Wavelets and multiresolution representations

Time meets frequency

Time-Frequency resolution

- Depends on the time-frequency spread of the wavelet atoms

Assuming that ψ is centred in t=0

Signal domain

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

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

Fourier domain

$$\eta = \frac{1}{2\pi} \int_{0}^{+\infty} \omega \, |\hat{\psi}(\omega)| \, d\omega$$

$$\hat{\psi}_{u,s}(\omega) = \sqrt{s}\,\psi(s\omega) e^{-i\omega u} \rightarrow \text{center frequency } \eta/s$$

Energy spread around η/s

$$\frac{\sigma_\omega^2}{s^2} = \frac{1}{2\pi} \int_{0}^{+\infty} \left(\omega - \frac{\eta}{s}\right)^2 \left|\hat{\psi}_{u,s}(\omega)\right|^2 d\omega$$

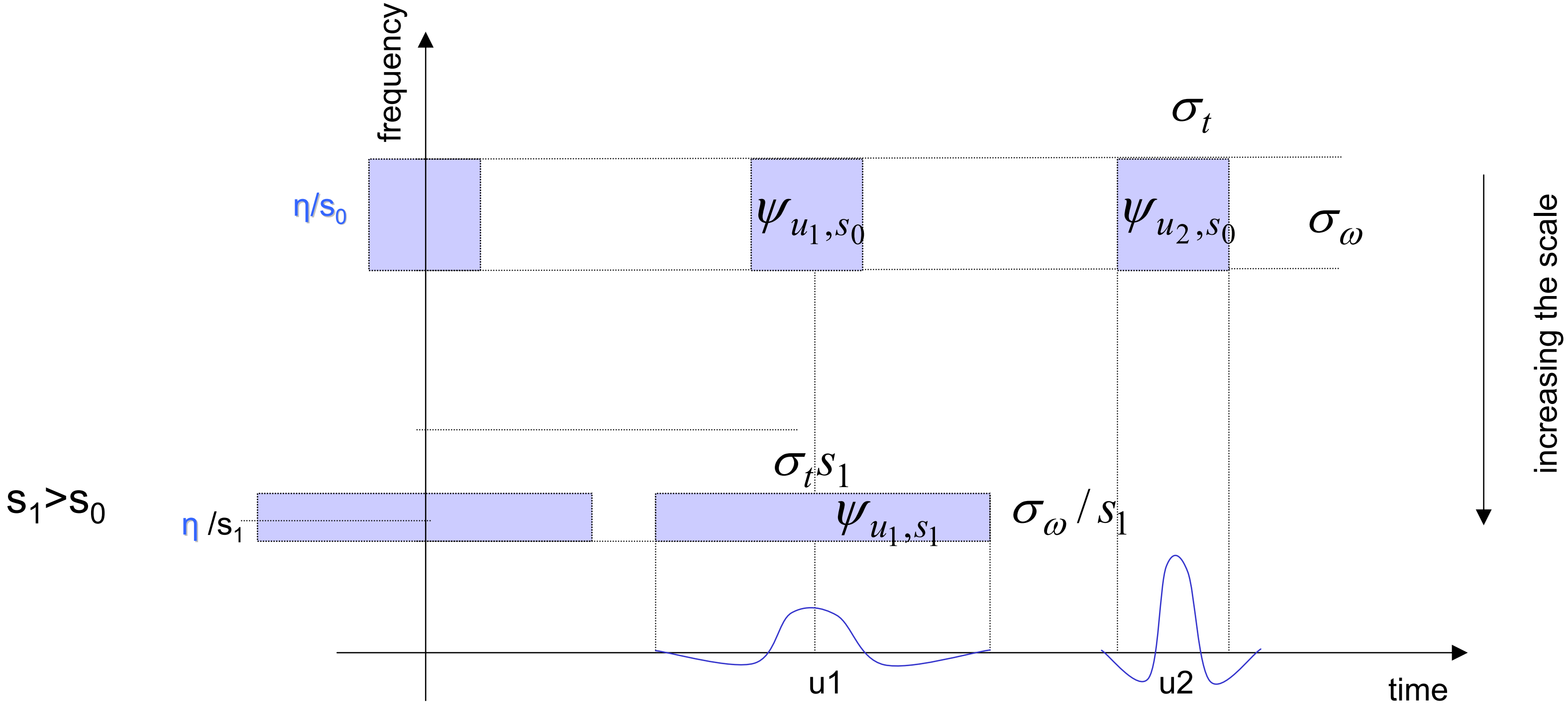
Time/frequency resolution

$$\sigma_{s,t}{}^2 = s^2 \sigma_t{}^2$$

$$\sigma_{s,\omega}{}^2 = \frac{\sigma_\omega{}^2}{s^2}$$

- The energy spread of a wavelet time-frequency atom corresponds to an Heisemberg box centred at $(u,\eta/s)$ of size $s\sigma_t$ along the time and σ_ω/s along the frequency.
- The area of the rectangle **remains equal to $\sigma_t\ \sigma_\omega$ at all scales**, while the resolution in time and frequency depends on s.
- A wavelet defines a local time-frequency energy density $P_W f$ which measures the energy in the Heisemberg box of each wavelet centred at $(u, \eta /s)$. This energy density is called scalogram

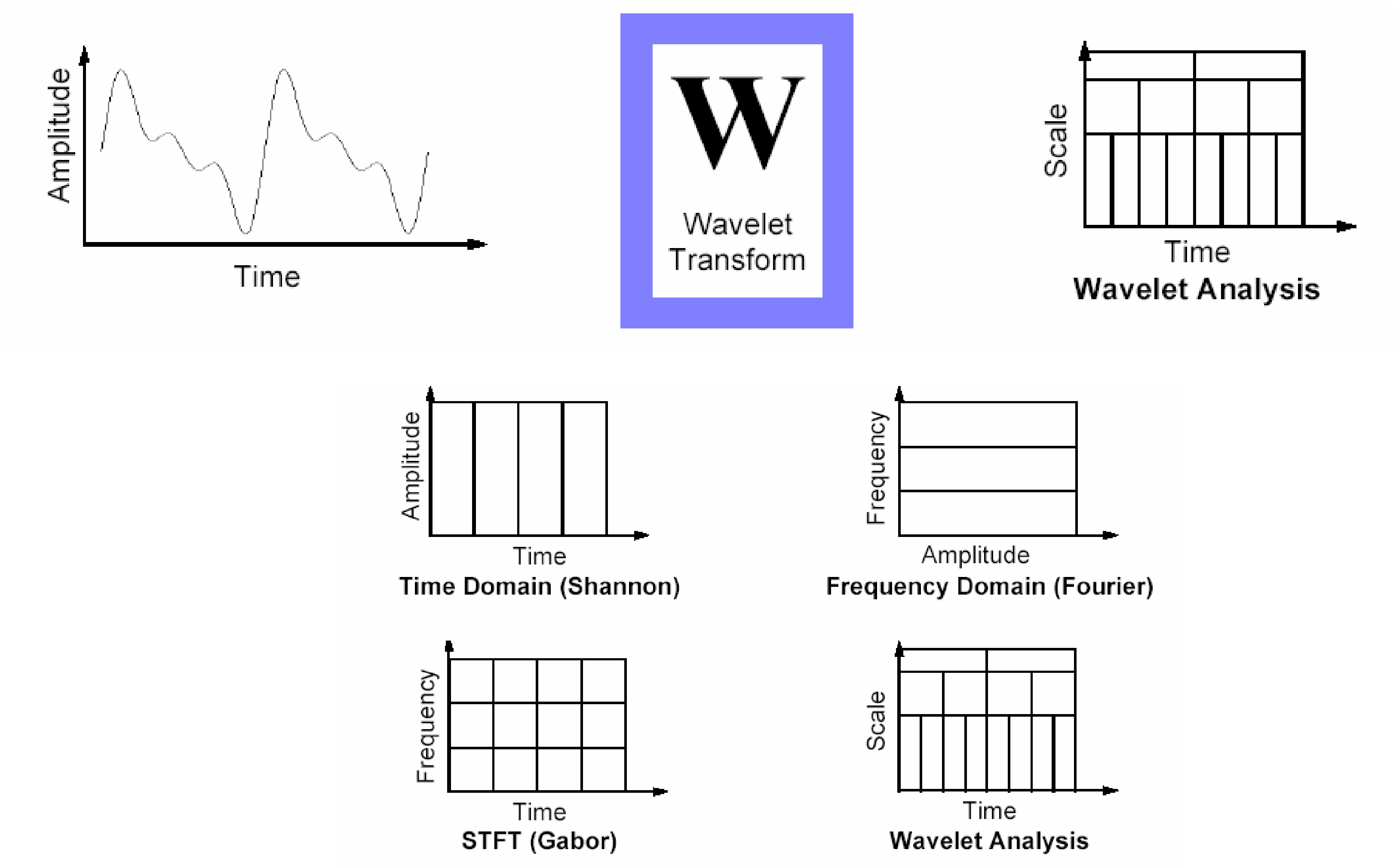
Time/frequency localization



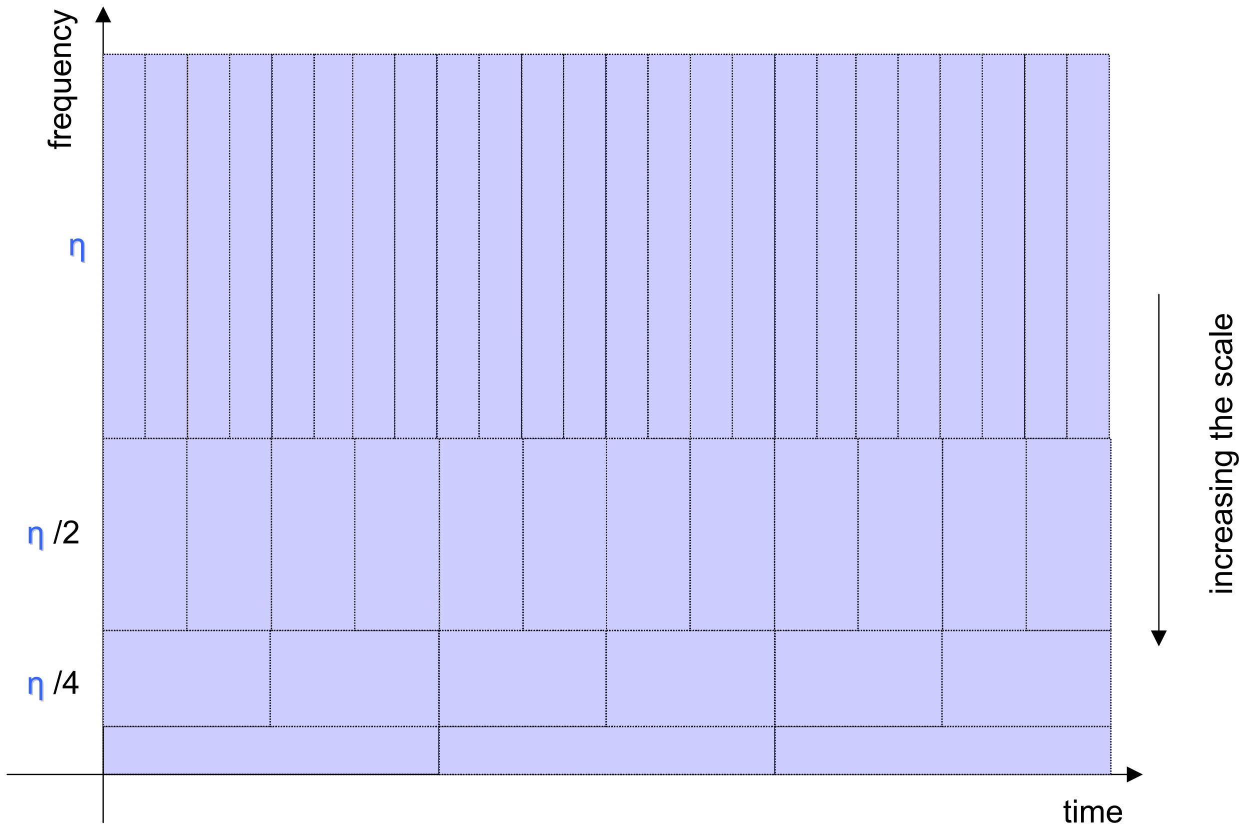
Increasing the scale (s gets larger) pushes the box towards low frequencies → frequency resolution increases, spatial resolution decreases

Time spread is proportional to scale
Frequency spread is proportional to 1/scale

Wavelet domain



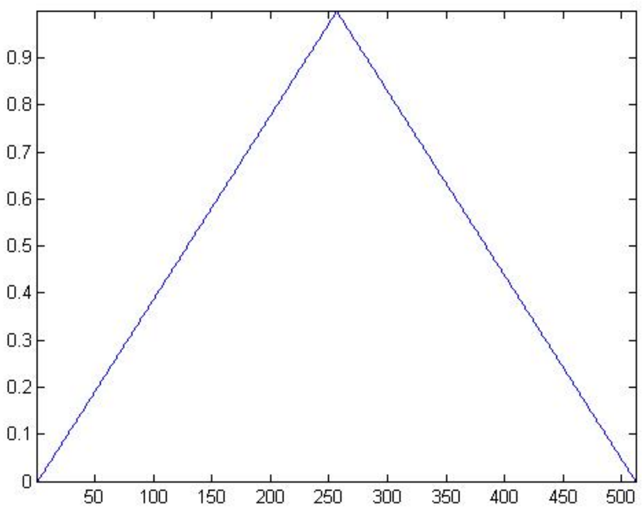
Dyadic Wavelets

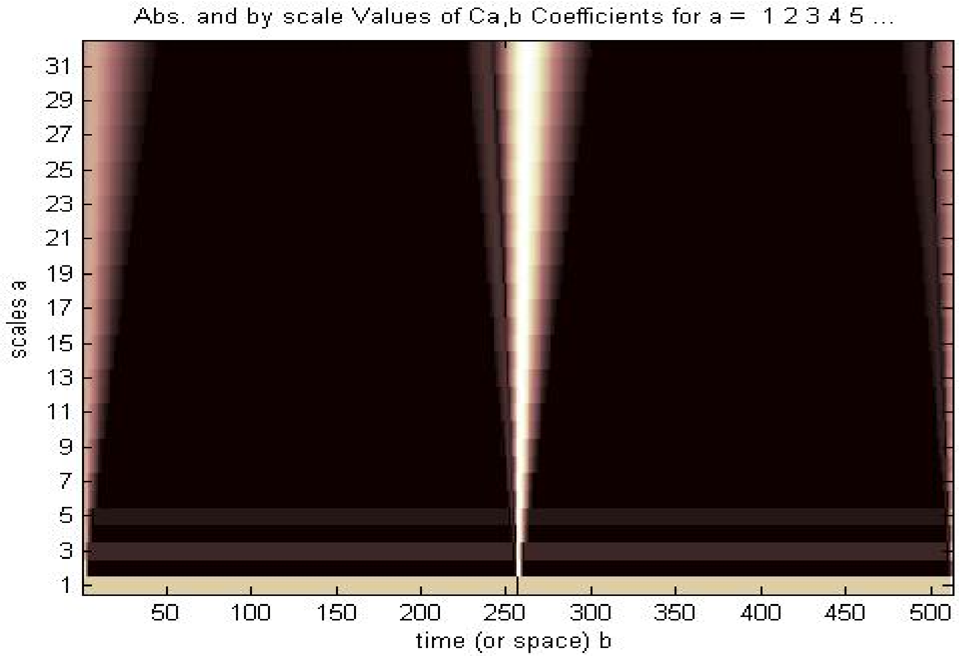


Scalogram

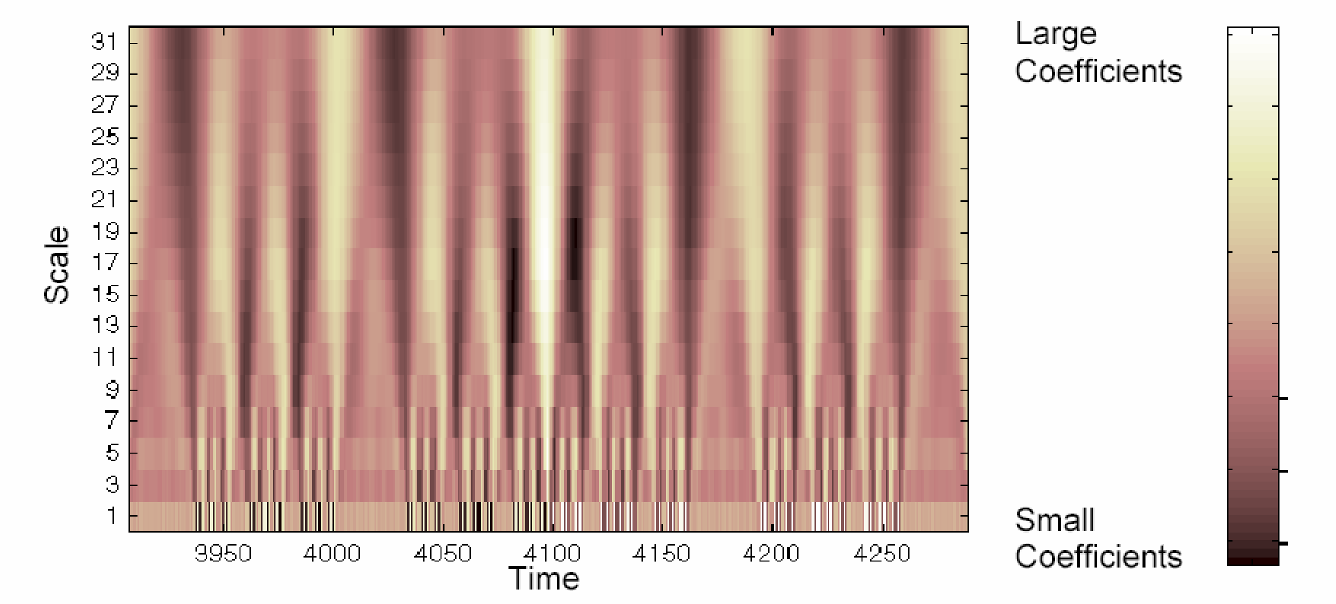
- The scalogram represents the local time/frequency energy density
  - Energy density in the Heisenberg box of each wavelet $\psi_{u,s}$

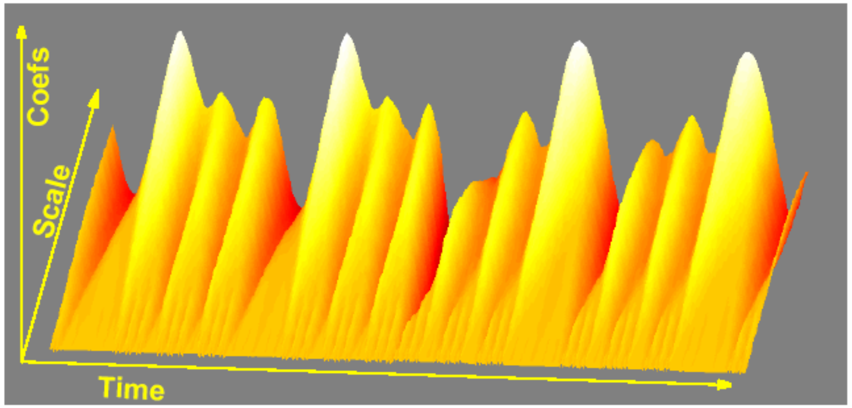
$$P_W f(u,\xi) = \left| Wf\left(u, \frac{\eta}{\xi}\right) \right|^2$$



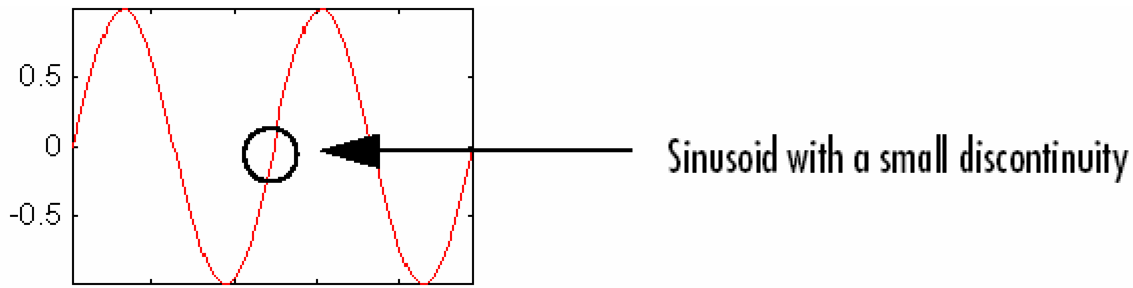


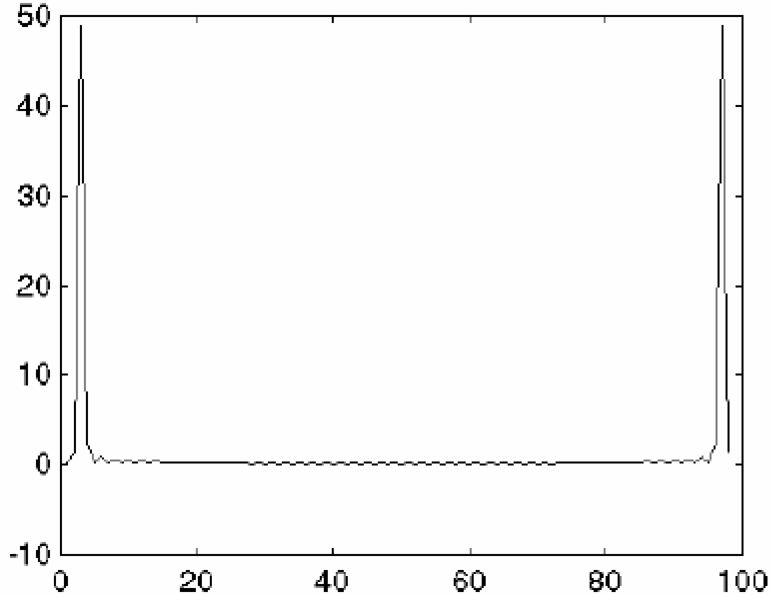
3D representation



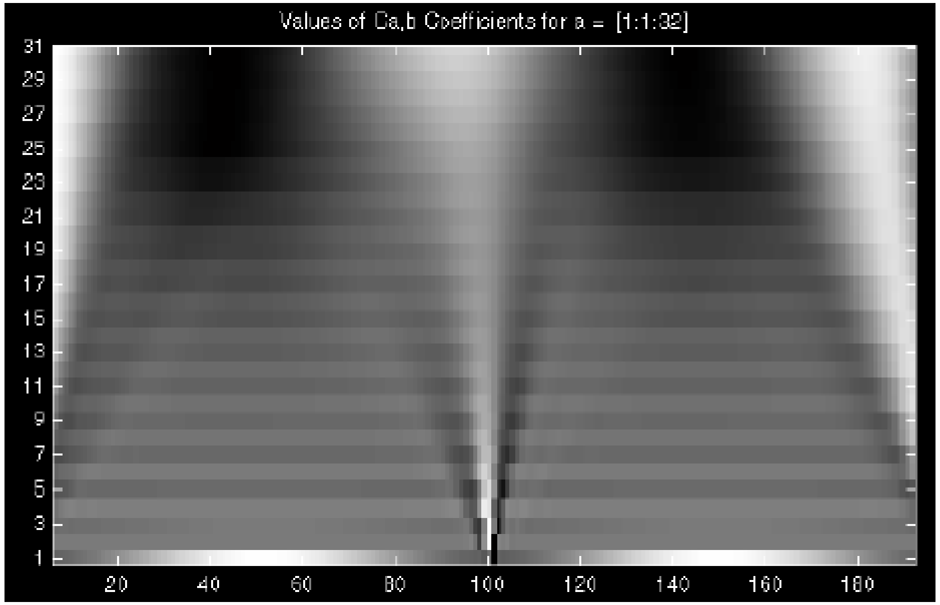


Local discontinuities





Fourier Coefficients



Wavelet Coefficients

Real Wavelets

- Detect sharp signal transitions

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

- Measures the variations of f in the neighborhood of u whose size is proportional to s
- A real WT is complete and maintains energy conservation as long as it satisfies a weak admissibility condition (Theorem 4.3, next slide)
- The decay of the coefficients as s goes to zero characterizes the regularity of f in the neighborhood of u

Real wavelets: Admissibility condition

- Theorem 4.3 (Calderon, Grossman, Morlet)

Let ψ in $L^2(R)$ be a real function such that

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

Admissibility condition

Any f in $L^2(R)$ satisfies

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

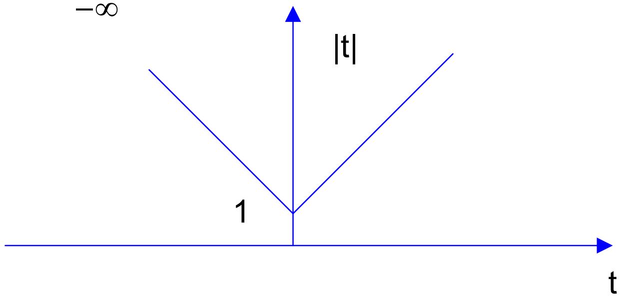
and

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

Admissibility condition

- Consequences
  - The integral is finite if the wavelet has zero average $\hat{\psi}(0) = 0$
    - This condition is nearly sufficient →
  - If $\hat{\psi}(0) = 0$ and $\hat{\psi}(\omega)$ is continuously differentiable, than the admissibility condition is satisfied
    - This happens if it has a sufficient time decay

$$\int_{-\infty}^{+\infty} (1+|t|)\,|\psi(t)|\,dt < +\infty$$



→ The wavelet function must decay **sufficiently fast** in both time and frequency

Wavelet families

$$f(\vec{x}) \leftrightarrow Wf(u,s;\vec{x}) = c_{u,s}(\vec{x})$$

- In general, there is a redundancy in the representation
- The amount of redundancy depends on the grids over which the u and s parameters are sampled

  u,s are real : Continuous WT (CWT, overcomplete representation)

  u in Z, $s=a^j$, j in Z : Wavelet Frames (DWF, DDWF, overcomplete)

  – $a=2$ Dyadic wavelet frames

  $u=k2^j$, $s=2^j$, k *in* I : Discrete Wavelet Transform (DWT) (*critically sampled*)

- Note: removing completely the redundancy leads to complete basis (*critically sampled*)

Wavelet bases

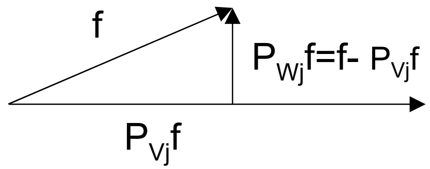
Mallat - Chapter VII

Wavelet *bases*

One can construct wavelets such that

$$\left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j n}{2^j} \right) \right\}_{j,n \in Z^2}$$

is an orthonormal basis for $L^2(R)$.



- Multiresolution approximations
  - The partial sum of wavelet coefficients giving $d_j(t)$ can be interpreted as the difference between two approximations of f at the scales 2^j and $2^{(j-1)}$
  - Multiresolution approximations compute the approximations of signals at various resolutions with orthogonal projections to different spaces $\{V_j\}_{j \text{ in } Z}$
  - The approximation of f at scale 2^j is specified by a discrete grid of samples that provides *local averages* of f on neighborhoods of size proportional to 2^j.
  - **A *multiresolution* consists of *embedded grids of approximations***

Orthogonal wavelet bases

- The search for orthogonal wavelets begins with multiresolution approximations

$$f \in L^2(\Re) \rightarrow \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \text{ difference bewteen two approximations}$$

at resolutions 2^{-j+1} and 2^{-j}

- Resolution = 1/scale
  - The larger the scale, the smaller the resolution
- Multiresolution approximations compute the approximation of signals at various resolutions with orthogonal projections on different spaces $\{V_j\}_{j \in Z}$
  - These are characterized by a one particular discrete filter that governs the loss of information across resolutions

Multiresolution approximations

- The approximation of a function *f* at a resolution 2^j is specified by a discrete grid of samples that provides local averages of *f* over neighborhoods of size proportional to 2^j.
- Thus, a multiresolution approximation is composed of *embedded grids of approximation*.
- More formally:
  - *the approximation of a function at a resolution 2^j is defined as an* **orthogonal projection** *on a space $\mathbf{V}_j \subset \mathbf{L}^2(R)$.*
  - *The space $\mathbf{V}_j$ regroups* ***all possible*** *approximations at the resolution 2^j .*
  - *The orthogonal projection of f is the function $f_j \in \mathbf{V}_j$ that minimizes $||f - f_j||$.*

Multiresolution approximations

Definition 7.1 A sequence $\{V_j\}_{j \, in \, Z}$ of closed subspaces of $L^2(R)$ is a multiresolution approximation if the following six conditions are satisfied

$$\forall (j,k) \in Z^2, f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j$$

V_j is invariant for translations proportional to the scale

$$\forall j \in Z, \quad V_{j+1} \subset V_j$$

The *finer* approximation subspace encloses all the information concerning the coarser one

$$\forall j \in Z, f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1}$$

Stratching the function by a factor 2 spans a coarser subspace

$$\lim_{j \to +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$$

When the resolution goes to zero all the details are lost

$$\lim_{j \to +\infty} \|P_{\mathbf{V}_j} f\| = 0.$$

$$\lim_{j \to -\infty} V_j = Closure\left(\bigcup_{j=-\infty}^{+\infty} V_j\right) = L^2(R)$$

When the resolution goes to infinity the approximation converges to the signal

$$\lim_{j \to -\infty} \|f - P_{\mathbf{V}_j} f\| = 0.$$

There exists ϑ such that $\{\vartheta(t-n)\}_{n \in Z}$ is a Riesz basis of V_0

Banach and Hilbert spaces

- A Hilbert space is an abstract vector space possessing the structure of an inner product that allows length and angle to be measured. Hilbert spaces are in addition required to be *complete*, a property that stipulates the existence of enough limits in the space to allow the techniques of calculus to be used.

Banach and Hilbert spaces

- Banach space

Signals are often considered as vectors. To define a distance, we work within a vector space $\mathbf{H}$ that admits a norm. A norm satisfies the following properties:

$$\forall f \in \mathbf{H}, \quad \|f\| \geqslant 0 \quad \text{and} \quad \|f\| = 0 \;\Leftrightarrow\; f = 0, \tag{A.3}$$

$$\forall \lambda \in \mathbb{C} \;\; \|\lambda f\| = |\lambda| \, \|f\|, \tag{A.4}$$

$$\forall f, g \in \mathbf{H}, \quad \|f + g\| \leqslant \|f\| + \|g\|. \tag{A.5}$$

With such a norm, the convergence of $\{f_n\}_{n\in\mathbb{N}}$ to f in $\mathbf{H}$ means that

$$\lim_{n\to+\infty} f_n = f \;\Leftrightarrow\; \lim_{n\to+\infty} \|f_n - f\| = 0.$$

To guarantee that we remain in $\mathbf{H}$ when taking such limits, we impose a completeness property, using the notion of *Cauchy sequences*. A sequence $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence if for any $\varepsilon > 0$, if n and p are large enough, then $\|f_n - f_p\| < \varepsilon$. The space $\mathbf{H}$ is said to be *complete* if every Cauchy sequence in $\mathbf{H}$ converges to an element of $\mathbf{H}$.

Banach and Hilbert spaces

- Hilbert space

Whenever possible, we work in a space that has an inner product to define angles and orthogonality. A *Hilbert space* **H** is a Banach space with an inner product. The inner product of two vectors $\langle f, g\rangle$ is linear with respect to its first argument:

$$\forall \lambda_1, \lambda_2 \in \mathbb{C}, \quad \langle \lambda_1 f_1 + \lambda_2 f_2, g\rangle = \lambda_1 \langle f_1, g\rangle + \lambda_2 \langle f_2, g\rangle. \qquad \textbf{(A.6)}$$

It has an Hermitian symmetry:

$$\langle f, g\rangle = \langle g, f\rangle^*.$$

Moreover,

$$\langle f, f\rangle \geqslant 0 \quad \text{and} \quad \langle f, f\rangle = 0 \Leftrightarrow f = 0.$$

One can verify that $\|f\| = \langle f, f\rangle^{1/2}$ is a norm. The positivity (A.3) implies the Cauchy-Schwarz inequality:

$$|\langle f, g\rangle| \leqslant \|f\| \, \|g\|, \qquad \textbf{(A.7)}$$

which is an equality if and only if f and g are linearly dependent.

We write $\mathbf{V}^\perp$ the orthogonal complement of a subspace $\mathbf{V}$ of $\mathbf{H}$. All vectors of $\mathbf{V}$ are orthogonal to all vectors of $\mathbf{V}^\perp$ and $\mathbf{V} \oplus \mathbf{V}^\perp = \mathbf{H}$.

Bases of Hilbert spaces

Orthonormal Basis

A family $\{e_n\}_{n\in\mathbb{N}}$ of a Hilbert space $\mathbf{H}$ is orthogonal if for $n \neq p$,

$$\langle e_n, e_p \rangle = 0.$$

If for $f \in \mathbf{H}$ there exists a sequence $a[n]$ such that

$$\lim_{N\to+\infty} \|f - \sum_{n=0}^{N} a[n]\, e_n\| = 0,$$

then $\{e_n\}_{n\in\mathbb{N}}$ is said to be an *orthogonal basis* of $\mathbf{H}$. The orthogonality implies that necessarily $a[n] = \langle f, e_n\rangle / \|e_n\|^2$, and we write

$$f = \sum_{n=0}^{+\infty} \frac{\langle f, e_n\rangle}{\|e_n\|^2}\, e_n. \tag{A.8}$$

A Hilbert space that admits an orthogonal basis is said to be *separable*.

The basis is *orthonormal* if $\|e_n\| = 1$ for all $n \in \mathbb{N}$. Computing the inner product of $g \in \mathbf{H}$ with each side of (A.8) yields a Parseval equation for orthonormal bases:

$$\langle f, g\rangle = \sum_{n=0}^{+\infty} \langle f, e_n\rangle\, \langle g, e_n\rangle^*. \tag{A.9}$$

Bases of Hilbert space

When $g = f$, we get an energy conservation called the *Plancherel formula*:

$$\|f\|^2 = \sum_{n=0}^{+\infty} |\langle f, e_n \rangle|^2. \tag{A.10}$$

The Hilbert spaces $\ell^2(\mathbb{Z})$ and $\mathbf{L}^2(\mathbb{R})$ are separable. For example, the family of translated Diracs $\{e_n[k] = \delta[k-n]\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $\ell^2(\mathbb{Z})$. Chapters 7 and 8 construct orthonormal bases of $\mathbf{L}^2(\mathbb{R})$ with wavelets, wavelet packets, and local cosine functions.

Riesz basis

Link to the discrete domain: the existance of a Riesz bases provides a discretization theorem

Definition: A family of vectors is a Riesz basis of a space H if

1. *it is linearly independent*
2. *there exist A,B>0 such that*

$$\forall y \in H \qquad \exists \lambda[n]: \qquad y = \sum_{n=0}^{+\infty} \lambda[n] e_n$$

$$\frac{1}{B}\|y\|^2 \leq \sum_{n=0}^{+\infty} |\lambda[n]|^2 \leq \frac{1}{A}\|y\|^2$$

The existance of a Riesz basis for V_0 provides a ***discretization theorem***

$$\forall f(t) \in V_0 \to f(t) = \sum_n a[n] \vartheta(t-n)$$

$$A\|f\|^2 \leq \sum_n |a[n]|^2 \leq B\|f\|^2$$

$$\left\{ \frac{1}{\sqrt{2^j}} \vartheta\left(\frac{t - 2^j n}{2^j} \right) \right\}_{n \in Z}$$ is a Riesz basis for V_j

Scaling function

- The scaling function is obtained by the orthogonalization of the Riesz basis

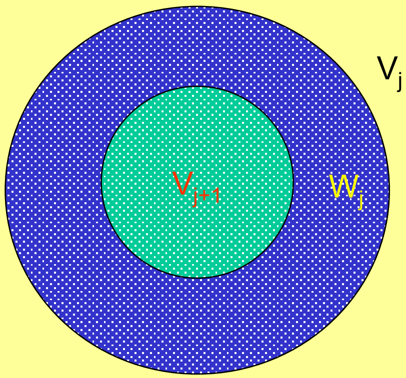
Theorem 7.1

Let V_j be a multiresolution approximation and $\boldsymbol{\varphi}$ be the scaling function whose FT is

$$\hat{\varphi}(\omega) = \frac{\hat{\vartheta}(\omega)}{\left(\sum_{k=-\infty}^{+\infty} \left|\hat{\vartheta}(\omega + 2k\pi)\right|^2\right)^{1/2}}$$

Let us denote

$$\varphi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \varphi\left(\frac{t - 2^j n}{2^j}\right)$$



The family $\{\boldsymbol{\varphi}_{j,n}\}_{n \text{ in } Z}$ is an orthonormal basis of V_j for all j in Z

Approximation

- The orthogonal projection of f onto V_j is obtained as an expansion in the scaling orthogonal basis

$$P_{V_j} f = \sum_{n=-\infty}^{+\infty} \langle f, \varphi_{j,n} \rangle \varphi_{j,n}$$

- The inner products $a_j[n]$ are the projection coefficients at scale 2^j

$$a_j[n] = \langle f, \varphi_{j,n} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^j}} \varphi\left(\frac{t - 2^j n}{2^j} \right) = f * \bar{\varphi}_j(2^j n)$$

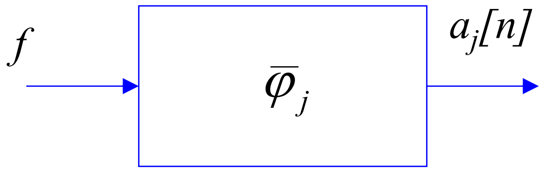
$$\bar{\varphi}_j(2^j n) = \frac{1}{\sqrt{2^j}} \varphi\left(-\frac{t}{2^j} \right)$$

  – The normalization factor at the denominator ensures that

$$\sum_{k=-\infty}^{\infty} \left| \hat{\varphi}(\omega + 2k\pi) \right|^2 = 1$$

Approximation

$$a_j[n] = f * \bar{\varphi}_j(2^j n)$$



only one sample out out of 2j is retained → downsampling

– The energy of φ_j is mostly concentrated in $[-\pi/2^j, \pi/2^j]$ which corresponds to low pass filtering

- The *signal approximation* is obtained by convolving f with a *low-pass filter* and downsampling by 2 -> any scaling function corresponds to a *conjugate mirror filter*
- A multiresolution is *completely characterized* by the scaling function

Wavelet representation

- Summarizing

$$A^{d}_{2^j} f = PV_j f = \sum_{n=-\infty}^{+\infty} \langle f, \varphi_{j,n} \rangle \varphi_{j,n}$$

discrete approximation at resolution j

$$a_j[n] = \langle f, \varphi_{j,n} \rangle$$

discrete approximation coefficients at resolution j

$$d_{2^j} f = PW_j f = \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}$$

details at resolution j

$$d_j[n] = \langle f, \psi_{j,n} \rangle$$

wavelet coefficients at resolution j

$$\left\{ A^{d}_{2^J} f, \left\{ d_{2^j} f \right\}_{1 \le j \le J} \right\}$$

wavelet representation

Wavelets and multiresolution representations

Scaling equation

- A multiresolution approximation is completely characterized by the function φ that generates the orthonormal bases for each V_j

→ We study the properties of φ which guarantee that all the spaces V_j satisfy all conditions of a multiresolution approximation.

→ It is proved that **any scaling function corresponds to a discrete filter called conjugate mirror filter**

- Procedure
  1. Link φ to the corresponding discrete filter *h[n]*
  2. Determine the properties of *h[n]* such that φ is a scaling function

Scaling equation

- From multiresolution conditions follows

$$V_j \subset V_{j-1}$$

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right) \subset V_1 \subset V_0$$

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} h[n]\varphi(t-n) \qquad (1)$$

$$h[n] = \left\langle \frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right), \varphi(t-n) \right\rangle \longrightarrow f(t)$$

- The **scaling equation** relates a dilation of φ by 2 to its integer translations.
- The sequence *h[n]* will be interpreted as a discrete filter

Scaling equation

- Taking the F-trasform of (1)

convolution product

$$\Im\left\{\frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right)\right\} = \Im\left\{\sum_{n=-\infty}^{+\infty} h[n]\varphi(t-n)\right\} \rightarrow$$

$$\hat{\phi}(2\omega) = \frac{1}{\sqrt{2}}\hat{h}(\omega)\hat{\phi}(\omega) \qquad (2)$$

  - where

$$\hat{h}(\omega) = \sum_{n=-\infty}^{+\infty} h[n]e^{-j\omega n}$$

- Next step is thus the expression of ^φ(ω) as a product of dilations of ^h(ω).
  - For any p≥0, (2) implies

$$\hat{\phi}\left(2^{-p+1}\omega\right) = \frac{1}{\sqrt{2}}\hat{h}\left(2^{-p}\omega\right)\hat{\phi}\left(2^{-p}\omega\right)$$

Scaling equation

Iterating:

$$\hat{\Phi}(2\omega) = \frac{1}{\sqrt{2}}\hat{h}(\omega)\hat{\Phi}(\omega) \rightarrow$$

$$\hat{\Phi}(\omega) = \frac{1}{\sqrt{2}}\hat{h}\left(\frac{\omega}{2}\right)\hat{\Phi}\left(\frac{\omega}{2}\right), \qquad \hat{\Phi}\left(\frac{\omega}{2}\right) = \frac{1}{\sqrt{2}}\hat{h}\left(\frac{\omega}{4}\right)\hat{\Phi}\left(\frac{\omega}{4}\right) \quad \rightarrow \ldots\ldots \hat{\Phi}\left(2^{-p+1}\omega\right) = \hat{h}\left(2^{-p}\omega\right)\hat{\Phi}\left(2^{-p}\omega\right)$$

replacing in the expression above for all values of p up to P:

$$\hat{\Phi}(\omega) = \left(\frac{1}{\sqrt{2}}\right)^2 \hat{\Phi}\left(\frac{\omega}{4}\right)\hat{h}\left(\frac{\omega}{4}\right)\hat{h}\left(\frac{\omega}{2}\right)$$

....

$$\hat{\Phi}(\omega) = \prod_{p=1}^{P}\frac{\hat{h}\left(2^{-p}\omega\right)}{\sqrt{2}}\hat{\Phi}(2^{-P}\omega)$$

If $\hat{\varphi}(\omega)$ is continuous at ω=0 then

$$\lim_{P\rightarrow+\infty}\left(\hat{\Phi}\left(2^{-p}\omega\right)\right) = \hat{\Phi}(0) \rightarrow$$

$$\hat{\Phi}(\omega) = \prod_{p=1}^{+\infty}\frac{\hat{h}\left(2^{-p}\omega\right)}{\sqrt{2}}\hat{\Phi}(0)$$

→ find the necessary and sufficient conditions on ^h(ω) to guarantee that this infinite product is the F-transform of a scaling function

Conjugate Mirror Filters

Teorem 7.2 (Mallat&Meyer)

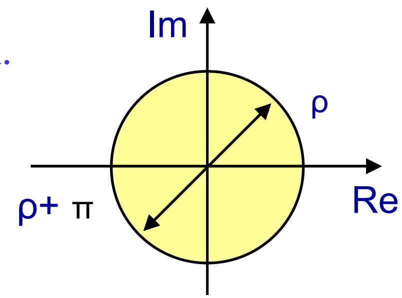
Let ϕ *in* $L^2(R)$ be an integrable scaling function. The F-series of *h[n]* satisfies

$$(2) \qquad \forall\omega \qquad \left|\hat{h}(\omega)\right|^2 + \left|\hat{h}(\omega+\pi)\right|^2 = 2 \qquad and \qquad \hat{h}(0) = \sqrt{2} \qquad \text{CMF}$$

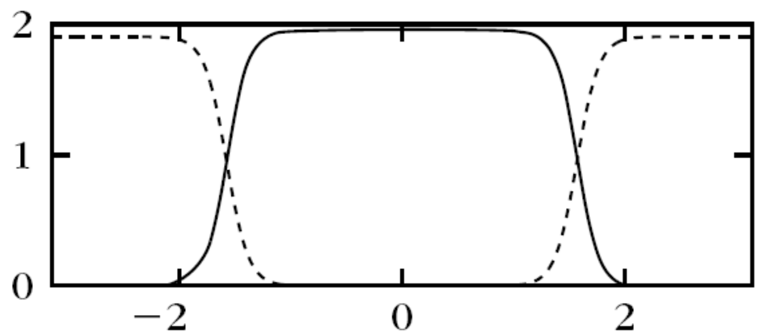
Conversely, if $\hat{h}(\omega)$ is 2π periodic and continuously differentiable in a neighborhood of $\omega=0$, if it satisfies (2) and if

$$\inf_{\omega\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]} \left|\hat{h}(\omega)\right| > 0$$

Then, $\hat{\Phi}(\omega) = \prod_{p=1}^{+\infty} \frac{\hat{h}\left(2^{-p}\omega\right)}{\sqrt{2}} \hat{\Phi}(0)$ is the F-transform of a scaling function.



CMF property



The solid line gives $|\hat{h}(\omega)|^2$ on $[-\pi,\pi]$ for a cubic spline multiresolution. The dotted line corresponds to $|\hat{g}(\omega)|^2$.

Conjugate mirror filters

Table 7.1 Conjugate Mirror Filters $h[n]$ for Linear Splines $m = 1$ and Cubic Splines $m = 3$

	n	$h[n]$		n	$h[n]$
$m = 1$	0	0.817645956	$m = 3$	5, −5	0.042068328
	1, −1	0.397296430		6, −6	−0.017176331
	2, −2	−0.069101020		7, −7	−0.017982291
	3, −3	−0.051945337		8, −8	0.008685294
	4, −4	0.016974805		9, −9	0.008201477
	5, −5	0.009990599		10, −10	−0.004353840
	6, −6	−0.003883261		11, −11	−0.003882426
	7, −7	−0.002201945		12, −12	0.002186714
	8, −8	0.000923371		13, −13	0.001882120
	9, −9	0.000511636		14, −14	−0.001103748
	10, −10	−0.000224296		15, −15	−0.000927187
	11, −11	−0.000122686		16, −16	0.000559952
$m = 3$	0	0.766130398		17, −17	0.000462093
	1, −1	0.433923147		18, −18	−0.000285414
	2, −2	−0.050201753		19, −19	−0.000232304
	3, −3	−0.110036987		20, −20	0.000146098
	4, −4	0.032080869			

Note: The coefficients below 10^{-4} are not given.

What about wavelets?

- Orthonormal wavelets carry the details needed to increase the resolution of a signal approximation.
- The approximations of f at scales 2^j and $2^{(j+1)}$ are respectively equal to its orthogonal projections in V_j and V_{j+1}
- We know that V_{j+1} is included in V_j
- Let W_{j+1} be the *orthogonal complement* of V_{j+1} in V_j

$$V_{j-1} = V_j \oplus W_j$$

- The orthogonal projection of f on V_j can be decomposed as follows

$$PV_{j-1}f = PV_j f + PW_j f$$

- The complement $PW_{j+1}f$ provides the details that appear at scale j but disappear at the next coarser scale.
- Next theorem will show that basis for W_j can be constructed by scaling and translating a wavelet ψ

Corresponding orthogonal wavelet family

- Theorem 7.3 [Mallat&Meyer]

Let ϕ be a scaling function and h the corresponding CMF. Let Ψ be such that

$$\hat{\Psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{\omega}{2}\right)\hat{\Phi}\left(\frac{\omega}{2}\right)$$

with

$$\hat{g}(\omega) = e^{-j\omega}\hat{h}^*(\omega+\pi)$$

Let us denote

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j n}{2^j}\right)$$

For any scale, $\{\Psi_{j,n}\}_{j \text{ in } Z}$ is an orthonormal basis for W_j.

For all j, $\left\{\psi_{j,n}\right\}_{j,n\in\mathbb{Z}^2}$ is an orthonormal basis for L^2.

Signal domain $\hat{g}(\omega) = e^{-j\omega}\hat{h}^*(\omega+\pi) \leftrightarrow g(z) = z^{-1}h(-z^{-1}) \leftrightarrow g[n] = (-1)^{1-n}h[1-n]$

Corresponding orthogonal wavelet family

- Lemma 7.1. The family $\{\psi_{j,n}\}_{n\in Z}$ is an orthonormal basis for W_j iif

$$\left|\hat{g}(\omega)\right|^2+\left|\hat{g}(\omega+\pi)\right|^2=2$$
and
$$\hat{g}(\omega)\hat{h}^*(\omega)+\hat{g}(\omega+\pi)\hat{h}^*(\omega+\pi)=2$$

- Furthermore

$$V_{j-1}=V_j+W_j\rightarrow\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right)\in W_1\subset V_0$$

since $\{\varphi(t-n)\}_{n\in Z}$ is an ortonormal basis of $V_0\rightarrow$

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right)=\sum_{n=-\infty}^{+\infty}g[n]\varphi(t-n)$$ with

$$g[n]=\left\langle\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right),\varphi(t-n)\right\rangle$$

  - The orthogonal wavelets **carry the details lost going from scale j to scale j+1**
  - Wavelets are the **basis functions for W_j**
  - The details at scale j are obtained by **projecting the signal onto the wavelet family $\psi_{j,n}$**

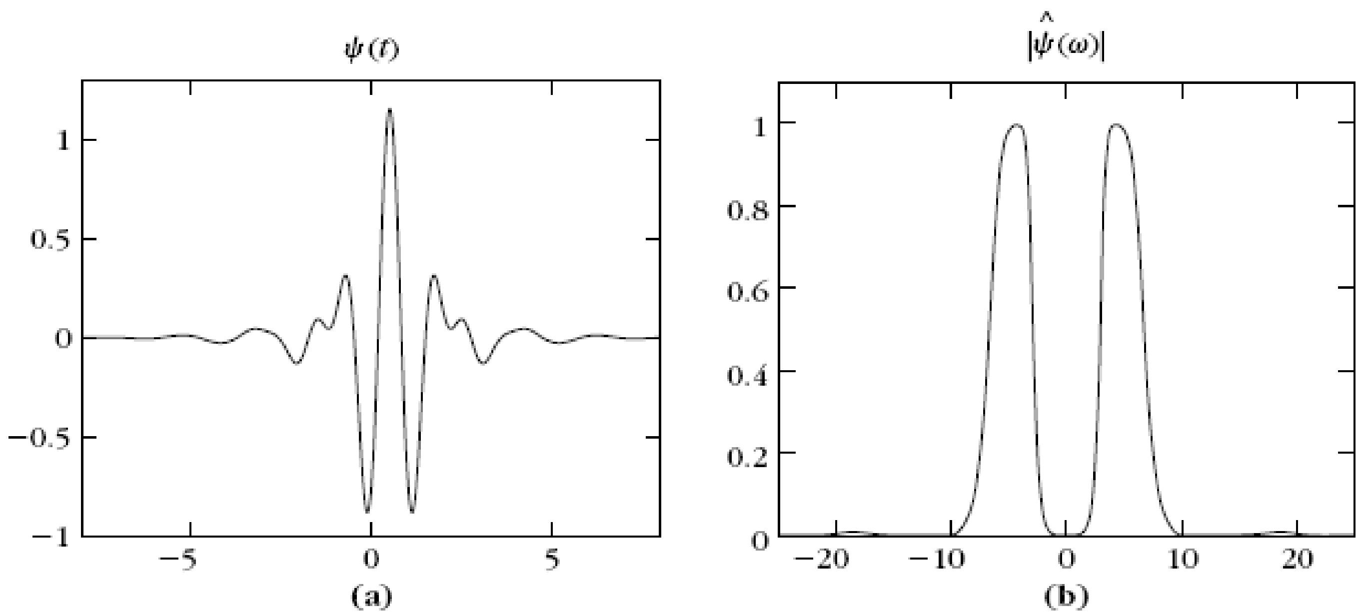
Summary

- Approximation function at scale 2^j: $$P_{V_j} f = \sum_{n=-\infty}^{+\infty} \langle f, \varphi_{j,n} \rangle \varphi_{j,n}$$
- Details ("residual" functions) at scale 2^j: $$P_{W_j} f = \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}$$
- Wavelet representation: $$f = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}$$
- If the basis is orthogonal, the scaling function characterizes the multi-resolution completely

Scaling function φ -> h[n] -> g[n] -> wavelet ψ

Example

- Battle-Lemarié cubic spline wavelet and its spectrum



Example

- Property: for any ψ that can generate an orthonormal family, one can verify that

$$\forall \omega \in \mathbb{R} - \{0\}, \quad \sum_{j=-\infty}^{+\infty} \left|\hat{\psi}\left(2^j \omega\right)\right|^2 = 1$$

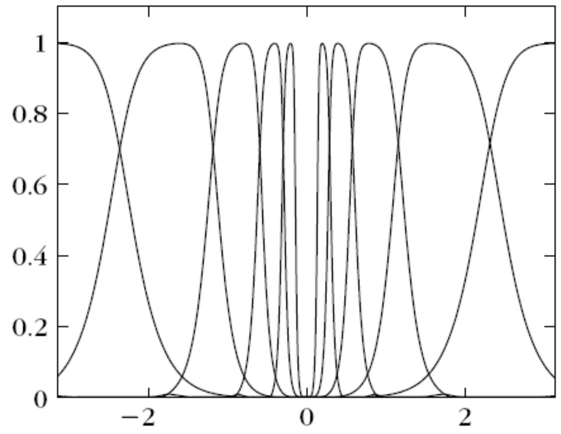


FIGURE 7.6

Graph of $|\hat{\psi}(2^j\omega)|^2$ for the cubic spline Battle-Lemarié wavelet, with $1 \leqslant j \leqslant 5$ and $\omega \in [-\pi, \pi]$.

Example of wavelet analysis

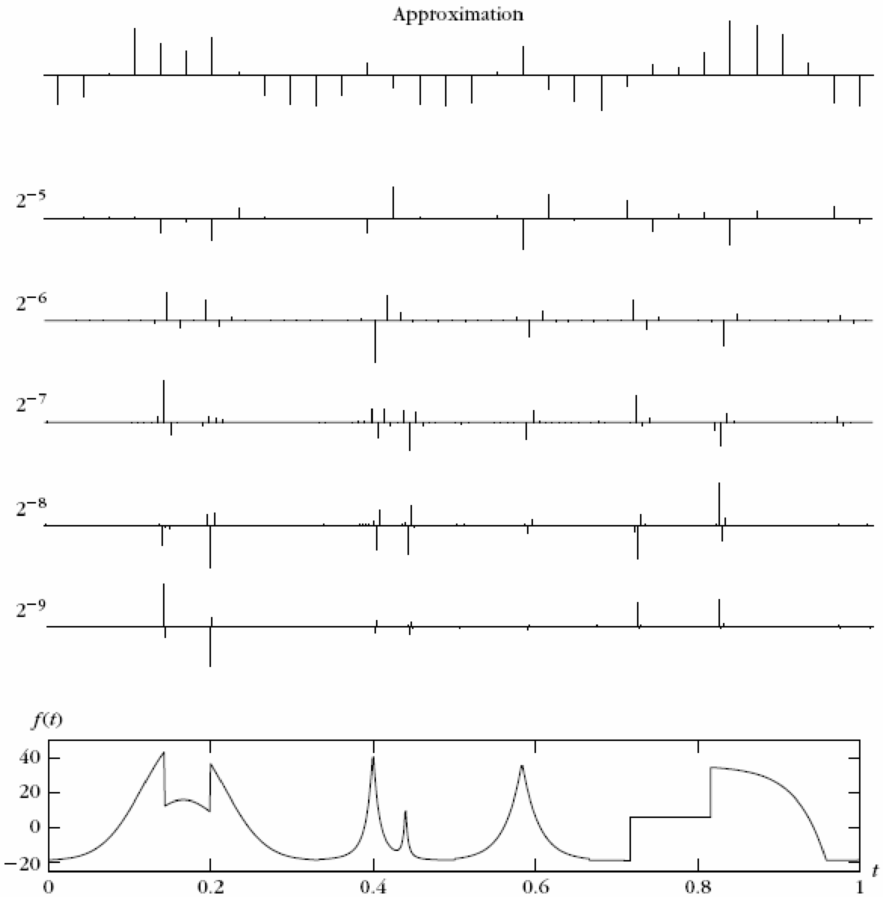


FIGURE 7.7

Wavelet coefficients $d_j[n] = \langle f, \psi_{j,n} \rangle$ calculated at scales 2^j with the cubic spline wavelet. Each up or down Dirac gives the amplitude of a positive or negative wavelet coefficient. At the top is the remaining coarse-signal approximation $a_J[n] = \langle f, \phi_{J,n} \rangle$ for $J = -5$.

Warning

- Each CMF generates a wavelet orthonormal bases
- Does any wavelet orthonormal bases correspond to a multiresolution approximation and CMF? It depends on the support:
  - If ψ has compact support than it corresponds to a multiresolution approximation [Lemarié]
  - However, there exists "pathological" wavelets that decay as $|t|^{-1}$ that cannot be derived from any multiresolution approximation

Classes of wavelet bases

- Wavelets are interesting for applications for their ability to represent signals with **few non zero coefficients**
- The best basis for an application is the one that maximizes the number of zero or close to zero coefficients. This depends on
  - The regularity of f
  - The number of vanishing moments of the wavelet
  - The size of its support
- The constraints on the wavelet translate to **design rules for the filter g[n], thus h[n]**
  - Thus, we need conditions on ^h(ω)

Wavelet properties

- Vanishing moments
  - The wavelet has p vanishing moments if

$$\int_{-\infty}^{+\infty} t^k \psi(t)\,dt = 0 \qquad \text{for} \qquad 0 \le k < p \qquad (3)$$

  - The number of vanishing moments is equal to the multiplicity of zeros of h^(ω) in π or, equivalently, the number of vanishing derivatives of ^ψ in zero
- Theorem 7.4: Vanishing moments

Let φ and ψ be a scaling function and a wavelet that generate an orthonormal basis. Suppose that $|\psi(t)| = O((1+t^2)^{-p/2-1})$ *and* $|\varphi(t)| = O((1+t^2)^{-p/2-1})$*. The four following statements are equivalent*

1. *The wavelet ψ has p vanishing moments*
2. $\hat{\psi}(\omega)$ *and its first p-1 derivatives are zero at ω=0*
3. *^h(ω) and its first p-1 derivatives are zero at ω=π*
4. *for any 0≤k<p* $q_k(t) = \sum_{n=-\infty}^{+\infty} n^k \varphi(t-n)$ is a polynomial of degree k

hints of the proof

- Point 1. The decay of $|\varphi(t)|$ *and* $|\psi(t)|$ *imply that* $|\hat{\varphi}(\omega)|$ *and* $|\hat{\psi}(\omega)|$ *are p-times differentiable*
- Point 2. The k-th order derivative of $\hat{\psi}^{(k)}(\omega)$ is the F-transform of $(-it)^k \psi(t)$ thus

$$\hat{\psi}^{(k)}(0) = \int_{-\infty}^{+\infty} (-it)^k \psi(t)\, dt. \qquad (4)$$

(4) is equivalent to (3), which proves 2.

- Point 3.

$$\hat{\Psi}(\omega) = \frac{1}{\sqrt{2}} \hat{g}\left(\frac{\omega}{2}\right) \hat{\Phi}\left(\frac{\omega}{2}\right) \qquad \hat{g}(\omega) = e^{-j\omega} \hat{h}^*(\omega+\pi) \qquad \text{thus}$$

$$\hat{\Psi}(2\omega) = \frac{1}{\sqrt{2}} \hat{g}(\omega) \hat{\Phi}(\omega) = e^{-i\omega} \hat{h}^*(\omega+\pi) \hat{\Phi}(\omega)$$

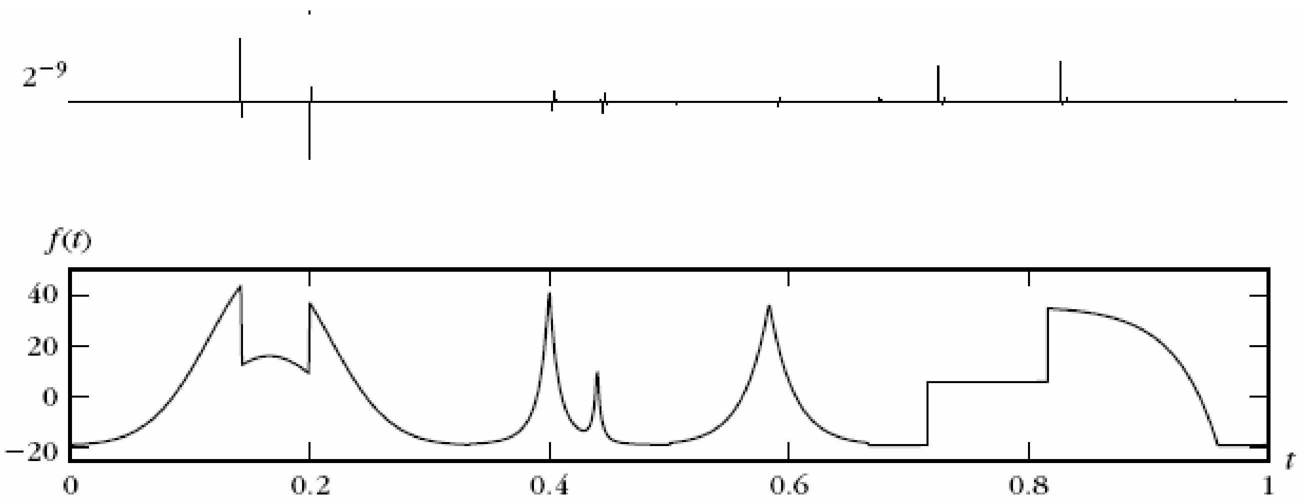
since $\hat{\Phi}(0) \neq 0$ by differentiating this expression we prove that 2. is equivalent to 3.

- Finally, it is proved that 4. is equivalent to 1. and viceversa.

hints of the proof

Let us now prove that (4) implies (1). Since ψ is orthogonal to $\{\phi(t-n)\}_{n\in\mathbb{Z}}$, it is also orthogonal to the polynomials q_k for $0 \leqslant k < p$. This family of polynomials is a basis of the space of polynomials of degree at most $p-1$. Thus, ψ is orthogonal to any polynomial of degree $p-1$ and in particular to t^k for $0 \leqslant k < p$. This means that ψ has p vanishing moments.

A wavelet with p vanishing moments **kills polynomials up to degree p**



Wavelet properties

- Support
  - The larger the support, the more the singularities will spread along scales: it should be **as short as possible**

  BUT a wavelet with *p* vm will have a support at least *2p-1* -> trade-off

- **Theorem 7.5:** *Compact Support*. The scaling function has a compact support if and only if *h* has a compact support and their supports are equal. If the support of *h* and φ is $[N_1,N_2]$, then the support of ψ is $[(N_1-N_2+1)/2,\ (N_1-N_2+1)/2]$.

$$h[n]=\frac{1}{\sqrt{2}}\left\langle \phi\left(\frac{t}{2}\right), \phi(t-n)\right\rangle,$$

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right)=\sum_{n=-\infty}^{+\infty} h[n]\,\phi(t-n).$$

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right)=\sum_{n=-\infty}^{+\infty} g[n]\,\phi(t-n)=\sum_{n=-\infty}^{+\infty}(-1)^{1-n}\,h[1-n]\,\phi(t-n).$$

If the supports of ϕ and h are equal to $[N_1,N_2]$, the sum on the right side has a support equal to $[N_1-N_2+1, N_2-N_1+1]$. Thus, ψ has a support equal to $[(N_1-N_2+1)/2, (N_2-N_1+1)/2]$. ■

Properties

- Support
  - To minimize the size of the support of the wavelet, we must synthesize conjugate mirror filters with *as few nonzero coefficients as possible*
  - However, the constraints imposed on orthogonal wavelets imply that if has *p* vanishing moments, then its support is at least of size $2p\text{-}1$ → trade off
  - Daubechies wavelets are optimal in the sense that they have a minimum size support for a given number of vanishing moments
    - If f has **few isolated singularities** and is very regular between singularities, we must choose a wavelet with **many** vanishing moments to produce a large number of small wavelet coefficients $<f, \psi_{j,n}>$. If the density of singularities increases, it might be better to decrease the size of its support at the cost of reducing the number of vanishing moments. Indeed, **wavelets that overlap the singularities create high-amplitude coefficients**.
- Regularity
  - The regularity or *smoothness* has mostly a cosmetic influence on the error introduced by quantizing or thresholding the coefficients. Such operation introduces a noise which is less visible if it is smooth. Better quality is reached with smoother wavelets
    - The Haar wavelet is not a good choice

Popular wavelet families

- Shannon, Meyer, Haar, and Battle-Lemarié Wavelets
  - Starting point

$$\hat{\Psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{\omega}{2}\right)\hat{\Phi}\left(\frac{\omega}{2}\right) \qquad \hat{g}(\omega) = e^{-j\omega}\hat{h}^{*}(\omega+\pi)$$

$$\hat{\Psi}(2\omega) = \frac{1}{\sqrt{2}}\hat{g}(\omega)\hat{\Phi}(\omega) = e^{-i\omega}\hat{h}^{*}(\omega+\pi)\hat{\Phi}(\omega)$$

$$\Downarrow$$

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) = \frac{1}{\sqrt{2}}\exp\left(\frac{-i\omega}{2}\right)\hat{h}^{*}\left(\frac{\omega}{2}+\pi\right)\hat{\phi}\left(\frac{\omega}{2}\right). \qquad (7.82)$$

Shannon wavelets

Shannon Wavelet

The Shannon wavelet is constructed from the Shannon multiresolution approximation, which approximates functions by their restriction to low-frequency intervals. It corresponds to $\hat{\phi} = \mathbf{1}_{[-\pi,\pi]}$ and $\hat{h}(\omega) = \sqrt{2}\,\mathbf{1}_{[-\pi/2,\pi/2]}(\omega)$ for $\omega \in [-\pi, \pi]$. We derive from (7.82) that

$$\hat{\psi}(\omega) = \begin{cases} \exp(-i\omega/2) & \text{if } \omega \in [-2\pi, -\pi] \cup [\pi, 2\pi] \\ 0 & \text{otherwise,} \end{cases} \qquad \textbf{(7.83)}$$

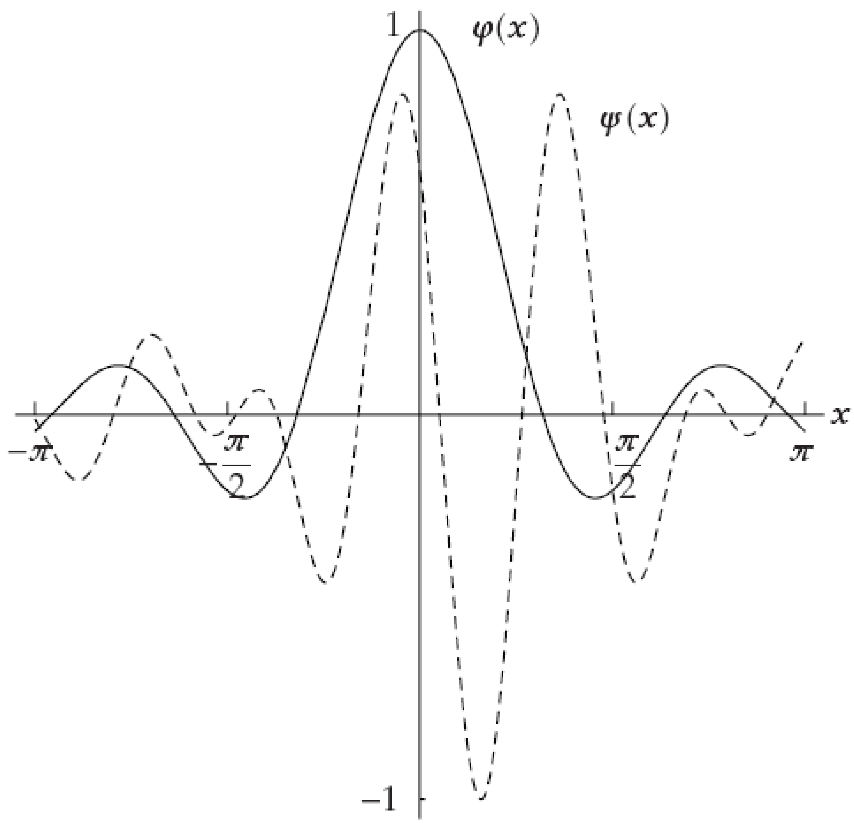
and thus,

$$\psi(t) = \frac{\sin 2\pi(t-1/2)}{2\pi(t-1/2)} - \frac{\sin \pi(t-1/2)}{\pi(t-1/2)}.$$

This wavelet is $\mathbf{C}^{\infty}$ but has a slow asymptotic time decay. Since $\hat{\psi}(\omega)$ is zero in the neighborhood of $\omega = 0$, all its derivatives are zero at $\omega = 0$. Thus, Theorem 7.4 implies that ψ has an infinite number of vanishing moments.

Since $\hat{\psi}(\omega)$ has a compact support we know that $\psi(t)$ is $\mathbf{C}^{\infty}$. However, $|\psi(t)|$ decays only like $|t|^{-1}$ at infinity because $\hat{\psi}(\omega)$ is discontinuous at $\pm\pi$ and $\pm 2\pi$.

Shannon wavelets



Shannon scaling function (continuous) and wavelet (dashed) lines.

Meyer wavelets

Meyer Wavelets

A Meyer wavelet [375] is a frequency band-limited function that has a Fourier transform that is smooth, unlike the Fourier transform of the Shannon wavelet. This smoothness provides a much faster asymptotic decay in time. These wavelets are constructed with conjugate mirror filters $\hat{h}(\omega)$ that are $\mathbf{C}^n$ and satisfy

$$\hat{h}(\omega) = \begin{cases} \sqrt{2} & \text{if } \omega \in [-\pi/3, \pi/3] \\ 0 & \text{if } \omega \in [-\pi, -2\pi/3] \cup [2\pi/3, \pi]. \end{cases} \tag{7.84}$$

The only degree of freedom is the behavior of $\hat{h}(\omega)$ in the transition bands $[-2\pi/3, -\pi/3] \cup [\pi/3, 2\pi/3]$. It must satisfy the quadrature condition

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2, \tag{7.85}$$

and to obtain $\mathbf{C}^n$ junctions at $|\omega| = \pi/3$ and $|\omega| = 2\pi/3$, the n first derivatives must vanish at these abscissa. One can construct such functions that are $\mathbf{C}^\infty$.

The scaling function $\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} 2^{-1/2}\, \hat{h}(2^{-p}\omega)$ has a compact support and one can verify that

$$\hat{\phi}(\omega) = \begin{cases} 2^{-1/2}\, \hat{h}(\omega/2) & \text{if } |\omega| \leqslant 4\pi/3 \\ 0 & \text{if } |\omega| > 4\pi/3. \end{cases} \tag{7.86}$$

Meyer wavelets

The resulting wavelet (7.82) is

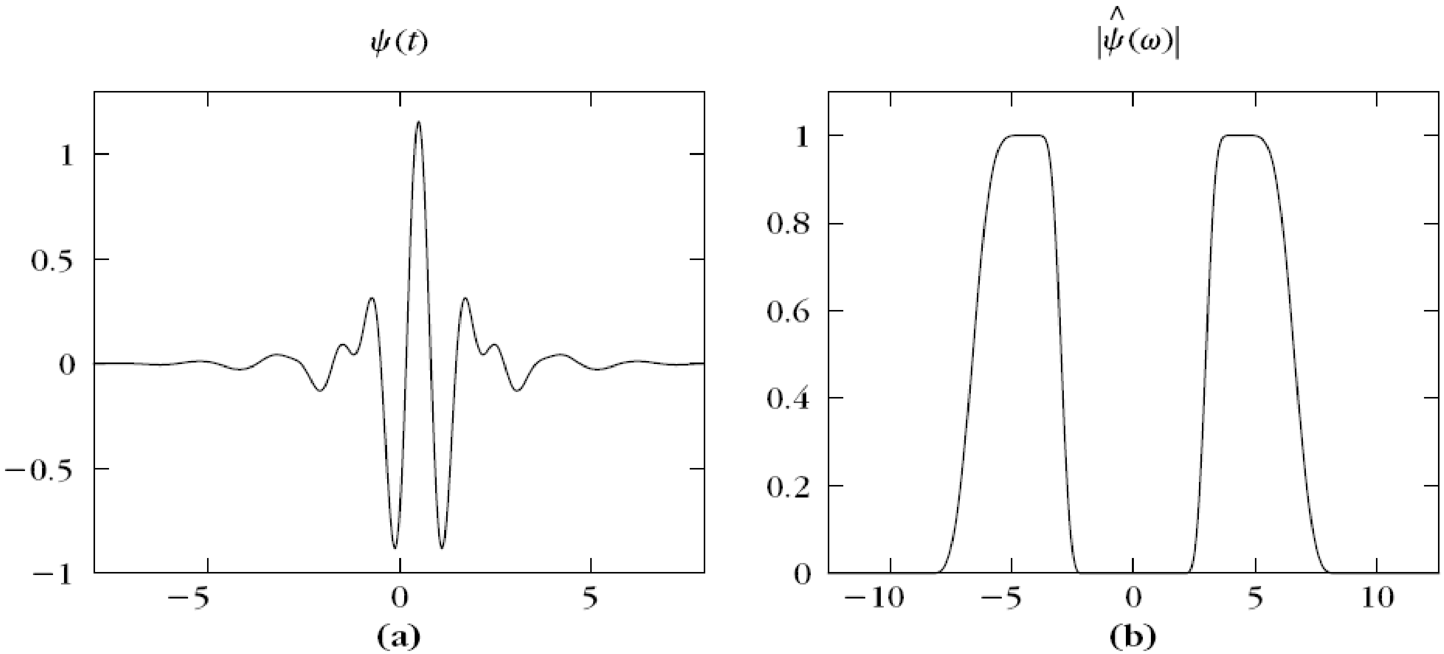
$$\hat{\psi}(\omega)=\begin{cases}0 & \text{if } |\omega|\leqslant 2\pi/3\\ 2^{-1/2}\,\hat{g}(\omega/2) & \text{if } 2\pi/3\leqslant|\omega|\leqslant 4\pi/3\\ 2^{-1/2}\ \exp(-i\omega/2)\,\hat{h}(\omega/4) & \text{if } 4\pi/3\leqslant|\omega|\leqslant 8\pi/3\\ 0 & \text{if } |\omega|>8\pi/3.\end{cases} \tag{7.87}$$

The functions ϕ and ψ are $\mathbf{C}^{\infty}$ because their Fourier transforms have a compact support. Since $\hat{\psi}(\omega)=0$ in the neighborhood of $\omega=0$, all its derivatives are zero at $\omega=0$, which proves that ψ has an infinite number of vanishing moments.

If $\hat{h}$ is $\mathbf{C}^n$, then $\hat{\psi}$ and $\hat{\phi}$ are also $\mathbf{C}^n$. The discontinuities of the $(n+1)^{\text{th}}$ derivative of $\hat{h}$ are generally at the junction of the transition band $|\omega|=\pi/3\,,\ 2\pi/3$, in which case one can show that there exists A such that

$$|\phi(t)|\leqslant A\,(1+|t|)^{-n-1} \quad \text{and} \quad |\psi(t)|\leqslant A\,(1+|t|)^{-n-1}.$$

Meyer wavelet: example



Haar wavelets

Haar Wavelets

The Haar basis is obtained with a multiresolution of piecewise constant functions. The scaling function is $\phi = \mathbf{1}_{[0,1]}$. The filter $h[n]$ given in (7.46) has two nonzero coefficients equal to $2^{-1/2}$ at $n = 0$ and $n = 1$. Thus,

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} (-1)^{1-n}\, h[1-n]\, \phi(t-n) = \frac{1}{\sqrt{2}}\Big(\phi(t-1) - \phi(t)\Big),$$

so

$$\psi(t) = \begin{cases} -1 & \text{if } 0 \leqslant t < 1/2 \\ \;\;1 & \text{if } 1/2 \leqslant t < 1 \\ \;\;0 & \text{otherwise.} \end{cases} \qquad \textbf{(7.90)}$$

The Haar wavelet has the shortest support among all orthogonal wavelets. It is not well adapted to approximating smooth functions because it has only one vanishing moment.

reminder: $$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} g[n]\, \phi(t-n) = \sum_{n=-\infty}^{+\infty} (-1)^{1-n}\, h[1-n]\, \phi(t-n).$$

Battle-Lemarié wavelets

Battle-Lemarié Wavelets

Polynomial spline wavelets introduced by Battle [99] and Lemarié [345] are computed from spline multiresolution approximations. The expressions of $\hat{\phi}(\omega)$ and $\hat{h}(\omega)$ are given, respectively, by (7.18) and (7.48). For splines of degree m, $\hat{h}(\omega)$ and its first m derivatives are zero at $\omega = \pi$. Theorem 7.4 derives that ψ has $m+1$ vanishing moments. It follows from (7.82) that

$$\hat{\psi}(\omega) = \frac{\exp(-i\omega/2)}{\omega^{m+1}} \sqrt{\frac{S_{2m+2}(\omega/2+\pi)}{S_{2m+2}(\omega)\, S_{2m+2}(\omega/2)}}.$$

This wavelet ψ has an exponential decay. Since it is a polynomial spline of degree m, it is $m-1$ times continuously differentiable. Polynomial spline wavelets are less regular than Meyer wavelets but have faster time asymptotic decay. For m odd, ψ is symmetric about $1/2$. For m even, it is antisymmetric about $1/2$. Figure 7.5 gives the graph of the cubic spline wavelet ψ corresponding to $m=3$. For $m=1$, Figure 7.9 displays linear splines ϕ and ψ. The properties of these wavelets are further studied in [15, 106, 164].

Battle-Lemarié: example

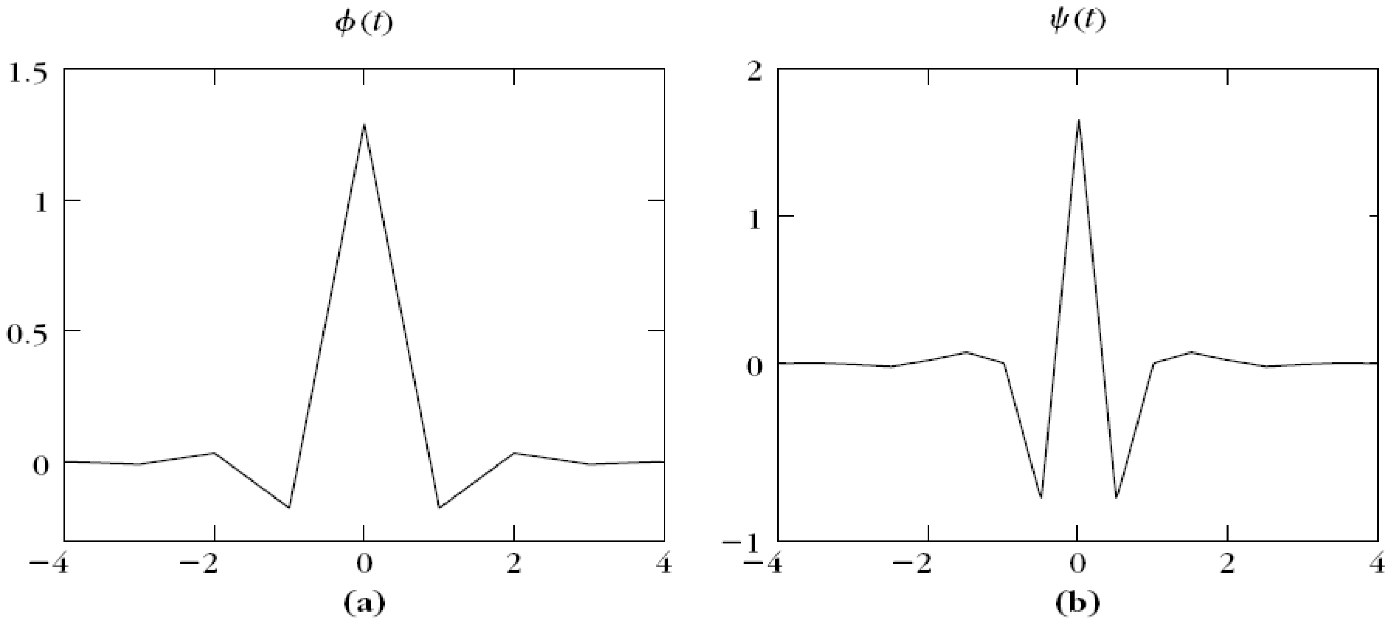


FIGURE 7.9

Linear spline Battle-Lemarié scaling function ϕ **(a)** and wavelet ψ **(b)**.

Daubechies compactly supported wavelets

7.2.3 Daubechies Compactly Supported Wavelets

Daubechies wavelets have a support of minimum size for any given number p of vanishing moments. Theorem 7.5 proves that wavelets of compact support are computed with finite impulse-response conjugate mirror filters h. We consider real causal filters $h[n]$, which implies that $\hat{h}$ is a trigonometric polynomial:

$$\hat{h}(\omega) = \sum_{n=0}^{N-1} h[n]\, e^{-in\omega}.$$

To ensure that ψ has p vanishing moments, Theorem 7.4 shows that $\hat{h}$ must have a zero of order p at $\omega = \pi$. To construct a trigonometric polynomial of minimal size, we factor $(1+e^{-i\omega})^p$, which is a minimum-size polynomial having p zeros at $\omega = \pi$:

$$\hat{h}(\omega) = \sqrt{2}\left(\frac{1+e^{-i\omega}}{2}\right)^p R(e^{-i\omega}). \tag{7.91}$$

The difficulty is to design a polynomial $R(e^{-i\omega})$ of minimum degree m such that $\hat{h}$ satisfies

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega+\pi)|^2 = 2. \tag{7.92}$$

As a result, h has $N = m + p + 1$ nonzero coefficients. Theorem 7.7 by Daubechies [194] proves that the minimum degree of R is $m = p - 1$.

Daubechies compactly supported wavelets

- **Theorem 7.7:** *Daubechies*. A real conjugate mirror filter h, such that $\hat{h}(\omega)$ has p zeroes at π, has at least $2p$ nonzero coefficients. Daubechies filters have $2p$ nonzero coefficients.
- **Theorem 7.9:** *Daubechies*. If is a wavelet with p vanishing moments that generates an orthonormal basis of $\mathbf{L}^2(\mathbb{R})$, then it has a support of size larger than or equal to $2p+1$.

  A Daubechies wavelet has a *minimum-size support* equal to $[-p+1, p]$. The support of the corresponding scaling function is $[0, 2p-1]$.

Daubechies wavelets: example

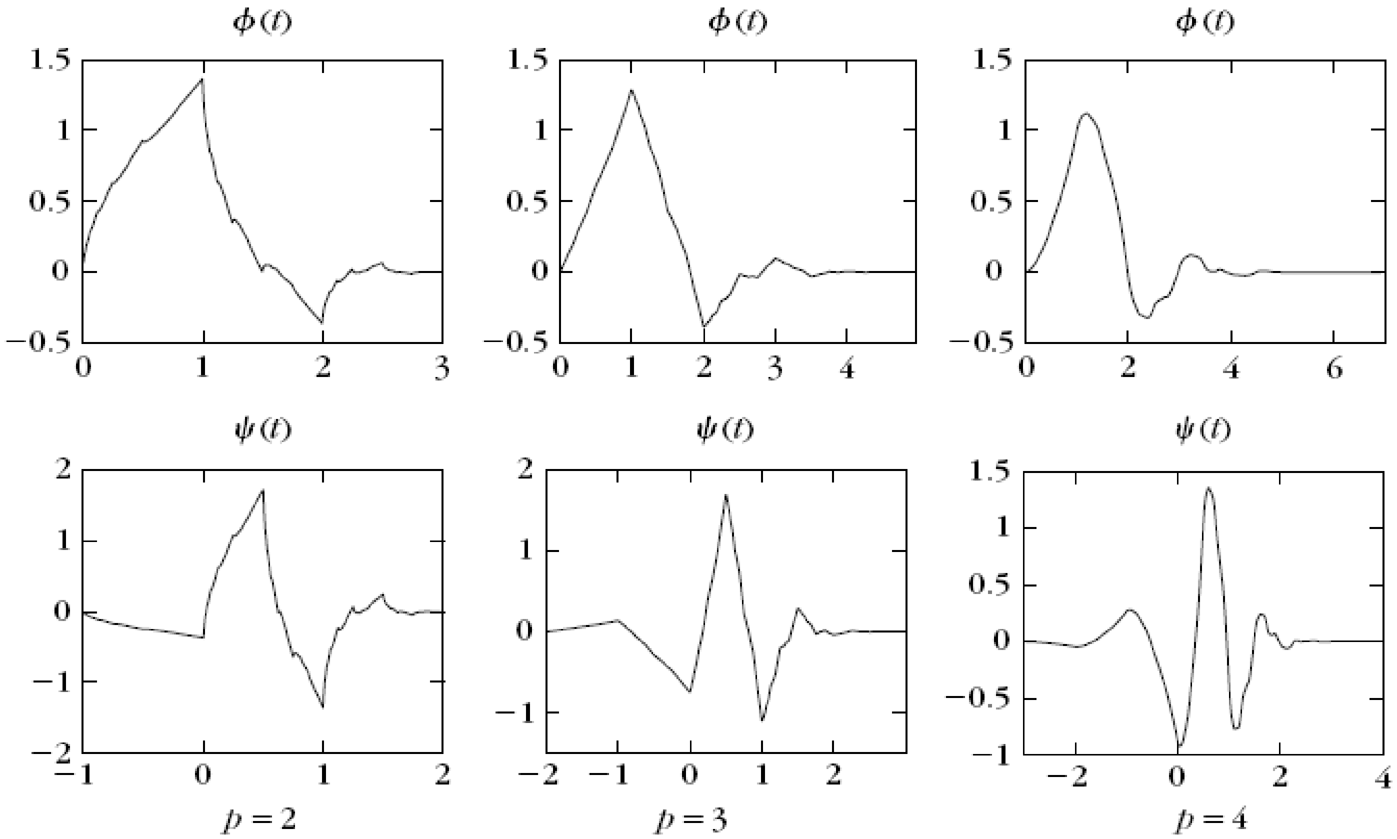


FIGURE 7.10

Daubechies scaling function ϕ and wavelet ψ with p vanishing moments.

Symlets

Symmlets

Daubechies wavelets are very asymmetric because they are constructed by selecting the minimum-phase square root of $Q(e^{-i\omega})$ in (7.97). One can show [51] that filters corresponding to a minimum-phase square root have their energy optimally concentrated near the starting point of their support. Thus, they are highly nonsymmetric, which yields very asymmetric wavelets.

To obtain a symmetric or antisymmetric wavelet, the filter h must be symmetric or antisymmetric with respect to the center of its support, which means that $\hat{h}(\omega)$ has a linear complex phase. Daubechies proved [194] that the Haar filter is the only real compactly supported conjugate mirror filter that has a linear phase. The Daubechies *symmlet* filters are obtained by optimizing the choice of the square root $R(e^{-i\omega})$ of $Q(e^{-i\omega})$ to obtain an almost linear phase. The resulting wavelets still have a minimum support $[-p+1, p]$ with p vanishing moments, but they are more symmetric, as illustrated by Figure 7.11 for $p = 8$. The coefficients of the symmlet filters are in WaveLab. Complex conjugate mirror filters with a compact support and a linear phase can be constructed [352], but they produce complex wavelet coefficients that have real and imaginary parts that are redundant when the signal is real.

Dubechies versus Symlets

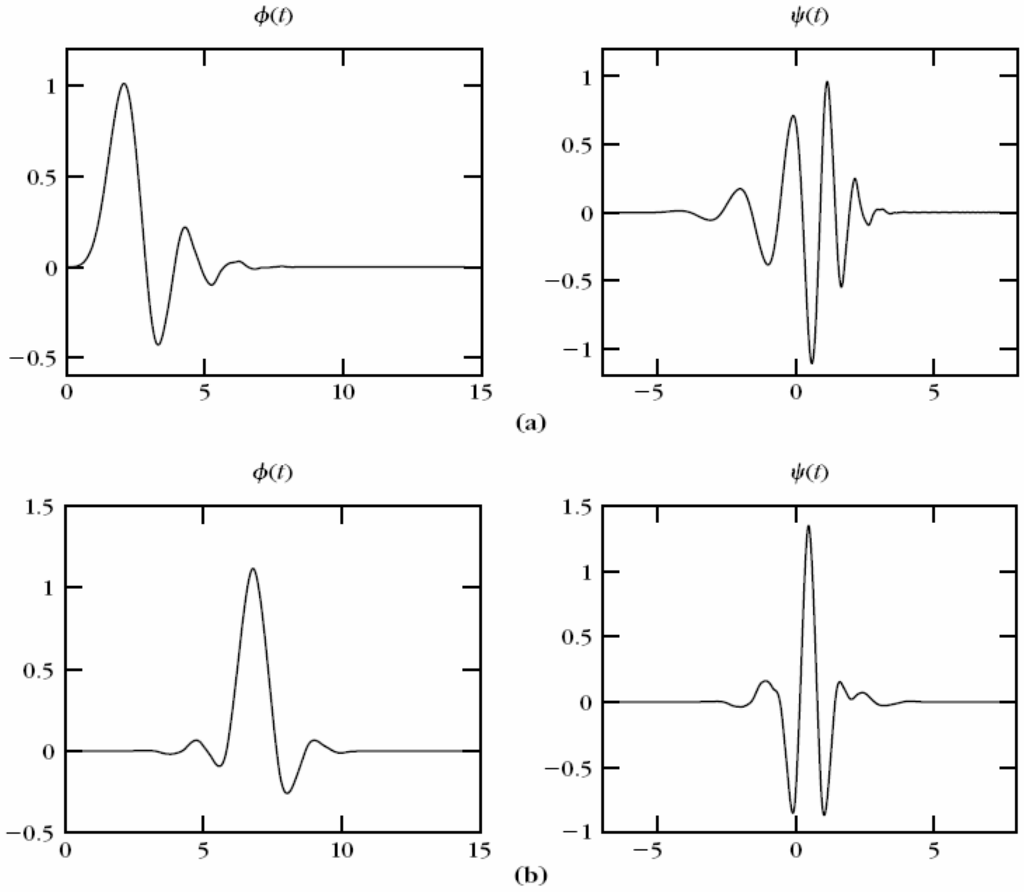


FIGURE 7.11

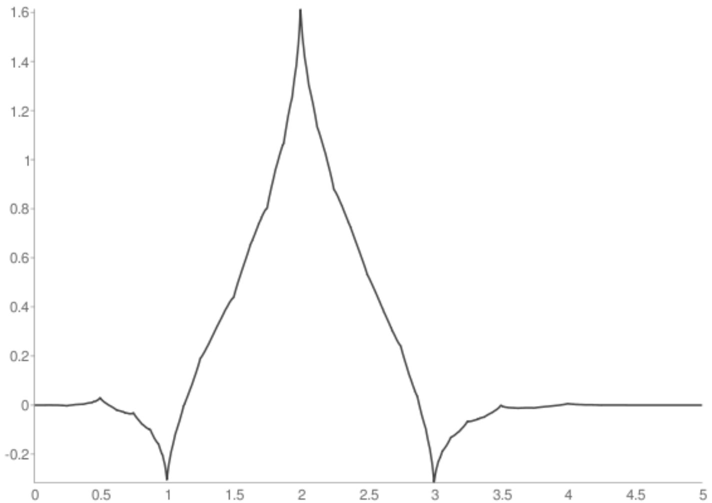
Daubechies **(a)** and symmlet **(b)** scaling functions and wavelets with $p = 8$ vanishing moments.

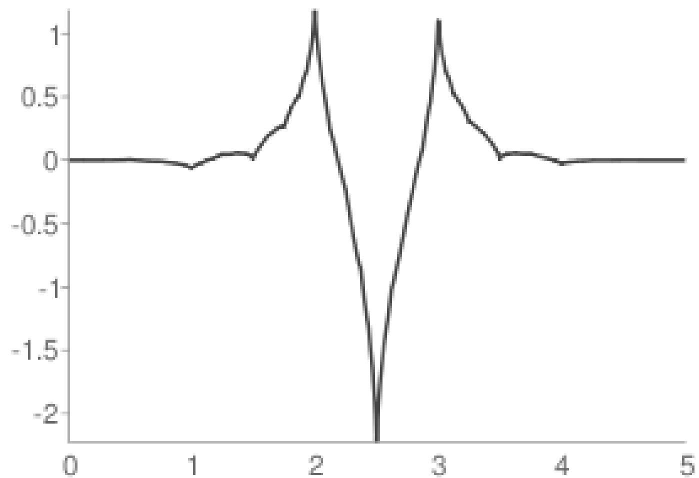
Coiflets

Coiflets

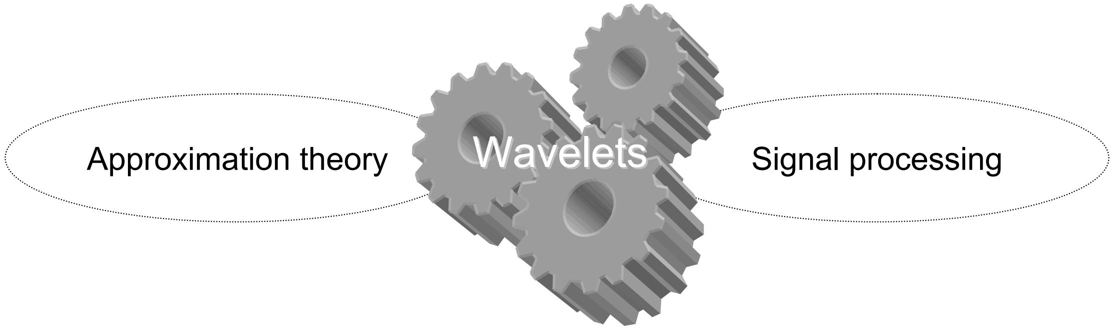
For an application in numerical analysis, Coifman asked Daubechies [194] to construct a family of wavelets ψ that have p vanishing moments and a minimum-size support, with scaling functions that also satisfy

$$\int_{-\infty}^{+\infty} \phi(t)\,dt = 1 \quad \text{and} \quad \int_{-\infty}^{+\infty} t^k\,\phi(t)\,dt = 0 \quad \text{for } 1 \leqslant k < p. \tag{7.99}$$





An approximation tour



- Linear approximation
  - Projects the signal f over M vectors of the ortho-normal basis B which are chosen *a-priori* among the basis B, say the first M

$$f_M = \sum_{n=0}^{M-1} \langle f, \phi_n \rangle \phi_n$$

  - Approximation error

$$\varepsilon[M] = \|f - f_M\|^2 = \sum_{n=M}^{+\infty} |\langle f, \phi_n \rangle|^2$$

choosing the first M vectors amounts to reconstruct f at a given resolution. The convergence properties similar as in the Fourier domain

- Non-linear approximations
  - The M vectors are chosen *a posteriori*

$$f_M = \sum_{n \in I_M} \langle f, \phi_n \rangle \phi_n$$

Approximation error

$$\varepsilon[M] = \|f - f_M\|^2 = \sum_{n \in I_M}^{+\infty} |\langle f, \phi_n \rangle|^2$$

The error can be minimized by choosing the vectors $\langle f, \phi_n \rangle$ corresponding to the highest

In wavelet basis this amounts to an adaptive approximation grid whose *resolution is locally increased where the signal is irregular!*

Adaptive basis choice

- Instead of choosing the basis a-priori, one could choose the *best* basis, depending on the signal
- The basis is chosen to minimize the non linear approximation error of f
- Same problem as the choice of the *optimal basis* for stimulus representation in visual perception
- The optimal basis could be chosen for *classes of signals*, considered as random processes
  - Gaussian processes → Karunen Loeve transform (KLT)
    - Diagonalization of the covariance matrix which removes the inter-dependencies among the samples and results in a set of independent coefficients (i.e. redundancy has been removed)
  - Other kind of processes → no golden rule
    - Images are not Gaussian and not stationary
    - In some cases wavelets do better

Adaptive basis

- Wavelet packets
  - The subband tree is progressively split according to the optimization of a cost function (i.e. rate/distortion)
- Matching pursuit
  - Vectors are progressively selected from a dictionary, while optimizing the signal approximation at each step

- Key issue: a good basis should be able to provide a good description (approximation properties) of the signal while being concise (sparseness properties)
  - Classical approaches: approximation theory, information theory, estimation in noise...
  - Perception based approaches: bring humans into the loop

Wavelet Packets