of Signal Processing.*

## 5.2 Detection of Singularities via Wavelet Transform Modulus Maxima

*Section 6.2.1 of A Wavelet Tour of Signal Processing.*

Theorems 31 and 32 prove that the local regularity of $f$ at $v$ depends on the decay of $|Wf(u,s)|$ as $s \to 0$ in the neighborhood of $v$. However, it is not necessary to measure this decay in the entire time-scale plane $(u,s) \in \mathbb{R}\times(0,\infty)$. Rather, $|Wf(u,s)|$ can be controlled from its local maxima values.

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 left or right neighborhood of $u_0$ to avoid having any local maxima when $|Wf(u,s_0)|$ is constant. We call any connected curve $(u, s(u))$ in the time-scale plane along which all points are modulus maxima a maxima line. See Figure 23, which computes the wavelet modulus maxima of the signal from Figure 13 using a wavelet $\psi = -\theta'$.

Compare to the wavelet ridges used to analyze instantaneous frequencies, which were defined as the local maxima of the scalogram $P_W f(u,s) = |Wf(u,s)|^2$. In that case the ridge points followed the instantaneous frequencies of the signal over time; here the wavelet modulus maxima trace back to the isolated singularities at a fixed time.

Recall from Theorem 30 that if $\psi$ has exactly $n$ vanishing moments and a fast decay, then there exists $\theta$ with fast decay such that

$$\psi = (-1)^n \theta^{(n)}, \quad \widehat{\theta}(0) \neq 0$$

in which case the wavelet transform can be rewritten as

$$Wf(u,s) = s^n \frac{d^n}{du^n}(f * \widetilde{\theta}_s)(u)$$

If the wavelet $\psi$ has only one vanishing moment, then

$$Wf(u,s) = sf' * \widetilde{\theta}_s(u)$$

and so the wavelet modulus maxima are the maxima of the derivative of $f$ after being smoothed by $\widetilde{\theta}_s$. These maxima locate discontinuities and edges in

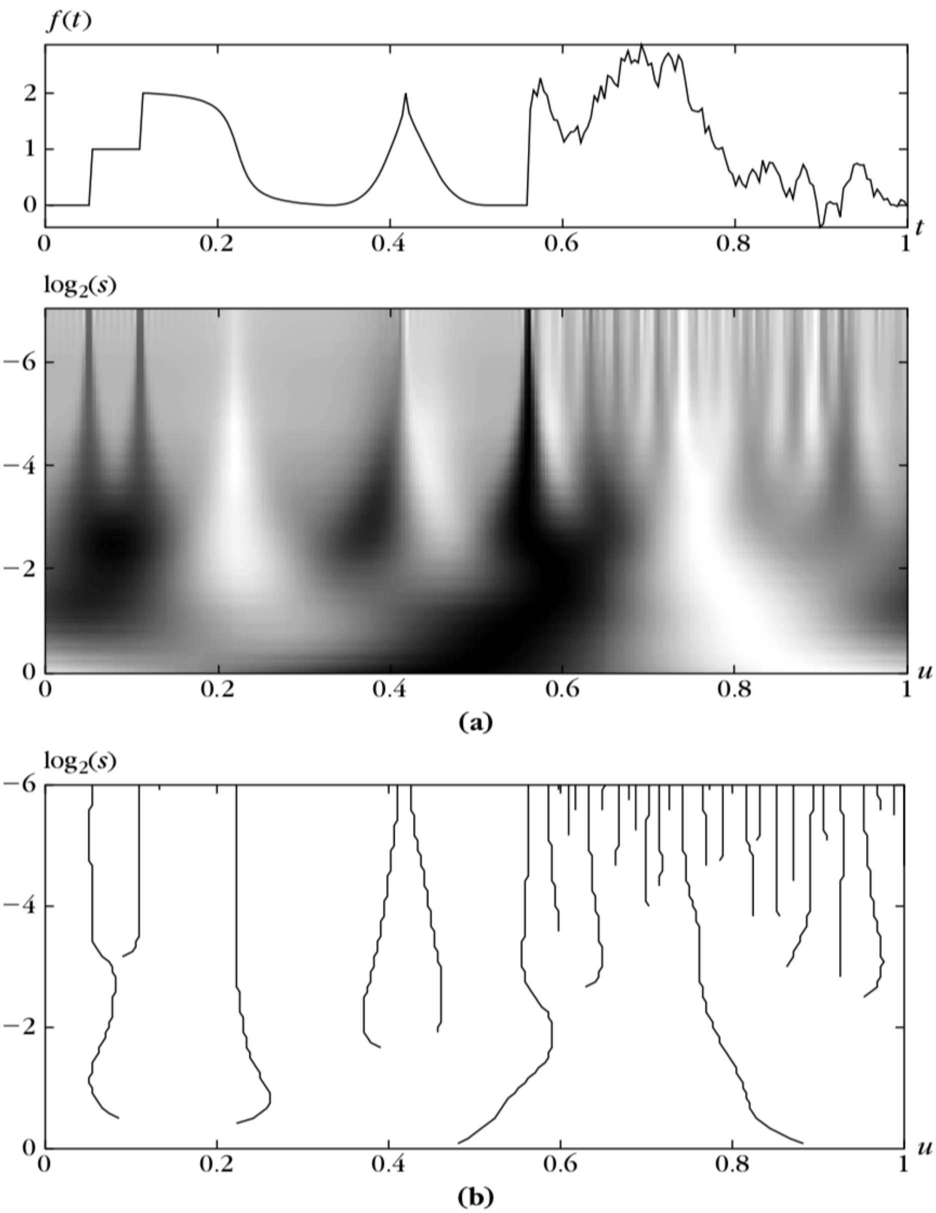


Figure 23: (a) Wavelet transform $Wf(u,s)$; the horizontal and vertical axes give $u$ and $\log_2 s$, respectively. (b) Modulus maxima of $Wf(u,s)$.

images. If the wavelet has two vanishing moments, then the wavelet modulus maxima correspond to high curvatures. See Figure 24 for an illustration. The next theorem proves that if $Wf(u,s)$ has no modulus maxima at fine scales, then $f$ is locally regular.

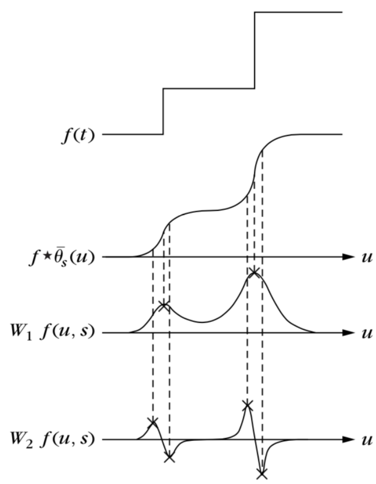


Figure 24: The convolution $f * \widetilde{\theta}_s(u)$ averages $f$ over a domain proportional to $s$. If $\psi = -\theta'$, then $W_1 f(u,s) = s\frac{d}{du}(f * \widetilde{\theta}_s)(u)$ has modulus maxima at sharp variation points of $f * \widetilde{\theta}_s(u)$. If $\psi = \theta''$, then the modulus maxima of $W_2 f(u,s) = s^2\frac{d^2}{du^2}(f * \widetilde{\theta}_s)(u)$ correspond to locally maximum curvatures.

**Theorem 33.** *Suppose that $\psi \in \mathbf{C}^n(\mathbb{R})$ with compact support, and $\psi = (-1)^n\theta^{(n)}$ with $\widehat{\theta}(0) \neq 0$. 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 $[a+\epsilon, b-\epsilon]$ for any $\epsilon > 0$.*

This theorem implies that $f$ can be singular at a point $v$ (i.e., $f$ is Lipschitz $\alpha < 1$ at $v$) only if there is a sequence of wavelet maxima points $(u_p, s_p)_{p\in\mathbb{N}}$ that converges to $v$ 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, however the theorem does guarantee that all singularities are detected by following the wavelet modulus maxima at fine scales. The maxima lines in Figure 23 illustrates the result.

Not all wavelets $\psi = (-1)^n\theta^{(n)}$ guarantee 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 $u = u_0$. This leads to spurious wavelet modulus maxima that make it harder to detect isolated singularities. This phenomena cannot occur, though, if $\theta$ is a Gaussian. Interestingly, in this case the wavelet transform can be written as the solution of the heat equation, where $s$ is proportional to the diffusion time, and the maximum principle proves that the maxima may not disappear.

We first recall the heat equation on the interval $[a, b]$ (which physically can be thought of as a rod). Let $g(x, \tau)$ be a function of space $x \in [a, b]$ and time $\tau \in [0, T]$, which measures the temperature of the rod at $x \in [a, b]$ at time $\tau$. We wish to find a $g$ that satisfies the *heat equation*,

$$\frac{\partial g}{\partial \tau} = \frac{\partial^2 g}{\partial x^2}$$
$$g(x, 0) = h(x)$$

for some initial temperature distribution given by $h(x)$. To make this problem well defined we need to specify the boundary condition of $g$ at $x = a, b$, which can be either Dirichlet, meaning that we enforce $g(a, \tau) = g(b, \tau) = 0$, or Neumann, meaning that $\partial_x g(a, \tau) = \partial_x g(b, \tau) = 0$. Since the theorem we are interested in, the maximum principle, holds regardless of boundary condition, we do not impose a specific one here.

For simplicity, suppose that $[a, b] = [0, 1]$, although this is not necessary. Define $R$ to be the space time rectangle

$$R = [0, 1] \times [0, T]$$

and define $B$ to be the boundary of $R$, not including the "top,"

$$B = \{(x, \tau) \in R : \tau = 0 \text{ or } x = 0 \text{ or } x = 1\}$$

The maximum principle proves that the maximum of $g(x, \tau)$ on $R$ must be attained somewhere on $B$.

**Theorem 34.** *If $g(x, \tau)$ satisfies the heat equation for $x \in [0, 1]$ and $\tau \in [0, T]$, then the maximum value of $g$ occurs at $\tau = 0$ (at the initial condition) or for $x = 0$ or $x = 1$ (at the ends of the rod). More precisely,*

$$\sup_{(x,\tau)\in R} g(x, \tau) = \sup_{(x,\tau)\in B} g(x, \tau)$$

*Proof.* Suppose that $(x_0, \tau_0)$ is an interior maximum point of $g$. In this case we must have

$$\partial_\tau g(x_0, \tau_0) = 0$$

Additionally, since maxima occur where the function is concave down, we must have that

$$\partial_{xx} g(x_0, \tau_0) \leq 0$$

The inequality cannot be strict, since $g$ is a solution of the heat equation and therefore

$$\partial_\tau g(x, \tau) = \partial_{xx} g(x, \tau), \quad \forall\, (x, \tau) \in R$$

In particular, it must be that $0 = \partial_\tau(x_0, \tau_0) = \partial_{xx} g(x_0, \tau_0)$.

Now define a new function

$$g_\varepsilon(x, \tau) = g(x, \tau) + \varepsilon x^2, \quad \varepsilon > 0$$

Since $g$ satisfies the heat equation it is easy to see that

$$\partial_\tau g_\varepsilon(x, \tau) - \partial_{xx} g_\varepsilon(x, \tau) = -2\varepsilon < 0 \tag{44}$$

We first show that $g_\varepsilon$ cannot achieve its maximum at an interior point $(x_0, \tau_0)$. Suppose that it does. By the same reasoning as above, $\partial_\tau g_\varepsilon(x_0, \tau_0) = 0$ and $\partial_{xx} g_\varepsilon(x_0, \tau_0) \leq 0$. But then

$$\partial_\tau g_\varepsilon(x_0, \tau_0) - \partial_{xx} g_\varepsilon(x_0, \tau_0) \geq 0$$

which contradicts (44).

Thus the maximum of $g_\varepsilon$ must occur on $\partial R$, the boundary of $R$. We now show that it cannot occur when $\tau = T$. Suppose that it does occur at a point $(x_0, T)$. Then again it must be that $\partial_{xx} g_\varepsilon(x_0, T) \leq 0$ and in this case, $\partial_\tau g_\varepsilon(x_0, T) \geq 0$ since

$$\partial_\tau g_\varepsilon(x_0, T) = \lim_{h \to 0^+} \frac{g_\varepsilon(x_0, T) - g_\varepsilon(x_0, T - h)}{h} \geq 0$$

where the inequality follows from the assumption that $(x_0, T)$ is the location of the maximum, which implies that $g_\varepsilon(x_0, T) \geq g_\varepsilon(x_0, T - h)$. But this would again imply that $\partial_\tau g_\varepsilon(x_0, T) - \partial_{xx} g_\varepsilon(x_0, T) \geq 0$, contradicting (44).

Therefore we must have

$$\begin{aligned}\sup_{(x,\tau)\in R} g_\varepsilon(x,\tau) = \sup_{(x,\tau)\in B} g_\varepsilon(x,\tau) &= \sup_{(x,\tau)\in B} [g(x,\tau)+\varepsilon x^2] \\ &\leq \sup_{(x,\tau)\in B} g(x,\tau) + \sup_{(x,\tau)\in B} \varepsilon x^2 \\ &= \sup_{(x,\tau)\in B} g(x,\tau) + \varepsilon\end{aligned}$$

On the other hand we also have:

$$\sup_{(x,\tau)\in R} g_\varepsilon(x,\tau) = \sup_{(x,\tau)\in R} [g(x,\tau)+\varepsilon x^2] \geq \sup_{(x,\tau)\in R} g(x,\tau)$$

It follows that

$$\sup_{(x,\tau)\in R} g(x,\tau) \leq \sup_{(x,\tau)\in B} g(x,\tau) + \varepsilon, \quad \forall\, \varepsilon > 0$$

Letting $\varepsilon \to 0$, we obtain

$$\sup_{(x,\tau)\in R} g(x,\tau) \leq \sup_{(x,\tau)\in B} g(x,\tau)$$

But since $g$ is continuous and $B \subset R$, we must have equality. □

Now we can prove the following theorem:

**Theorem 35.** *Let $\psi = (-1)^n \theta^{(n)}$, where $\theta$ 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 as the scale decreases.*

*Proof.* To simplify the proof we use a normalized Gaussian

$$\theta(t) = \frac{1}{2\sqrt{\pi}} e^{-t^2/4} \Longrightarrow \widehat{\theta}(\omega) = e^{-\omega^2}$$

Note as well that $\widetilde{\theta}_s = \theta_s$. Let $f^{(n)}$ be the $n^{\text{th}}$ derivative of $f$, which is defined in the sense of distributions if $f^{(n)}$ is not defined otherwise. Theorem 30 proves that

$$Wf(u,s) = s^n \frac{d^n}{du^n}(f * \theta_s)(u) = s^n f^{(n)} * \theta_s(u)$$

Consider the heat equation

$$\partial_\tau g(u,\tau) = \partial_{xx} g(u,\tau), \quad g(u,0) = h(u)$$

We can compute the solution by taking the Fourier transform of both sides with respect to $u$:

$$\partial_\tau \widehat{g}(\omega,\tau) = -\omega^2 \widehat{g}(\omega,\tau)$$

It follows that

$$\widehat{g}(\omega,\tau) = \widehat{h}(\omega) e^{-\tau\omega^2}$$

from which we derive:

$$g(u,\tau) = \frac{1}{\sqrt{\tau}} h * \theta_{\sqrt{\tau}}(u)$$

If we set $\tau = s^2$ and $h = f^{(n)}$, then we obtain

$$s^{n+1} g(u,s^2) = s^n f^{(n)} * \theta_s(u) = Wf(u,s) \tag{45}$$

Thus the wavelet transform is proportional to the solution of the heat equation with initial temperature distribution given by $f^{(n)}$.

Recall that a modulus maxima of $Wf(u,s)$ at $(u_0,s_0)$ is a local maxima of $|Wf(u,s)|$ for a fixed $s$ and varying $u$, which due to (45) corresponds to a local maxima of $|g(u,s^2)|$ for a fixed $s$ and varying $u$. Suppose that a curve of wavelet modulus maxima is interrupted at $(u_1,s_1)$ with $s_1 > 0$. Then one can show there exists $\varepsilon > 0$ such that on the domain $[u_1-\varepsilon, u_1+\varepsilon] \times [s_1-\varepsilon, s_1]$, the global maximum of $|g(u,s^2)|$ is at $(u_1,s_1)$. However, this contradicts the maximum principle, and so the curve cannot be interrupted. □

**Exercise 57.** Let $f(t) = \cos(\omega_0 t)$ and $\psi(t)$ a wavelet that is symmetric about zero.

(a) Verify that $Wf(u,s) = \sqrt{s}\widehat{\psi}(s\omega_0)\cos(\omega_0 t)$.

(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 $\psi$.

**Exercise 58** (20 points)**.** Prove that if $f(t) = \mathbf{1}_{[0,+\infty)}(t)$ 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 $\psi$.

# References

[1] Stéphane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd edition, 2008.

[2] Yves Meyer. *Wavelets and Operators*, volume 1. Cambridge University Press, 1993.

[3] Elias M. Stein and Rami Shakarchi. *Fourier Analysis: An Introduction*. Princeton Lectures in Analysis. Princeton University Press, 2003.

[4] Richard L. Wheeden and Antoni Zygmund. *Measure and Integral: An Introduction to Real Analysis*. Marcel Dekker, 1977.

[5] Elias M. Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, 1971.

[6] John J. Benedetto and Matthew Dellatorre. Uncertainty principles and weighted norm inequalities. *Contemporary Mathematics*, 693:55–78, 2017.

[7] Karlheinz Gröchenig. *Foundations of Time Frequency Analysis*. Springer Birkhäuser, 2001.