Computational Harmonic Analysis (Wavelet Tutorial) Part II

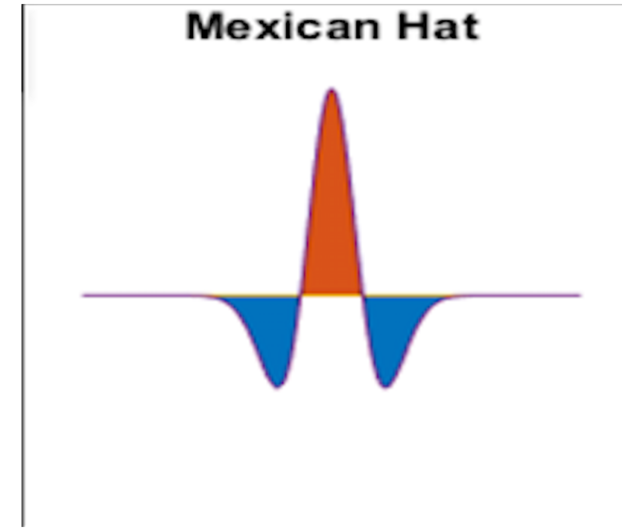
Understanding Many Particle Systems with Machine Learning

Tutorials

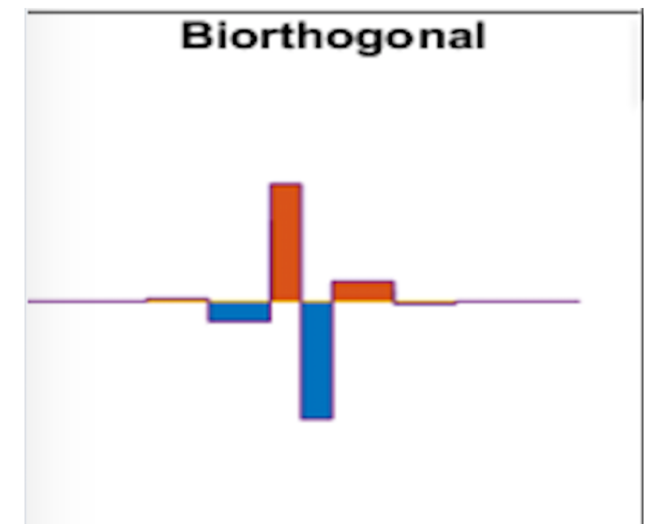
Matthew Hirn
Michigan State University
Department of Computational Mathematics, Science & Engineering
Department of Mathematics

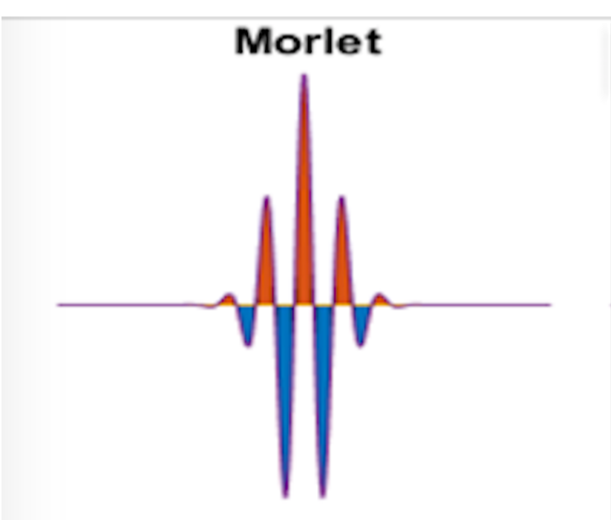
Wavelet Transform

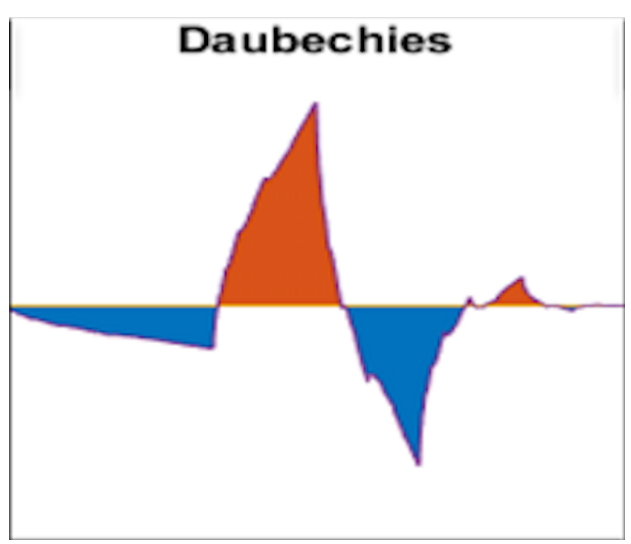
Wavelets

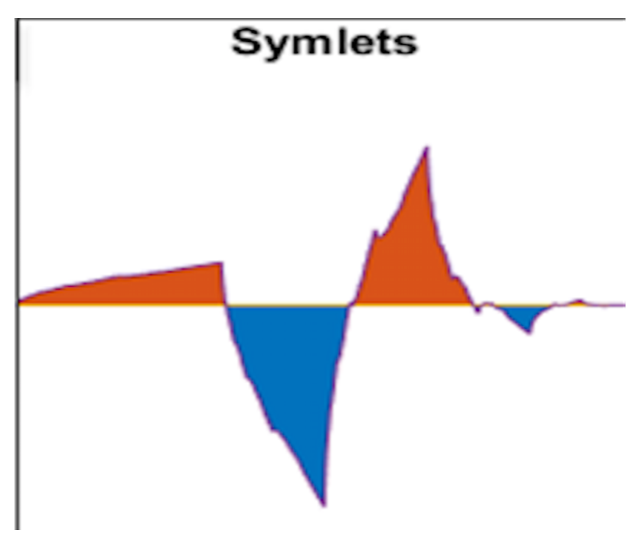


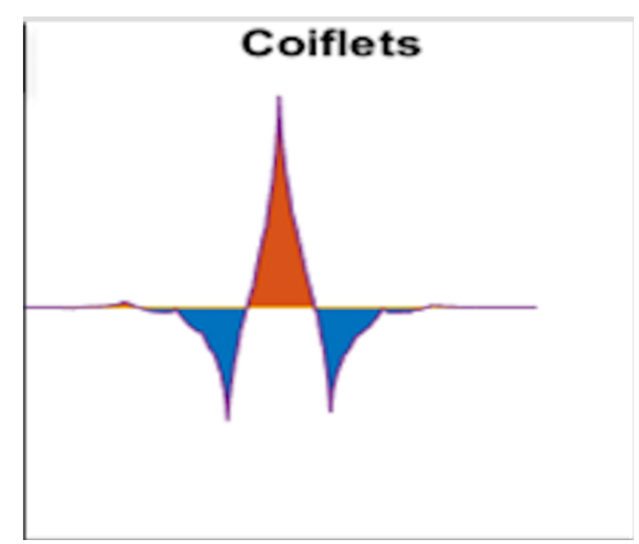
- Wavelet $\psi \in \mathbf{L}^2(\mathbb{R})$ satisfies:
  - Zero average: $\int \psi = 0$
  - Normalized: $\|\psi\|_2 = 1$
  - Centered around $t = 0$
  - Localized in time and frequency
  - Can be either real or complex valued











Wavelet Transform

- Wavelet dictionary obtained by scaling and translating ψ:

$$\mathcal{D} = \{\psi_{u,s}\}_{u \in \mathbb{R}, s \in \mathbb{R}+}, \quad \psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$

- Wavelet transform:

$$\begin{aligned} Wf(u,s) &= \langle f, \psi_{u,s} \rangle \\ &= \int_{-\infty}^{+\infty} f(t) s^{-1/2} \overline{\psi(s^{-1}(t-u))}\, dt \\ &= f * \tilde{\psi}_s(u) \end{aligned}$$

where

$$\tilde{\psi}_s(t) = s^{-1/2} \overline{\psi(s^{-1}t)}$$

- Note:

$$\widehat{\tilde{\psi}}_s(\omega) = \sqrt{s}\, \overline{\widehat{\psi}(s\omega)}$$

Thus, since:

$$\widehat{f * \tilde{\psi}_s}(\omega) = \widehat{f}(\omega) \widehat{\tilde{\psi}}_s(\omega)$$

the wavelet transform $Wf(u,s)$ captures the frequency information of f organized by the frequency bands of $\tilde{\psi}_s$.

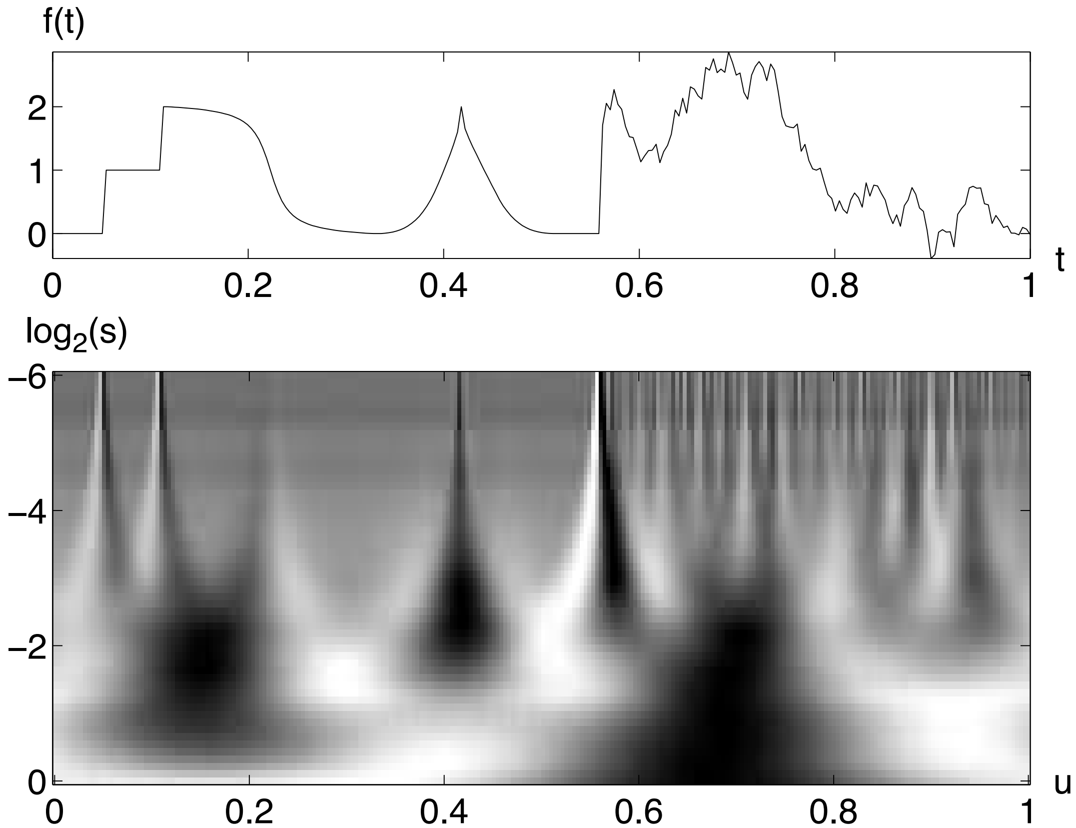


Fig. 4.7. A Wavelet Tour of Signal Processing, 3rd ed. Real wavelet transform $Wf(u,s)$ computed with a Mexican hat wavelet The vertical axis represents $\log_2 s$. Black, grey and white points correspond respectively to positive, zero and negative wavelet coefficients.

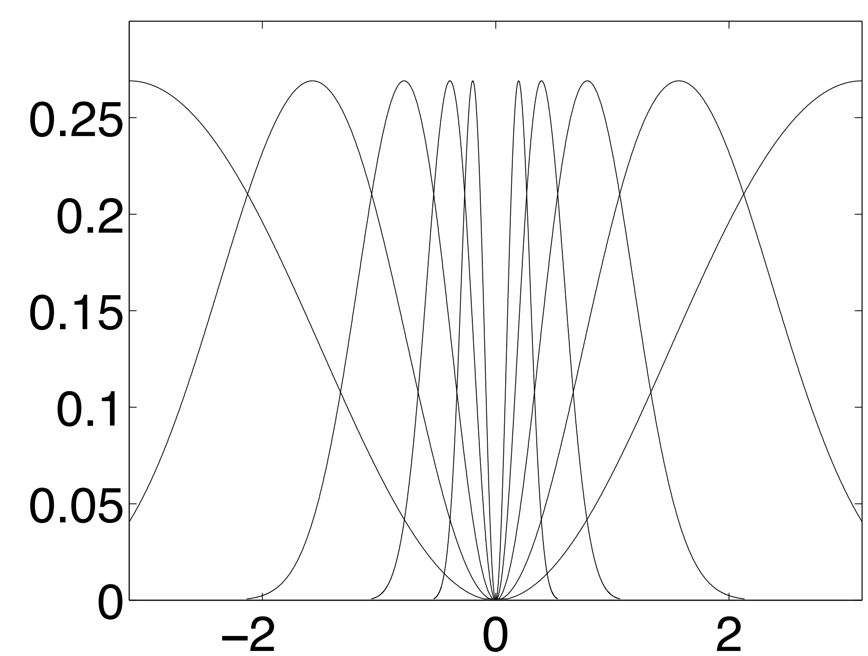


Fig. 5.1. A Wavelet Tour of Signal Processing, 3rd ed. Scaled Fourier transforms $|\hat{\psi}(2^j\omega)|^2$, for $1 \leqslant j \leqslant 5$ and $\omega \in [-\pi, \pi]$.

Real Wavelet Reconstruction

- Theorem (Calderón, Grossman and Morlet): Let $\psi \in \mathbf{L}^2(\mathbb{R})$ be a real function such that

  $$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega}\, d\omega < +\infty$$

  Then, for any $f \in \mathbf{L}^2(\mathbb{R})$:

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u,s) s^{-1/2} \psi(s^{-1}(t-u))\, du\, \frac{ds}{s^2}$$

$$\|f\|_2^2 = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} |Wf(u,s)|^2\, du\, \frac{ds}{s^2}.$$

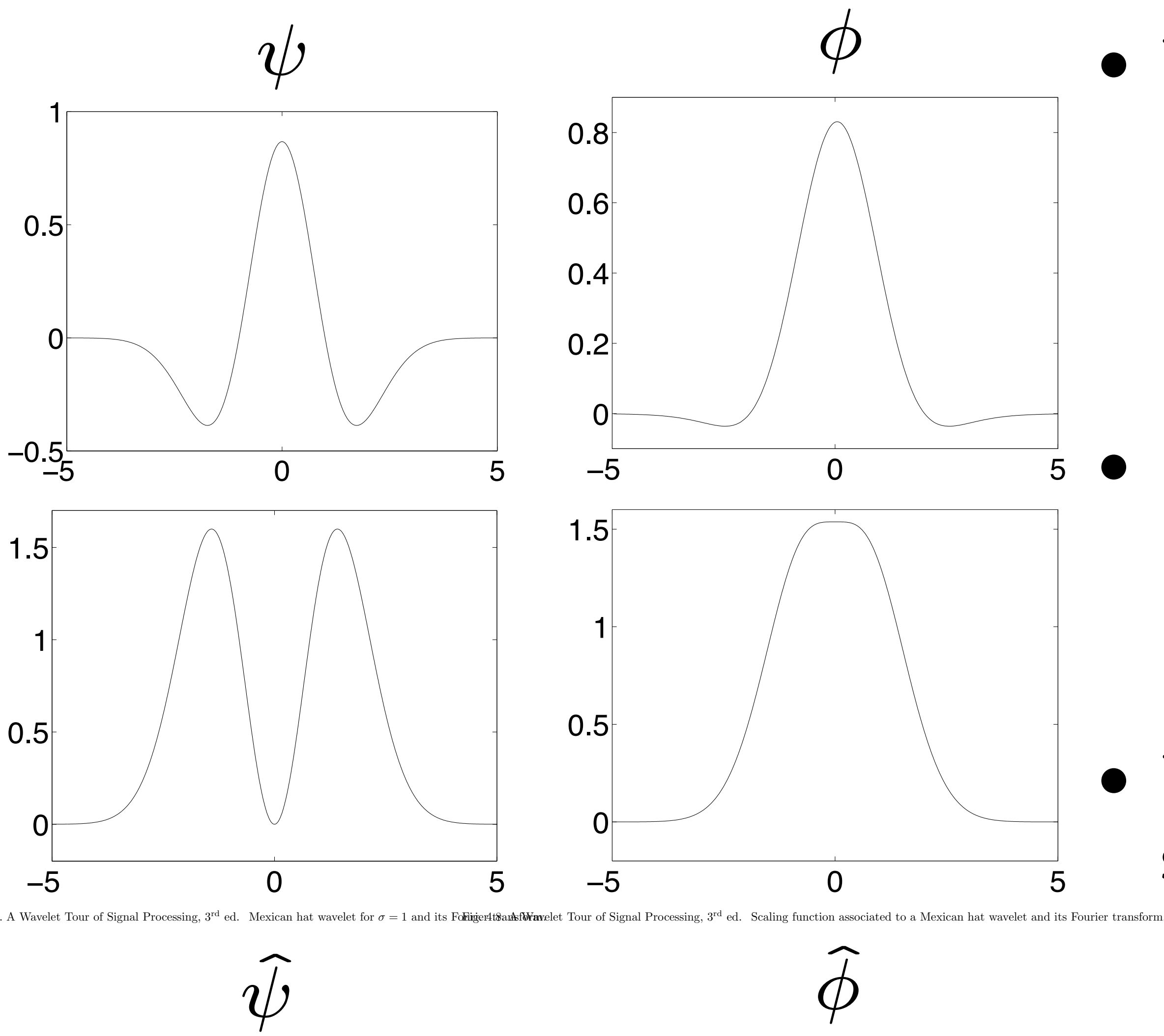
- $C_\psi < \infty$ is called the wavelet admissibility condition.

- $C_\psi < +\infty \Rightarrow \hat{\psi}(0) = 0$. This is almost sufficient.

- If additionally, $\hat{\psi} \in \mathbf{C}^1$, then $C_\psi < +\infty$. Can insure this with sufficient time decay:

  $$|\psi(t)| \leq \frac{K}{1 + |t|^{2+\epsilon}}$$

Scaling Function



- Numerically the wavelet transform is only computed up to scales $s < s_0$, which loses the low frequency information of f.

- The scaling function ϕ captures this information. Defined by:

$$|\hat{\phi}(\omega)|^2 = \int_1^{+\infty} |\hat{\psi}(s\omega)|^2 \frac{ds}{s}$$

- Denote:

$$\phi_s(t) = \frac{1}{\sqrt{s}}\phi\left(\frac{t}{s}\right) \text{ and } \tilde{\phi}_s(t) = \overline{\phi_s(-t)}$$

- The low frequency approximation of f at scale s is:

$$Af(u,s) = \langle f, \phi_{u,s} \rangle = f * \tilde{\phi}_s(u)$$

- Reconstruction still holds:

$$f(t) = \frac{1}{C_\psi}\int_0^{s_0} Wf(\cdot,s)*\psi_s(t)\,\frac{ds}{s^2} + \frac{1}{C_\psi s_0} Af(\cdot,s_0)*\phi_{s_0}(t)$$

Analytic Wavelets

- Complex valued, analytic wavelets admit a time-frequency analysis, like the windowed Fourier transform.
- The wavelet ψ is analytic if:

$$\forall\, \omega < 0, \quad \widehat{\psi}(\omega) = 0$$

- The wavelet transform $Wf(u,s)$ of an analytic wavelet satisfies very similar reconstruction and energy preservation formulas as the real wavelet transform.

Analytic Wavelet Construction

- Let g be a real, symmetric window.
- Define a wavelet ψ as:

$$\psi(t) = g(t)e^{i\eta t} \Rightarrow \widehat{\psi}(\omega) = \widehat{g}(\omega - \eta)$$

- Thus if $\widehat{g}(\omega) = 0$ for $|\omega| > \eta$, then $\widehat{\psi}(\omega) = 0$ for $\omega < 0$, and ψ is analytic.
- ψ is centered in time at $t = 0$ and in frequency at $\omega = \eta$.
- Gabor wavelets use a Gaussian window, and so are not strictly analytic and do not have precisely zero average. However $\widehat{\psi}(\omega) \approx 0$ for $\omega \leq 0$.
- Morlet wavelets also use a Gaussian window, but subtract a constant in order to have zero average:

$$\psi(t) = g(t)(e^{i\eta t} - C)$$

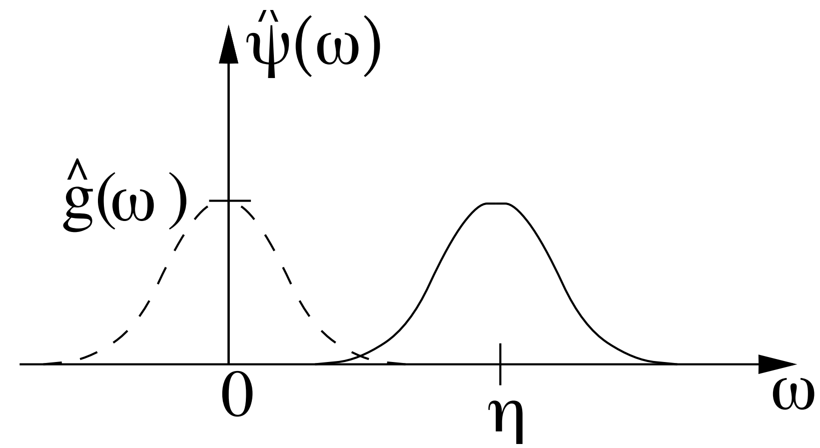
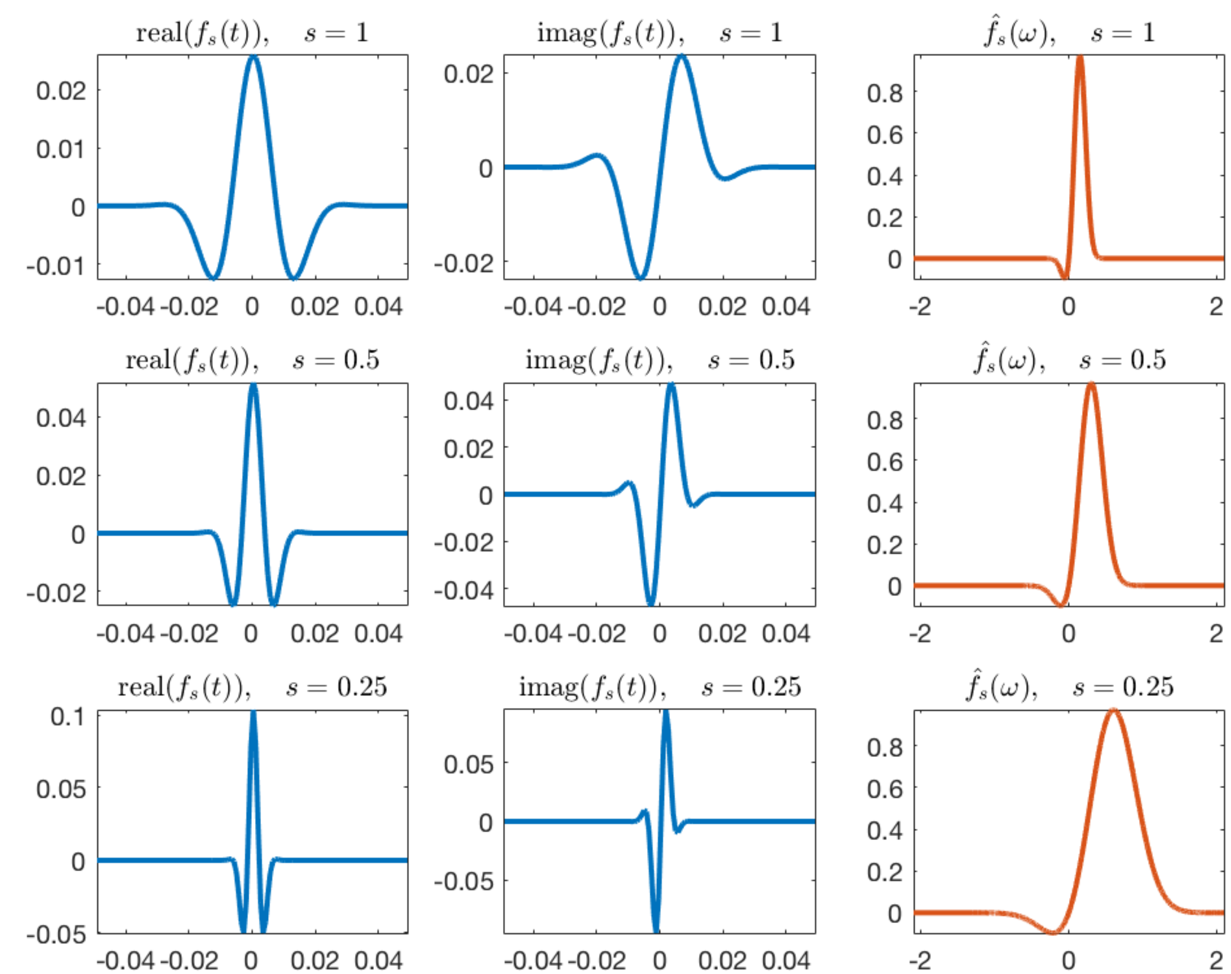


Fig. 4.10. A Wavelet Tour of Signal Processing, 3rd ed. Fourier transform $\hat{\psi}(\omega)$ of a wavelet $\psi(t) = g(t)\exp(i\eta t)$.



Analytic Wavelet Heisenberg Boxes

- Suppose ψ is centered at $t = 0$ with central frequency $\omega = \eta$.

- The time variance σ_t^2 and frequency variance σ_ω^2 of ψ are:

$$\sigma_t^2 = \int_{-\infty}^{+\infty} t^2 |\psi(t)|^2 \, dt$$

$$\sigma_\omega^2 = \frac{1}{2\pi} \int_0^{+\infty} (\omega - \eta)^2 |\hat{\psi}(\omega)|^2 \, d\omega$$

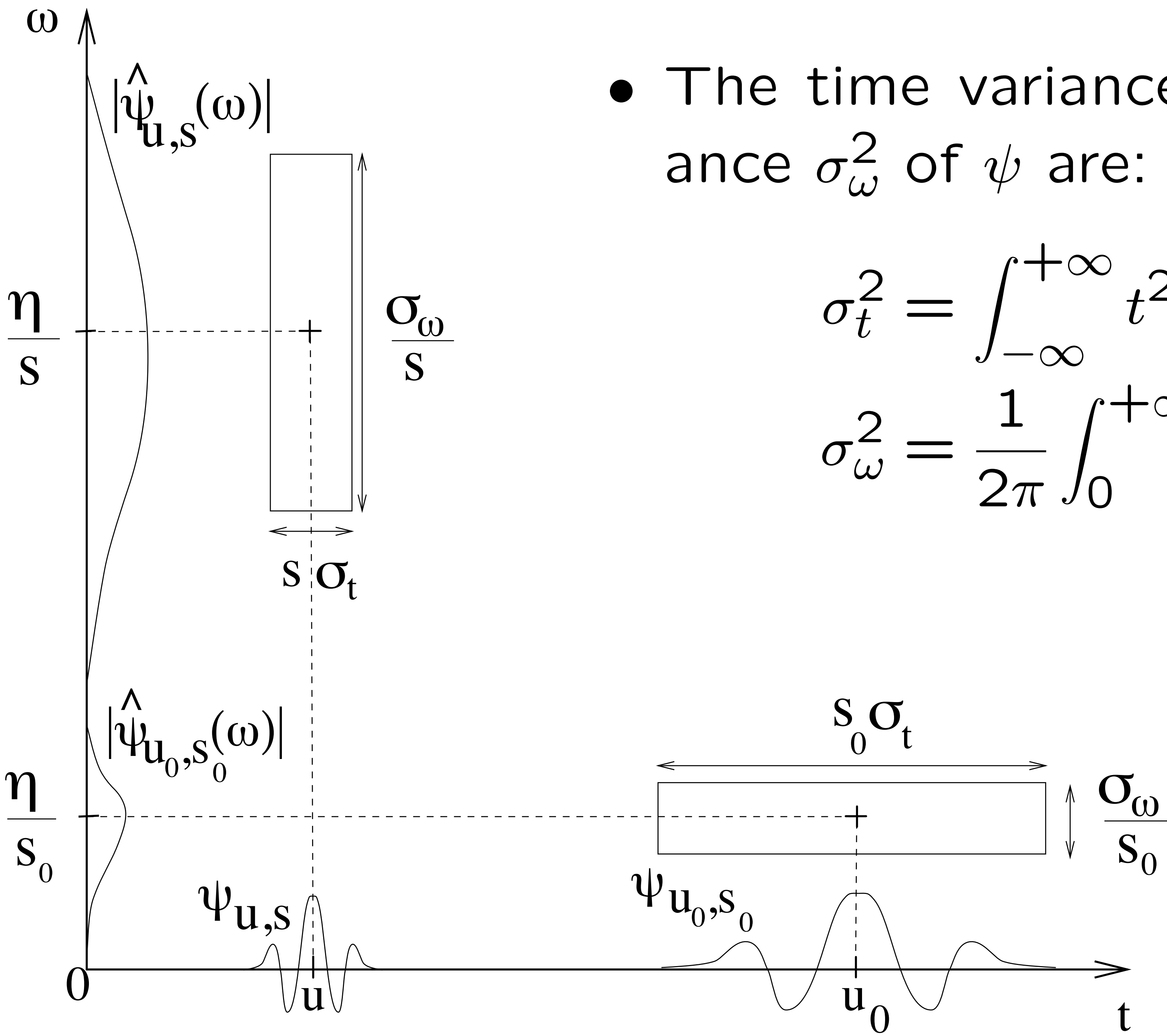


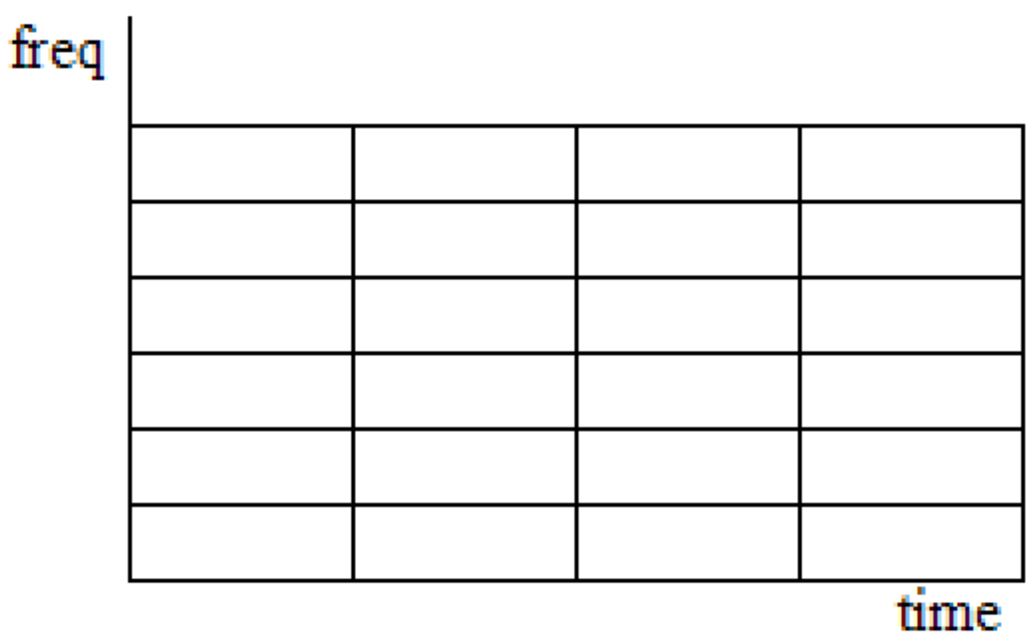
Fig. 4.9. A Wavelet Tour of Signal Processing, 3rd ed. Heisenberg boxes of two wavelets. Smaller scales decrease the time spread but increase the frequency support, which is shifted towards higher frequencies.

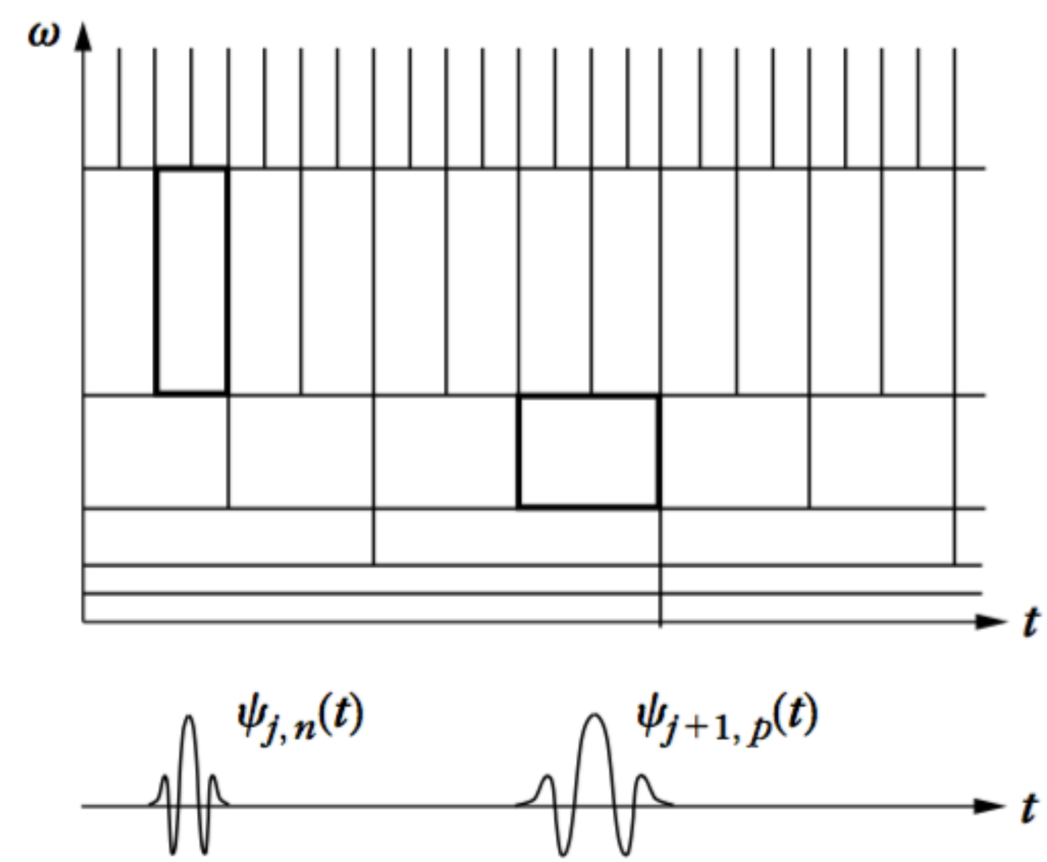
- Scalogram:

$$P_W f(u, \eta/s) = |W f(u, s)|^2$$

Time-Frequency Plane: Wavelets vs. Windowed Fourier

Comparison of time-frequency tilings:

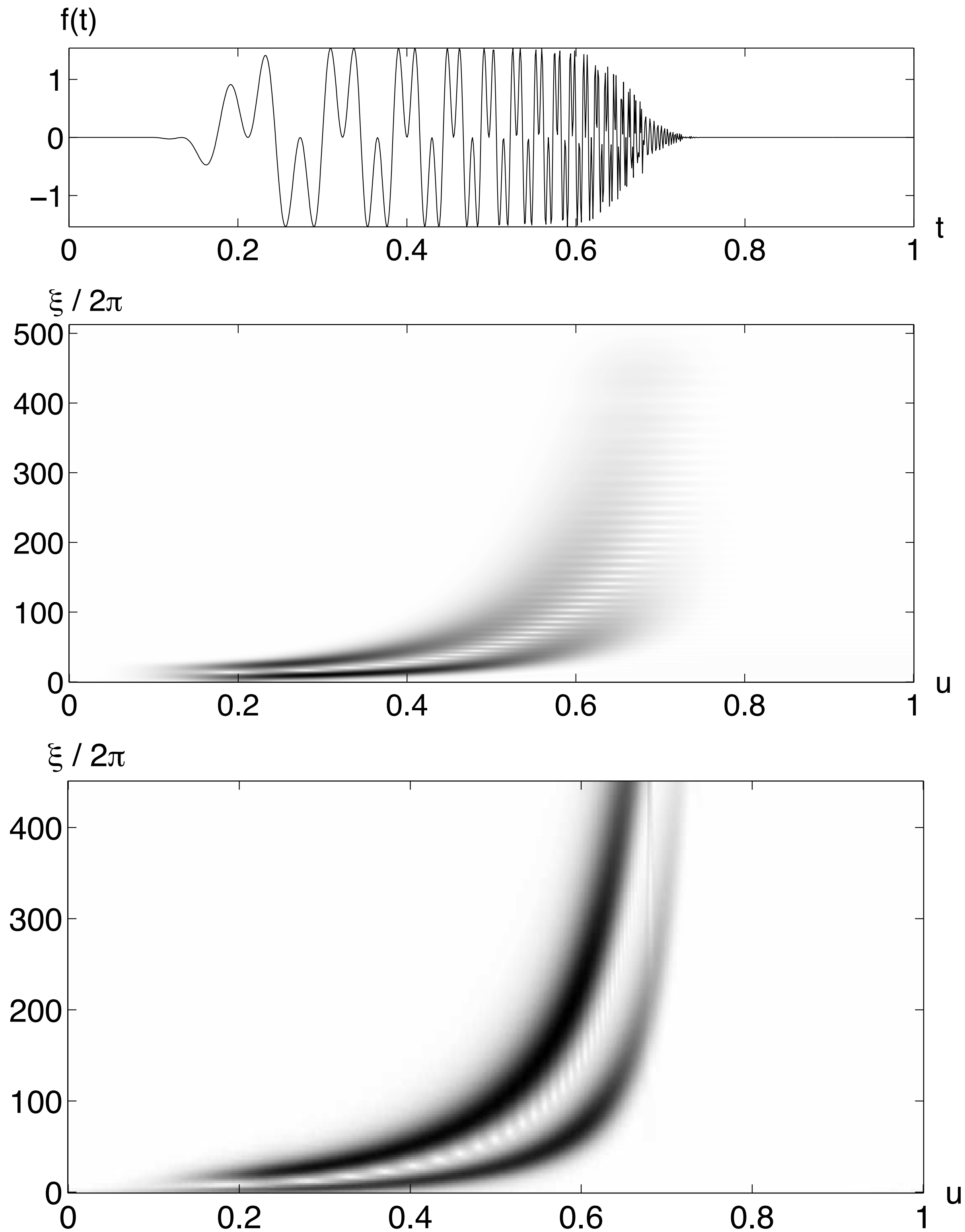




Windowed Fourier Transform

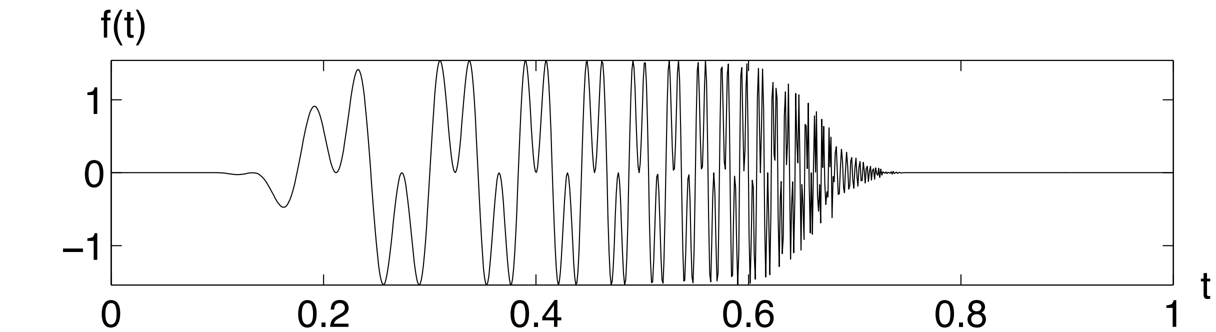
Wavelet Transform

Hyperbolic Chirp Revisited

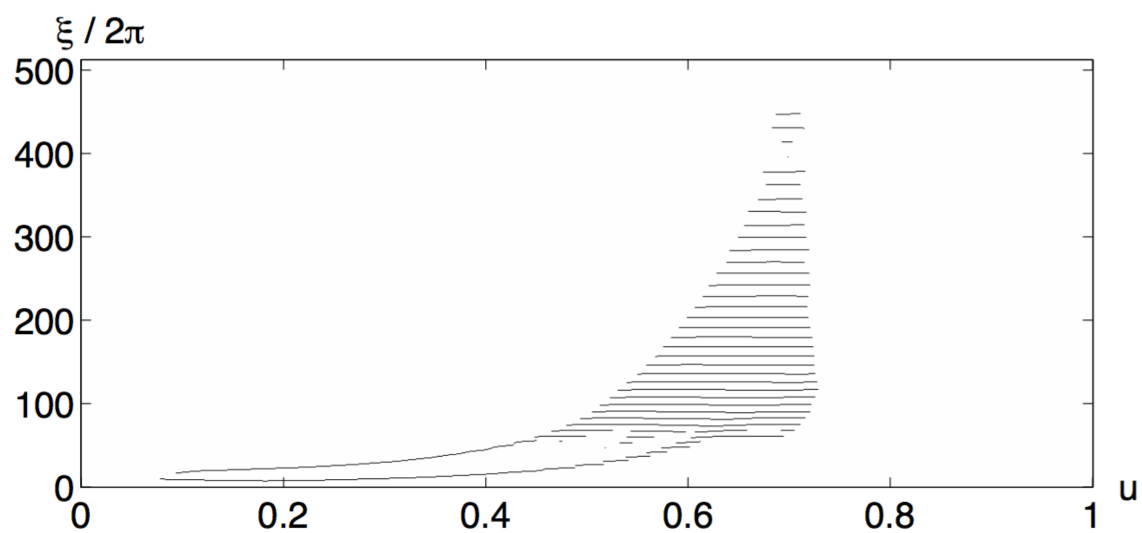


- $f(t) = a_1 \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + a_2 \cos\left(\frac{\alpha_2}{\beta_2 - t}\right)$
- Spectrogram $P_S f(u, \xi)$ of windowed Fourier transform
- Scalogram $P_W f(u, \eta/s)$ of analytic wavelet transform

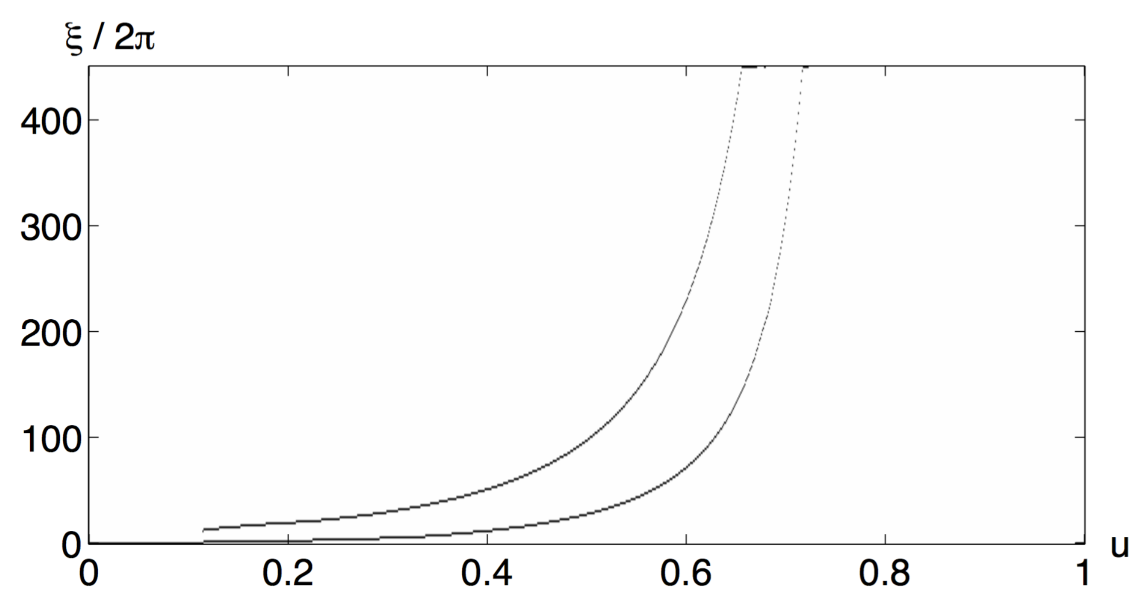
Hyperbolic Chirp Revisited



- $f(t) = a_1 \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + a_2 \cos\left(\frac{\alpha_2}{\beta_2 - t}\right)$



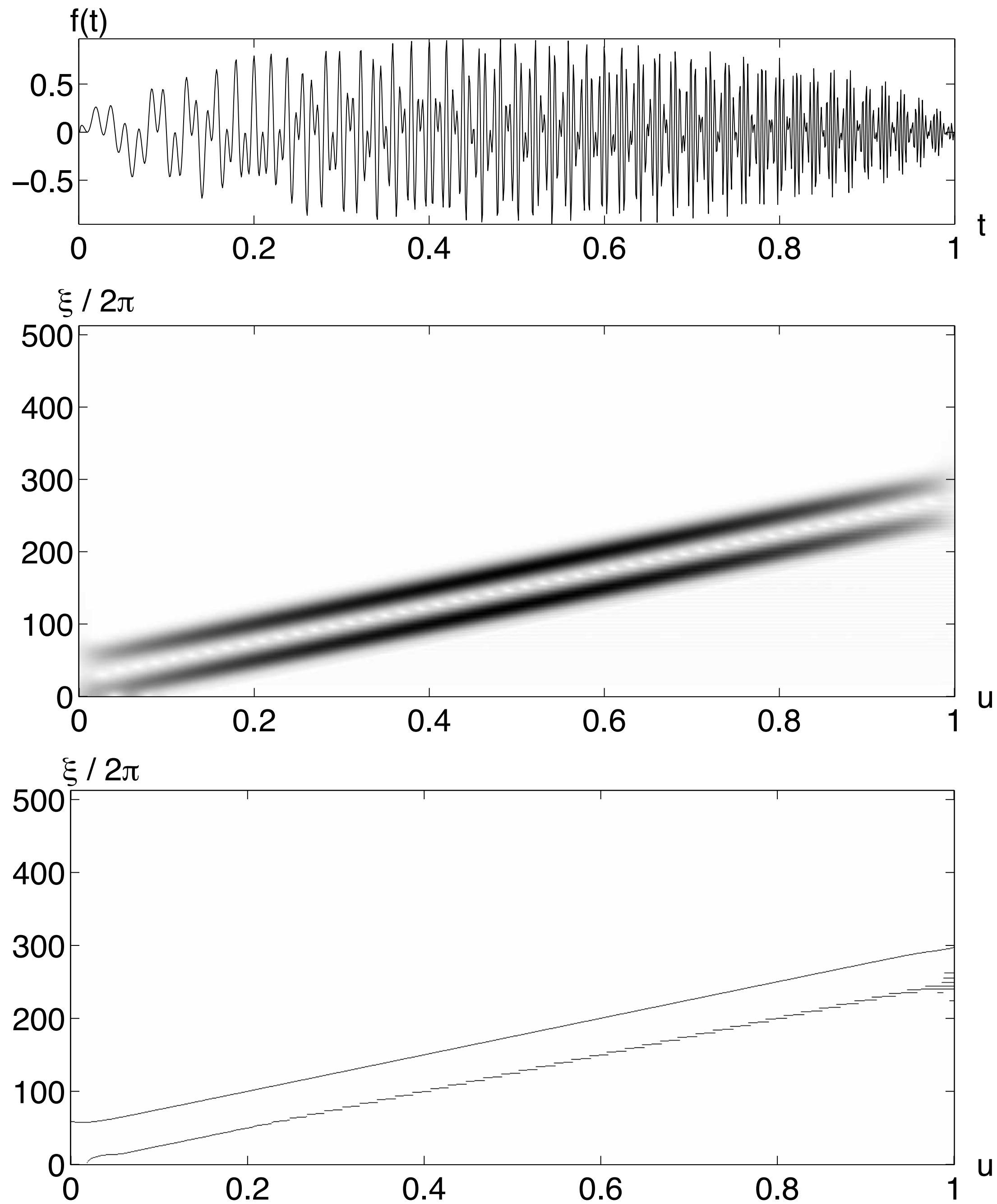
- Local maxima of spectrogram $P_S f(u, \xi)$



- Local maxima of scalogram $P_W f(u, \eta/s)$

Parallel Linear Chirps

- $f(t) = a_1 \cos(bt^2 + ct) + a_2 \cos(bt^2)$



velet Tour of Signal Processing, 3rd ed. Sum of two parallel linear chirps. (a): Spectrogram $P_S f(u,\xi) = |Sf(u,\xi)|^2$. (b): Ridge support calculated from the spectrogram.

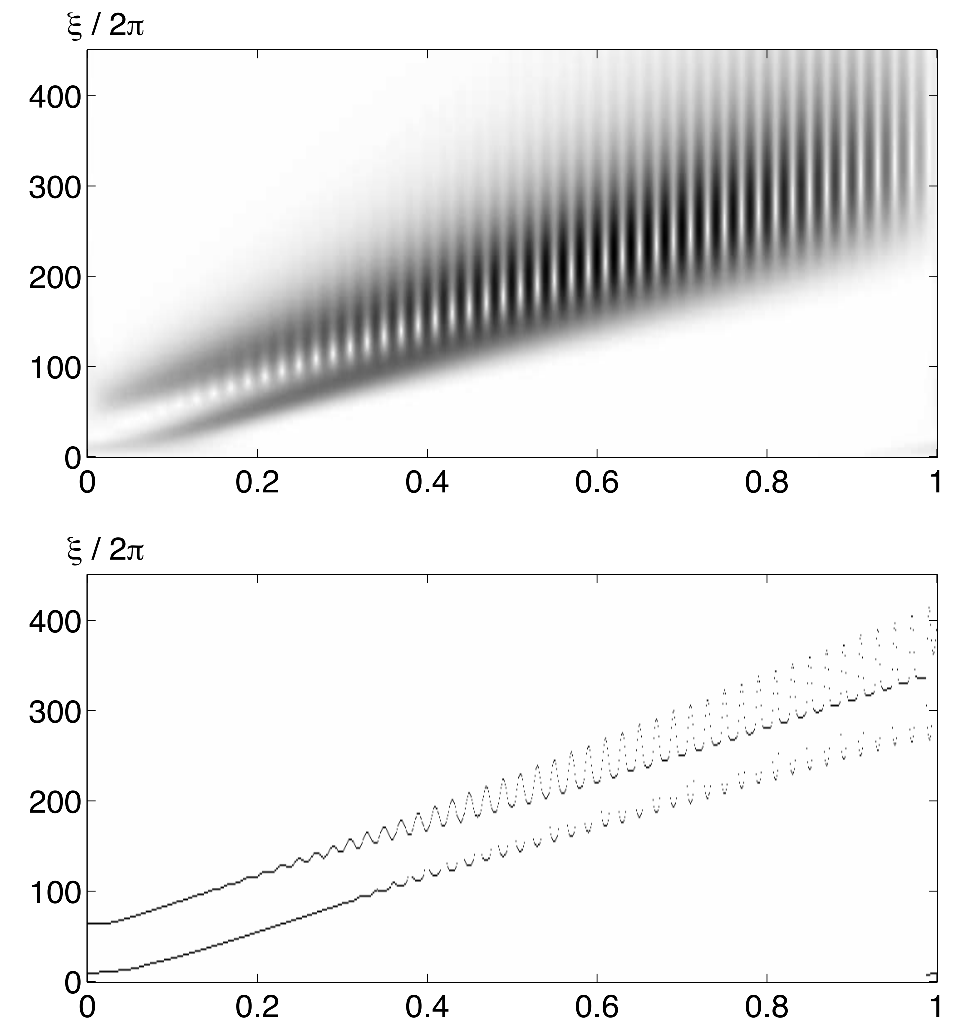


Fig. 4.16. A Wavelet Tour of Signal Processing, 3rd ed. (a): Normalized scalogram $\eta^{-1}\xi P_W f(u,\xi)$ of two parallel linear chirps. (b): Wavelet ridges.

- Spectrogram: $P_S f(u,\xi)$
- Scalogram: $P_W f(u,\eta/s)$

Sparsity and Time-Frequency Resolution

- Lesson: Best transform depends on the signal f time-frequency properties.
- A transform that is adapted to the signal time-frequency property has fewer local maxima, and is thus sparser.
- Transforms that are not adapted to the signal diffuse the signal's energy over many atoms, leading to more local maxima and a less sparse representation.
- Thus sparsity is a natural criterion to guide the construction of time-frequency transforms.

Wavelet Zoom

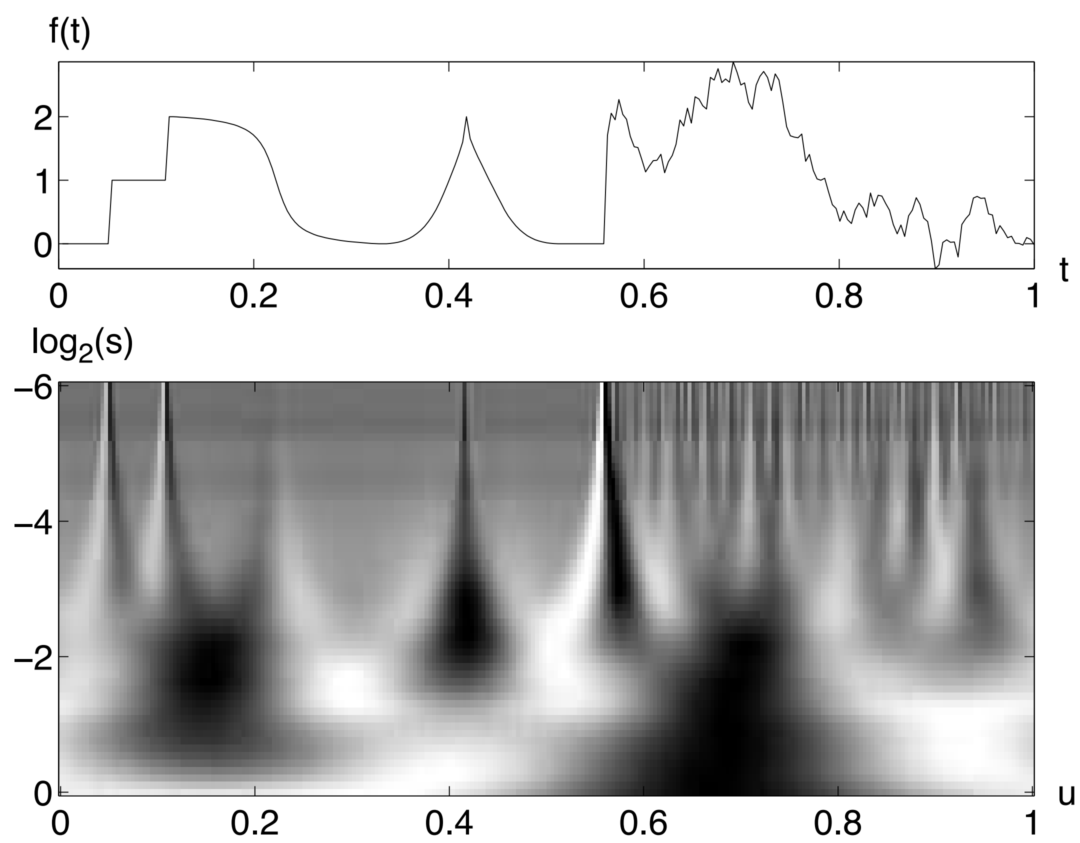


Fig. 4.7. A Wavelet Tour of Signal Processing, 3[rd] ed. Real wavelet transform $Wf(u,s)$ computed with a Mexican hat wavelet The vertical axis represents $\log_2 s$. Black, grey and white points correspond respectively to positive, zero and negative wavelet coefficients.

Taylor's Theorem

- We now turn to measuring the local regularity of f at a point v.

- Suppose f is m times differentiable in $[v-h, v+h]$.

- Let p_v be the Taylor polynomial of f in the neighborhood of v:

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

- Taylor's Theorem: The residual $\varepsilon_v(t) = f(t) - p_v(t)$ satisfies $\forall\, t \in [v-h, v+h]$:

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

Lipschitz Regularity

- Lipschitz Regularity: A function f is point wise Lipschitz (Hölder) $\alpha \geq 0$ at v, if there exists $K > 0$ and a polynomial p_v of degree $m = \lfloor \alpha \rfloor$ such that

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

- f is uniformly Lipschitz α over $[a, b]$ if it satisfies the above for all $v \in [a, b]$ with a K independent of v.

- Global Lipschitz regularity and the Fourier transform: A function f is bounded and uniformly Lipschitz α over $\mathbb{R}$ if:

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

Wavelet Vanishing Moments

- A wavelet has n vanishing moments if:

$$\forall\, 0 \le k < n, \quad \int_{-\infty}^{+\infty} t^k \psi(t)\, dt = 0$$

- Wavelet transform kills polynomials p with $\deg(p) \le n-1$: $Wp(u,s) = 0$.

- Let f be Lipschitz $\alpha < n$ at v, so that:

  $f(t) = p_v(t) + \varepsilon_v(t)$ with $|\varepsilon_v(t)| \le K|t-v|^\alpha$

  Then:

$$Wf(u,s) = W\varepsilon_v(u,s)$$

- We are going to measure α from $|Wf(u,s)|$, with u close to v.

Multiscale Differential Operator

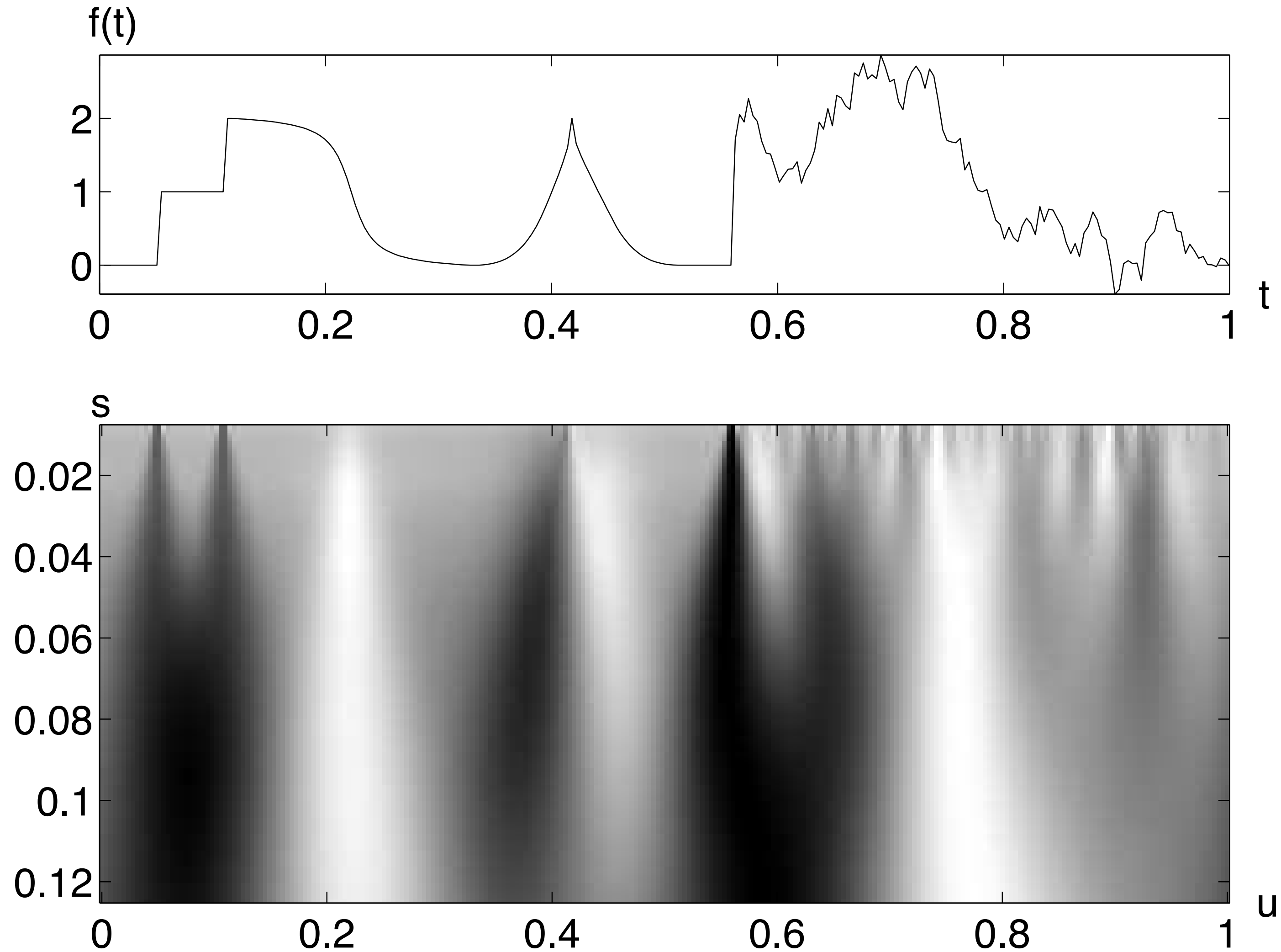


Fig. 6.1. A Wavelet Tour of Signal Processing, 3rd ed. Wavelet transform $Wf(u,s)$ calculated with $\psi = -\theta'$ where θ is a Gaussian, for the signal f shown above. The position parameter u and the scale s vary respectively along the horizontal and vertical axes. Black, grey and white points correspond respectively to positive, zero and negative wavelet coefficients. Singularities create large amplitude coefficients in their cone of influence.

- Wavelet transform $Wf(u,s)$ with wavelet ψ with one vanishing moment
  - Black: positive
  - White: negative
  - Grey: zero
- Singularities create large amplitude wavelet coefficients
- Notice that the coefficients give information regarding the derivative of f - this is not an accident!

Multiscale Differential Operator

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

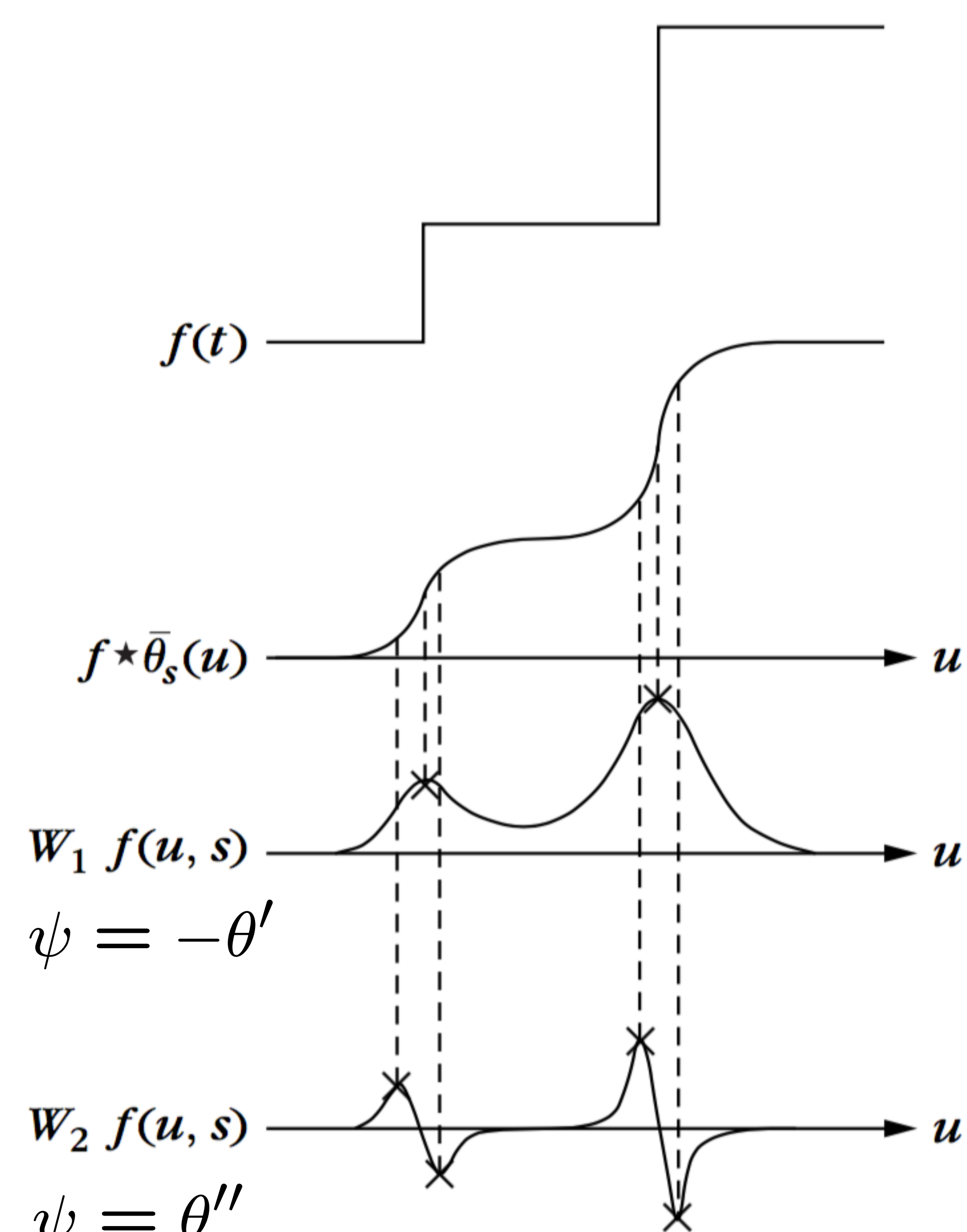
  $$\psi(t) = (-1)^n \frac{d^n \theta(t)}{dt^n}$$

  As a consequence:

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

  with

  $$\tilde{\theta}_s(t) = s^{-1/2}\theta(-t/s)$$



Wavelet Zoom on an Interval

- Let $\psi \in \mathbf{C}^n(\mathbb{R})$ have n vanishing moments and derivatives that have fast decay.

- Theorem:

  – If $f \in \mathbf{L}^2(\mathbb{R})$ is uniformly Lipschitz $\alpha \leq n$ over $[a, b]$, then there exists $A > 0$ such that $\forall\,(u, s) \in [a, b] \times \mathbb{R}^+$,

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

  – Conversely, suppose f is bounded and $|Wf(u, s)| \leq A s^{\alpha+1/2}\ \forall\,(u, s) \in [a, b] \times \mathbb{R}^+$ for an $\alpha < n$, $\alpha \notin \mathbb{Z}$. Then f is uniformly Lipschitz α on $[a + \epsilon, b - \epsilon]$ for any $\epsilon > 0$.

Wavelet Zoom at a Point

- Let $\psi \in \mathbf{C}^n(\mathbb{R})$ have n vanishing moments and derivatives that have fast decay.
- Theorem (Jaffard):
  - If $f \in \mathbf{L}^2(\mathbb{R})$ is Lipschitz $\alpha \leq n$ at v, then there exists $A > 0$ such that $\forall\,(u,s) \in \mathbb{R} \times \mathbb{R}^+$,

$$|Wf(u,s)| \leq A s^{\alpha+1/2}\left(1 + \left|\frac{u-v}{s}\right|^{\alpha}\right)$$

  - Conversely, if $\alpha < n$, $\alpha \notin \mathbb{Z}$ and there exists $A > 0$ and $\alpha' < \alpha$ such that $\forall(u,s) \in \mathbb{R} \times \mathbb{R}^+$,

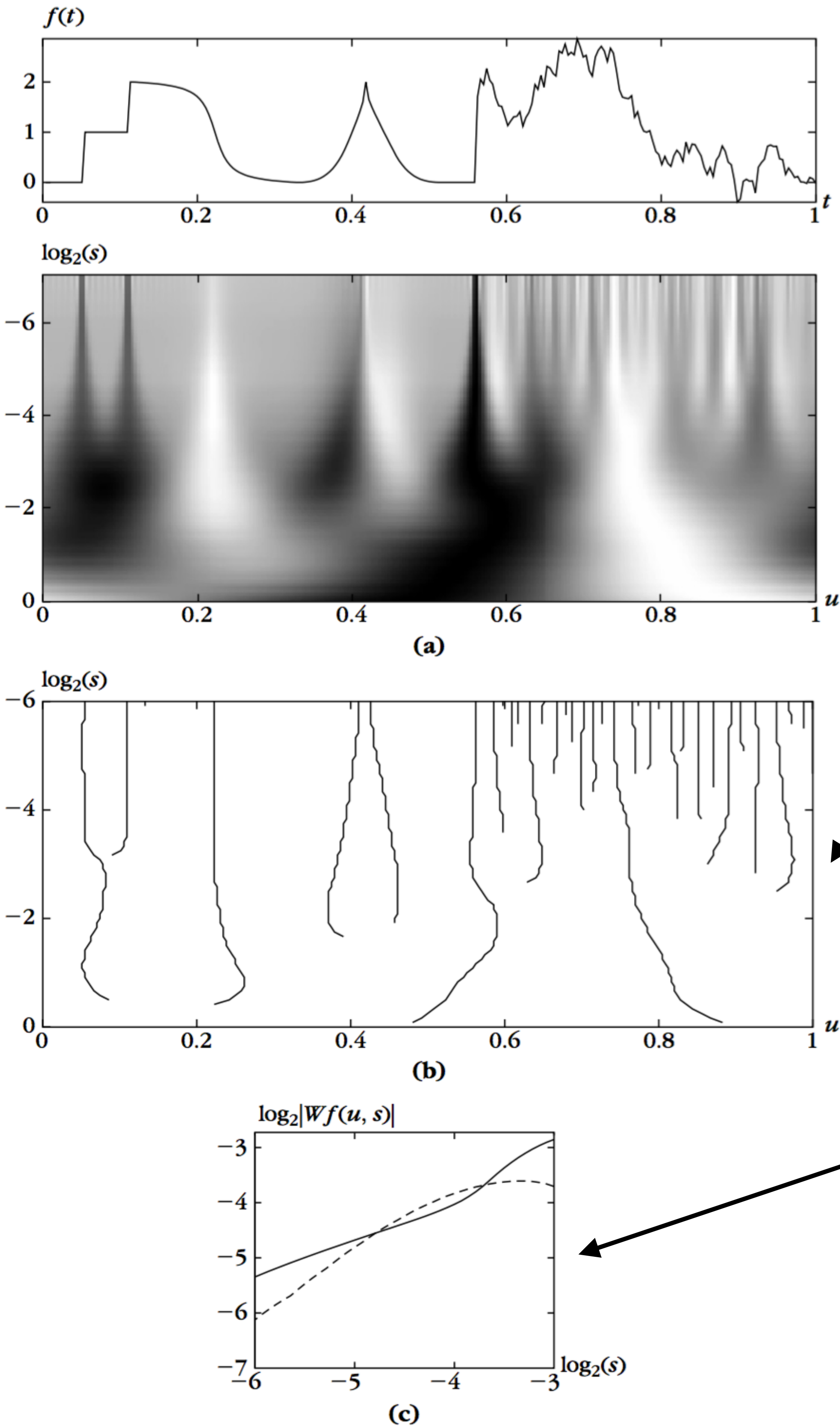
$$|Wf(u,s)| \leq A s^{\alpha+1/2}\left(1 + \left|\frac{u-v}{s}\right|^{\alpha'}\right)$$

    then f is Lipschitz α at v.

Wavelet Modulus Maxima

- Previous two theorems show that the local Lipschitz regularity of f at v depends on the decay of $|Wf(u,s)|$ as $s \to 0$.
- In fact, we only need to look at the local maxima of $|Wf(u,s)|$ to detect and characterize singularities of f.
- Wavelet modulus maximum is a point (u_0, s_0) such that $|Wf(u, s_0)|$ is locally maximum at $u = u_0$.

Maxima Propagation

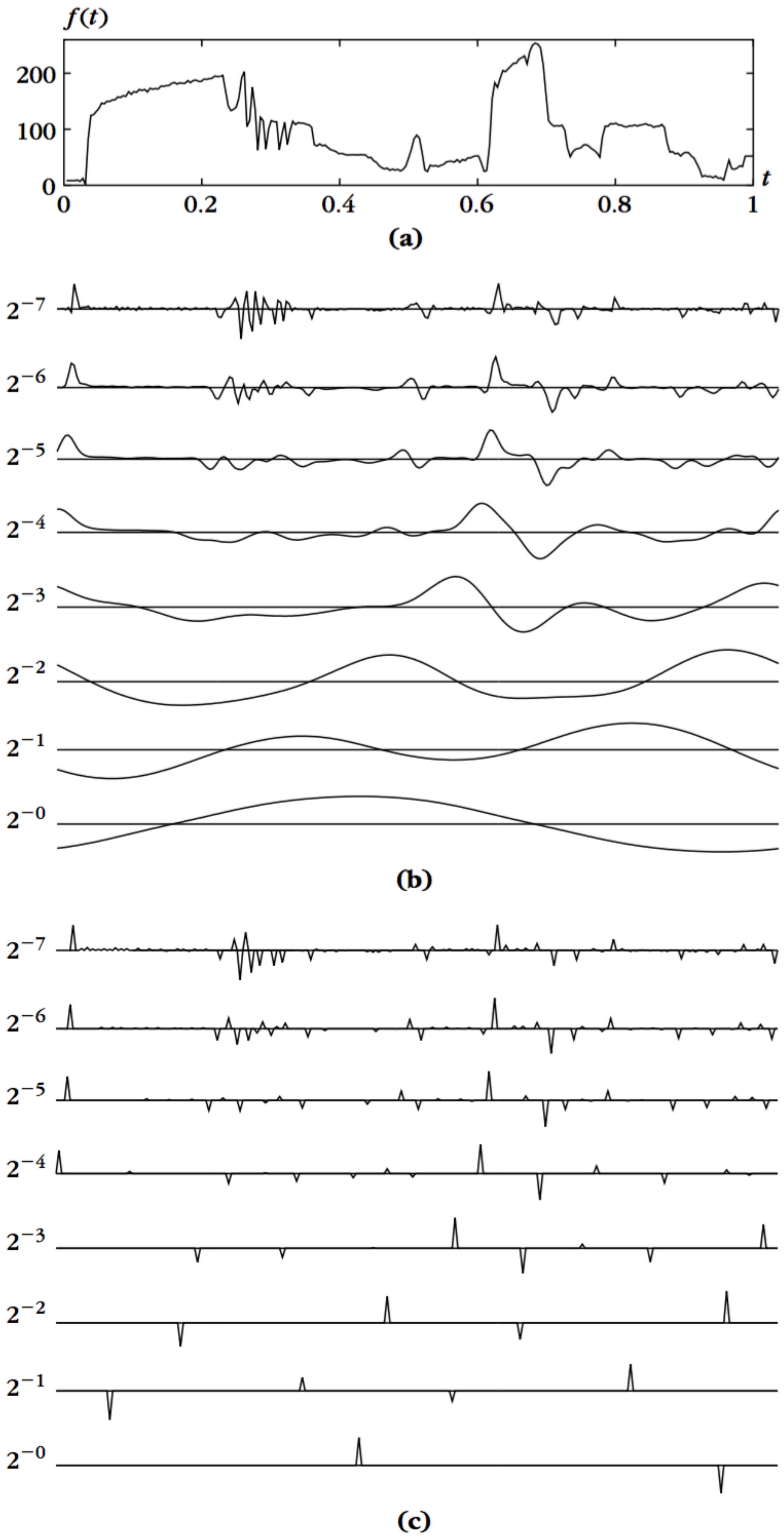


- Wavelet modulus maxima
- Theorem (Hwang, Mallat): f is singular at a point v only if there is a sequence of wavelet modulus maxima (u_p, s_p) that converges to v at fine scales:

$$\lim_{p \to +\infty} (u_p, s_p) = (v, 0)$$

- Theorem (Hummel, Poggio, Yuille): If $\psi = (-1)\theta^{(n)}$ for θ a Gaussian, then the wavelet modulus maxima belong to connected curves that are not interrupted as $s \to 0$.
- The maximum slope of $\log_2 |Wf(u,s)|$ as a function of $\log_2 s$ along the maximum line converging to v is $\alpha + 1/2$.
- Full line: Decay of $\log_2 |Wf(u,s)|$ along maxima line converging to $t = 0.05$.
- Dashed line: Maxima line converging to $t = 0.42$.

Dyadic Wavelet Transform and Maxima

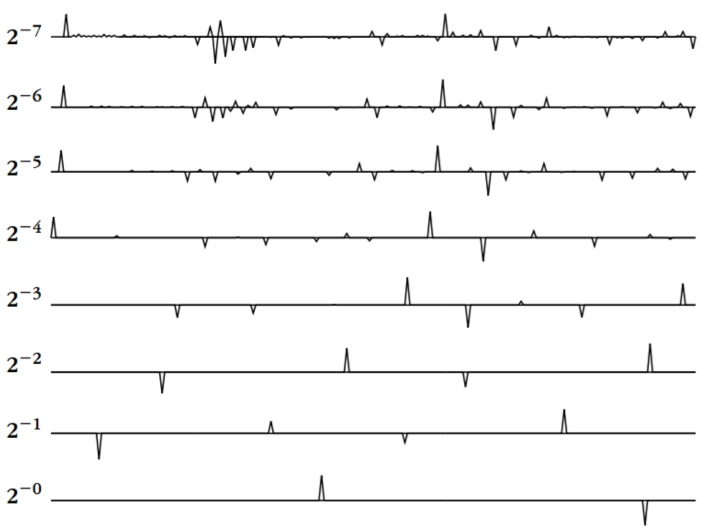


- Dyadic wavelet transform:

$$Wf(u, 2^j) = f * \tilde{\psi}_{2^j}(u)$$

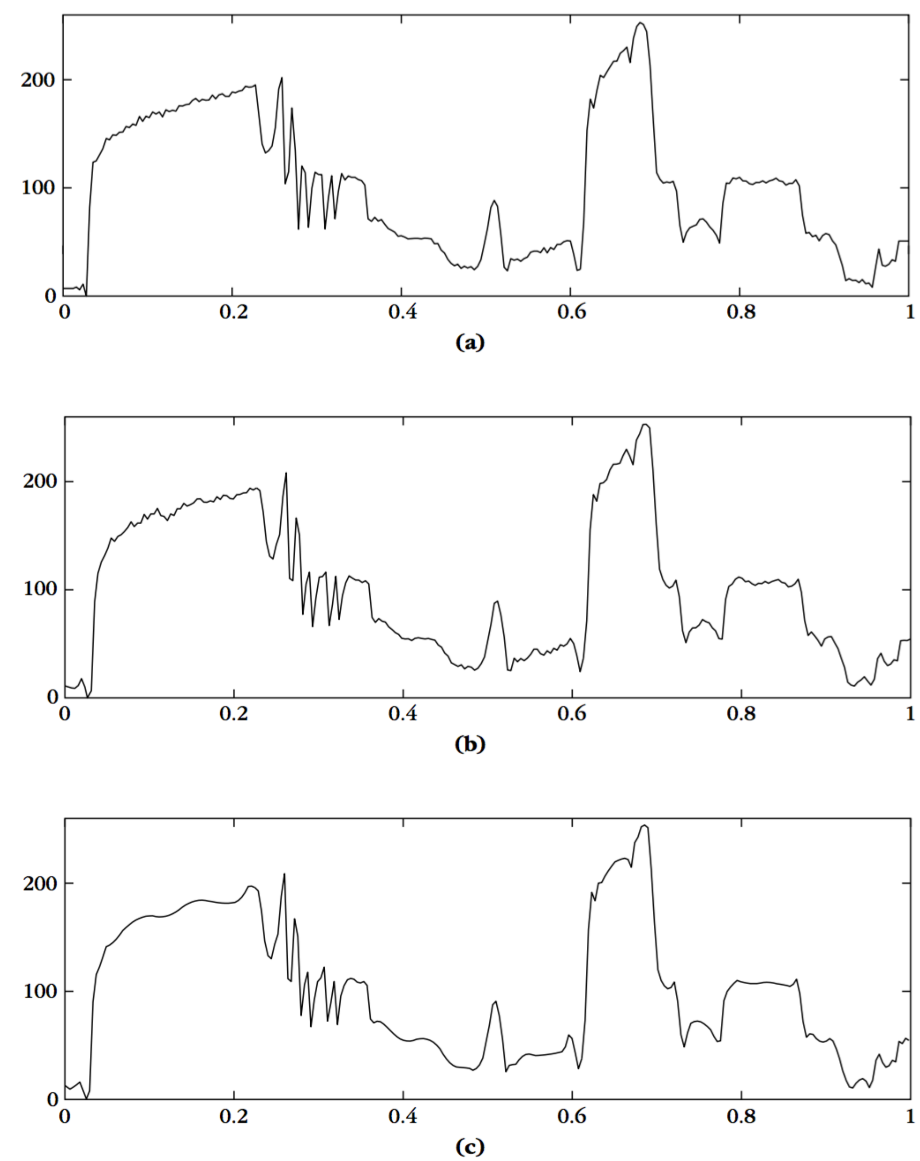
- Wavelet maxima (keeping the sign)

Wavelet Maxima Approximation in 1D



Analysis

Synthesis

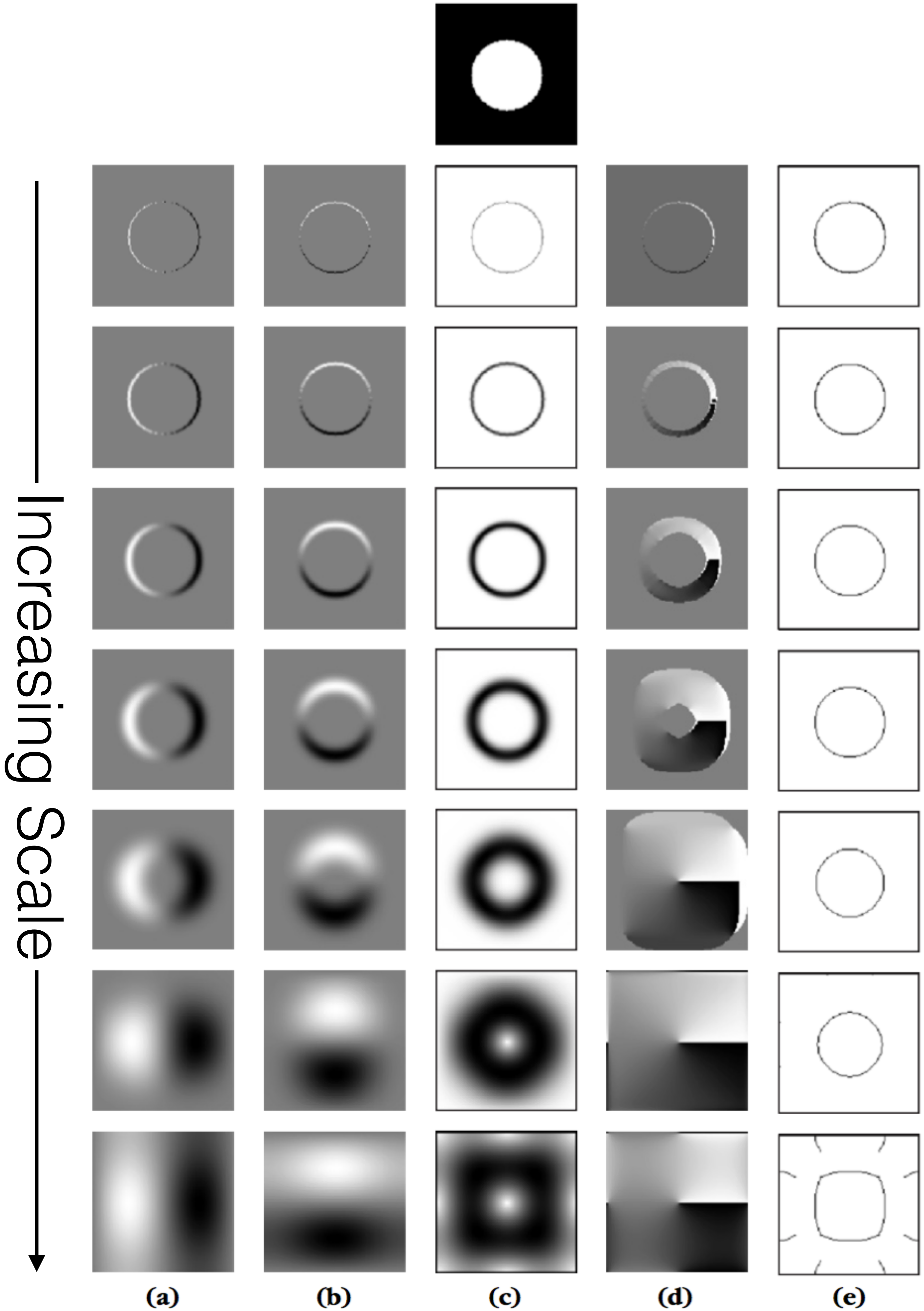


- $f(t)$

- Approximation of $f(t)$ with 100% wavelet maxima

- Approximation of $f(t)$ with 50% wavelet maxima

Wavelet Transform and Modulus Maxima in 2D



(a) Wavelet transform in horizontal direction

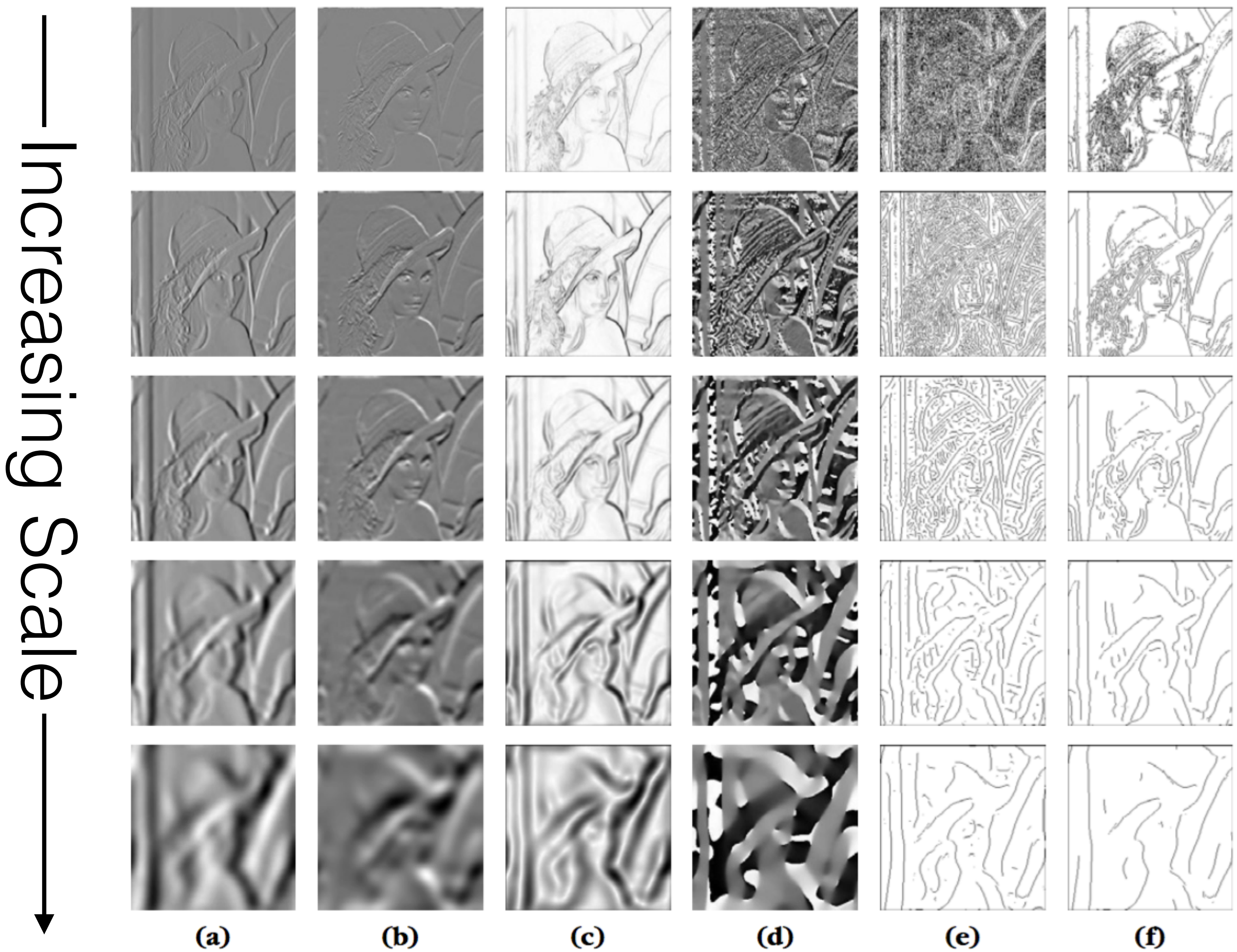
(b) Wavelet transform in vertical direction

(c) Wavelet transform modulus

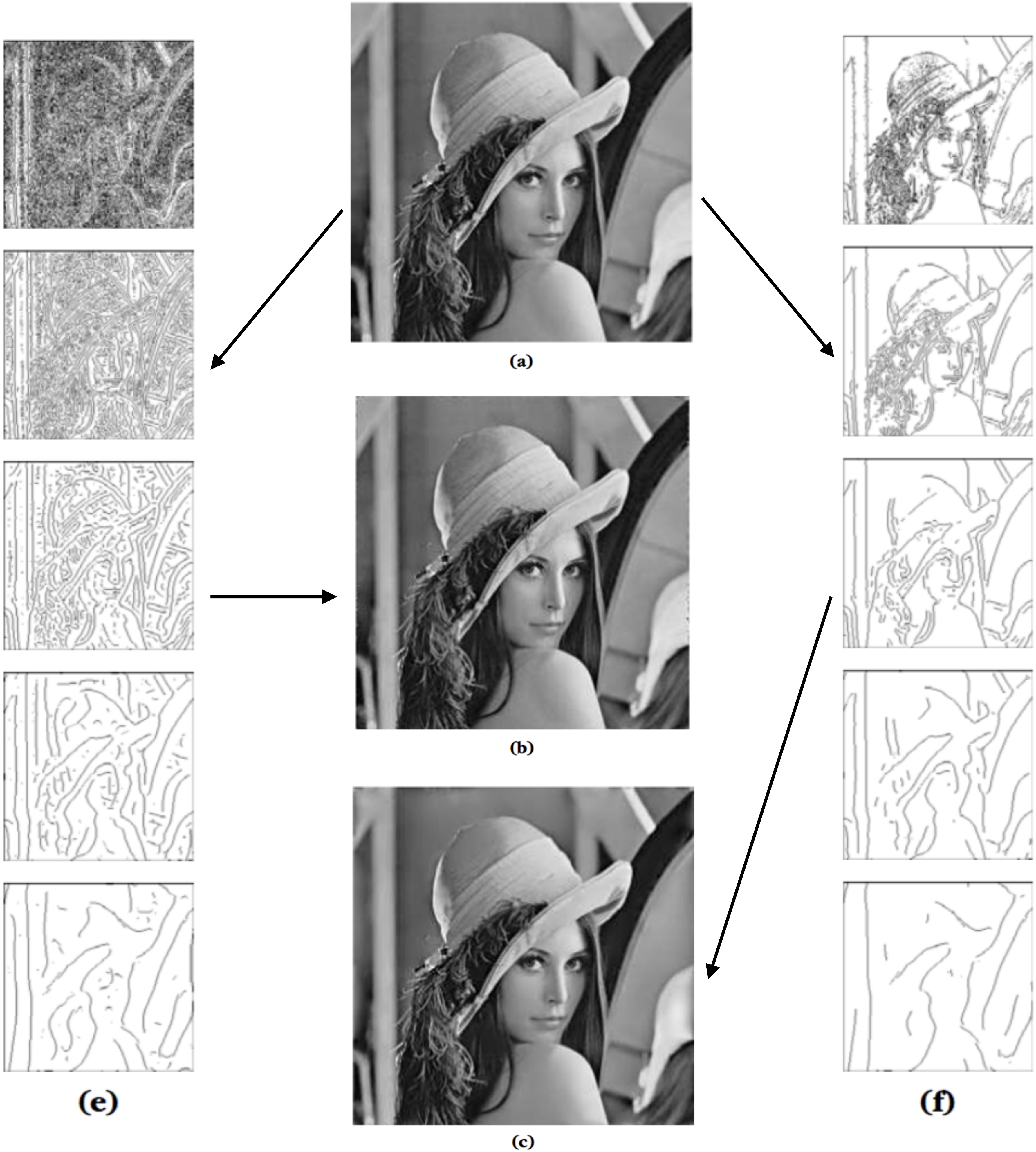
(d) Angles

(e) Wavelet modulus maxima

Wavelet Transform and Modulus Maxima in 2D



- (a) Wavelet transform in horizontal direction
- (b) Wavelet transform in vertical direction
- (c) Wavelet transform modulus
- (d) Angles
- (e) Wavelet modulus maxima
- (e) Wavelet modulus maxima above a threshold



Wavelet Maxima Approximation in 2D

(a) Original Image

(b) Approximation from 100% wavelet maxima (e)

(c) Approximation from thresholded wavelet maxima (f)

Dyadic Wavelet Frames

Translation Invariant Frames

- Recall translation invariant dictionary:

$$\mathcal{D} = \{\phi_{u,\gamma}\}_{\gamma \in \Gamma, u \in \mathbb{R}}, \quad \phi_{u,\gamma}(t) = \phi_\gamma(t-u)$$

and the (frame) operator:

$$\Phi f(u,\gamma) = \langle f, \phi_{u,\gamma} \rangle = f * \tilde{\phi}_\gamma(u), \quad \tilde{\phi}_\gamma(t) = \overline{\phi_\gamma(-t)}$$

- A translation invariant dictionary is a frame for $\mathbf{L}^2(\mathbb{R})$ if there exists $B \geq A > 0$ such that for all $f \in \mathbf{L}^2(\mathbb{R})$,

$$A\|f\|_2^2 \leq \sum_\gamma \|\Phi f(\cdot,\gamma)\|_2^2 \leq B\|f\|_2^2$$

where

$$\|\Phi f(\cdot,\gamma)\|_2^2 = \int_{-\infty}^{+\infty} |\Phi f(u,\gamma)|^2 \, du = \int_{-\infty}^{+\infty} |f * \tilde{\phi}_\gamma(u)|^2 \, du$$

- When $A = B$ the frame is tight.

- Frames are redundant.

Translation Invariant Frames

- Theorem: If there exists $B \geq A > 0$ such that for almost every $\omega \in \mathbb{R}$,

$$A \leq \sum_{\gamma} |\widehat{\phi}_{\gamma}(\omega)|^2 \leq B,$$

then

$$\mathcal{D} = \{\phi_{u,\gamma}\}_{\gamma \in \Gamma, u \in \mathbb{R}}, \quad \phi_{u,\gamma}(t) = \phi_{\gamma}(t-u)$$

is a frame for $\mathbf{L}^2(\mathbb{R})$.

- Define the generators $\{\varphi_{\gamma}\}_{\gamma}$ of the dual frame via:

$$\widehat{\varphi}_{\gamma}(\omega) = \frac{\widehat{\phi}_{\gamma}(\omega)}{\sum_{\gamma'} |\widehat{\phi}_{\gamma'}(\omega)|^2}$$

- We then have the following reconstruction formula:

$$f(t) = \sum_{\gamma} \Phi f(\cdot, \gamma) * \varphi_{\gamma}(t) = \sum_{\gamma} f * \tilde{\phi}_{\gamma} * \varphi_{\gamma}(t)$$

Dyadic Wavelet Frame

- A translation invariant dyadic wavelet dictionary is defined as:

$$\mathcal{D} = \left\{ \psi_{u,2^j}(t) = 2^{-j}\psi(2^{-j}(t-u)) \right\}_{u\in\mathbb{R}, j\in\mathbb{Z}}$$

- Dyadic wavelet transform:

$$Wf(u, 2^j) = f * \tilde{\psi}_{2^j}(u), \quad \tilde{\psi}_{2^j}(t) = 2^{-j}\overline{\psi(-2^{-j}t)}$$

- Corollary: If there exists $B \geq A > 0$ such that for all $\omega \in \mathbb{R} \setminus \{0\}$,

$$A \leq \sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j\omega)|^2 \leq B$$

  then the dyadic wavelet dictionary is a frame.

- If $A = B = 1$, then reconstruction is particularly simple:

$$f(t) = \sum_{j=-\infty}^{+\infty} f * \tilde{\psi}_{2^j} * \psi_{2^j}(t)$$

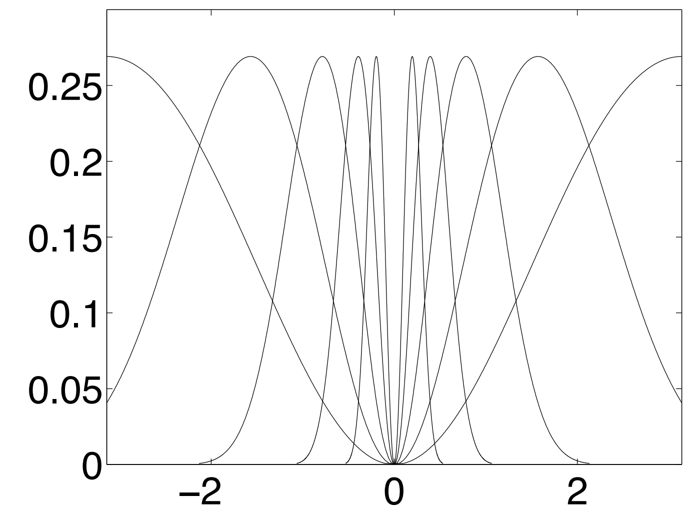
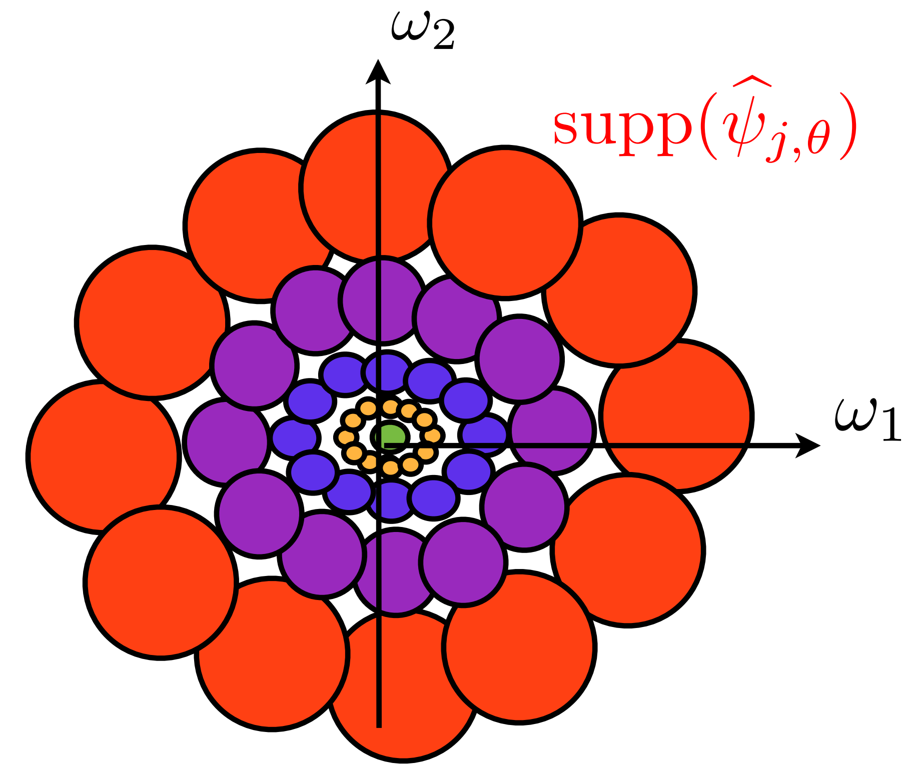
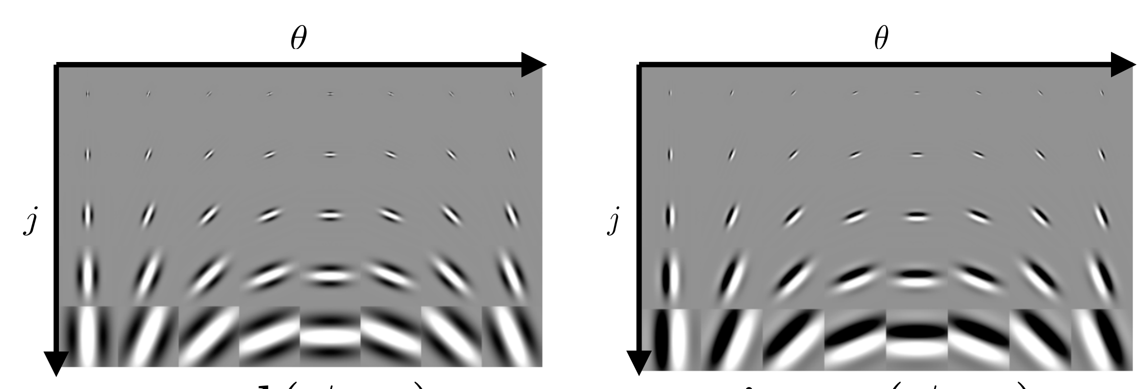


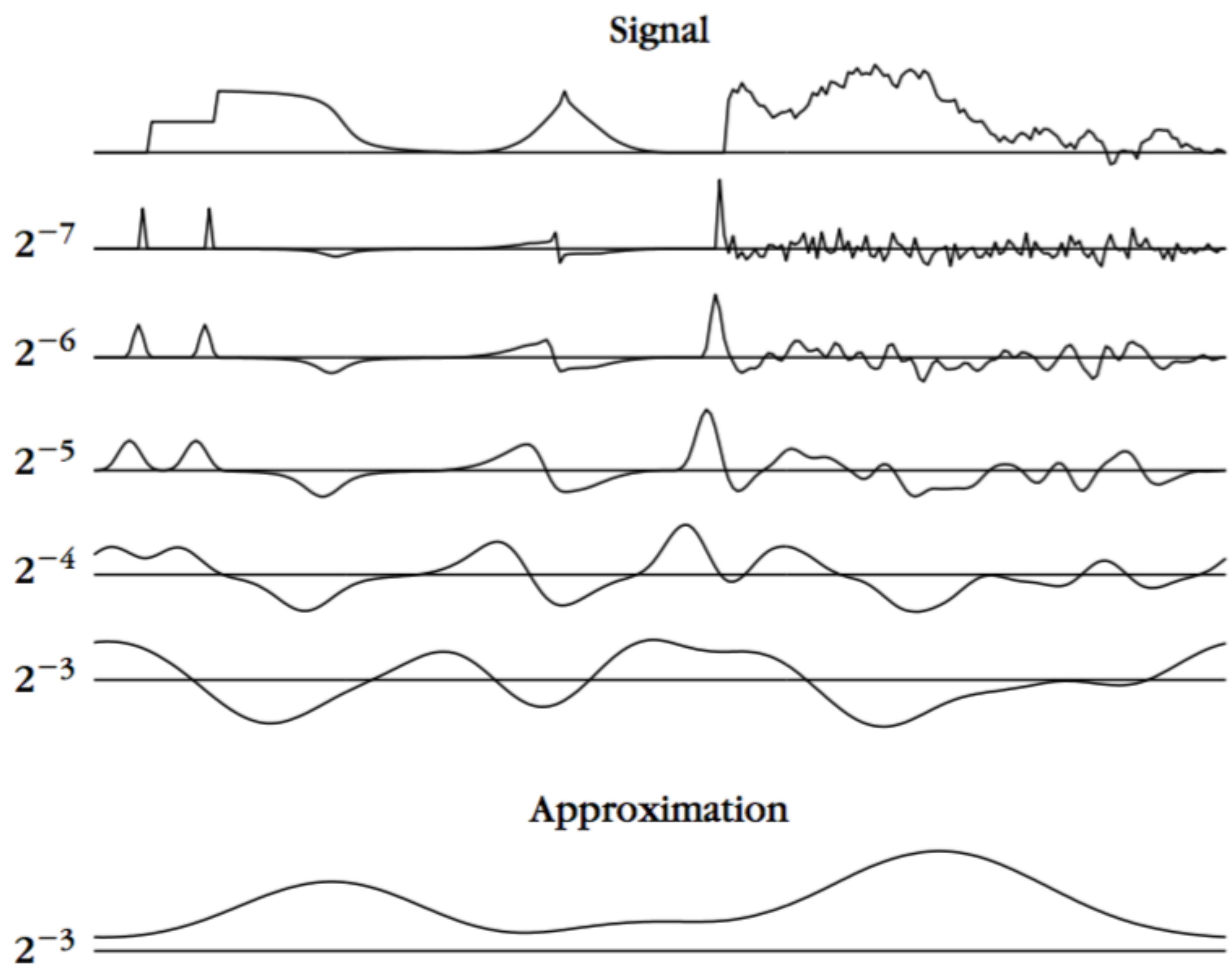
Fig. 5.1. A Wavelet Tour of Signal Processing, 3rd ed. Scaled Fourier transforms $|\hat{\psi}(2^j\omega)|^2$, for $1 \leqslant j \leqslant 5$ and $\omega \in [-\pi, \pi]$.





$\text{real}(\psi_{j,\theta})$ $\text{imag}(\psi_{j,\theta})$

1D Wavelet Transform at Different Scales



- $Wf(u, 2^j) = f * \tilde{\psi}_{2^j}(u)$ captures the details of f at the scale 2^j.

2D Wavelet Transform at Different Scales