The What, Why, and How of Wavelets

Chapter. 1

수학과 오서영

Introduction

Wavelet Transform (WT)

Data, functions, operator –- cut off --> Different frequency components

- Study each component with a resolution matched to its scale

- WT of signal evolving in time depends on two variables
  1. scale (or frequency), 2. time

Signal f(t), t : continuous variable

- Interest : frequency content locally in time
- **Ex)** Which notes to play at any given moment

1. Time-Frequency localization

Fourier Transform (FT)

$$(\mathcal{F}f)(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-i\omega t} f(t)\, dt$$ freq content of f

-> high frequency bursts cannot be read off easily from **Ff**

- **Time localization** can be achieved by first windowing **f**
  → Cut off only well-localized slice of **f** + taking its **FT**

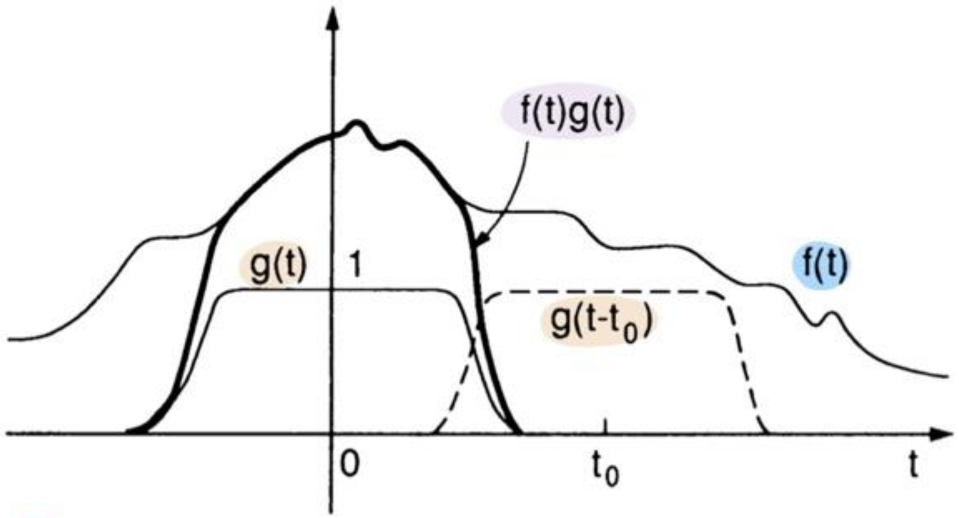
Windowed FT

$$(T^{win} f)(\omega, t) = \int ds\, f(s)\, g(s-t)\, e^{-i\omega s}$$ windowed FT

↪ discrete version where $t = n t_0$, $\omega = m\omega_0$ $m, n \in \mathbb{Z}$, $\omega_0, t_0 > 0$

$$: T^{win}_{m,n}(f) = \int ds\, f(s)\, g(s - n t_0)\, e^{-i m \omega_0 s}$$

1. Time-Frequency localization



— $T^{win}_{m,n}(f)$ (n : fixed)

⟷ Fourier coefficients of $f(\cdot)g(\cdot - nt_0)$

— Change n (shift the slices) nt_0

↪ recovery of all of f from the $T^{win}_{m,n}(f)$

↪ The windowed FT : 1 $f(t) \times g(t)$ window ft   2 Fourier coefficients of $f(t)g(t)$

1, 2 is repeated for translated versions of window $g(t-t_0)$

$g(t-2t_0)$ …

1. Time-Frequency localization

- Window function g

-> compact support and reasonable smoothness -- popular choice --> Gaussian g
-> g : well-concentrated in both time and frequency

- If g, $\hat{g}$ are concentrated around zero

  then T^{win}f can be interpreted loosely as the content of f near time t and freq w

→ The windowed FT provides a description of f in time-freq plane

X : topological space.
$f : X \to \mathbb{R}$ (function)
" $\operatorname{supp} f = cl\{x \in X \mid f(x) \neq 0\}$ "
if supp f is compact set,
f is a function with compact support

2. FT : Analogies and differences with the windowed FT

WT formula

$$(T^{wav} f)(a,b) = |a|^{-1/2} \int dt\, f(t)\, \psi\left(\frac{t-b}{a}\right) \quad \text{where} \int dt\, \psi(t) = 0$$

restrict $a = a_0^m$, $b = n b_0 a_0^m$ with $m, n \in \mathbb{Z}$, $a_0 > 1$, $b_0 > 0$

$$T^{wav}_{m,n}(f) = a_0^{-m/2} \int dt\, f(t)\, \psi(a_0^{-m} t - n b_0)$$

Similarity between WT and windowed FT

: Inner products of f with a family of functions

$$\text{windowed FT} : g^{w,t}(s) = e^{iws} g(s-t)$$

$$\text{WT} : \underbrace{\psi^{a,b}}_{\text{wavelets}}(s) = |a|^{-1/2} \psi\left(\frac{s-b}{a}\right)$$

2. FT : Analogies and differences with the windowed FT

⊕ Typical Choice of ψ : $\psi(t) = (1-t^2)\exp\left(-\frac{t^2}{2}\right)$ second derivative of Gaussian

↪ well localized in both time & frequency = mexican hat function

① As a changes, $\psi^{a,0}(s) = |a|^{-1/2}\,\psi\left(\frac{s}{a}\right)$ cover different frequency ranges

↪ $|a|$: large $\longleftrightarrow$ freq : small

scaling parameter small $\longleftrightarrow$ large

② As b changes, move the time localization center

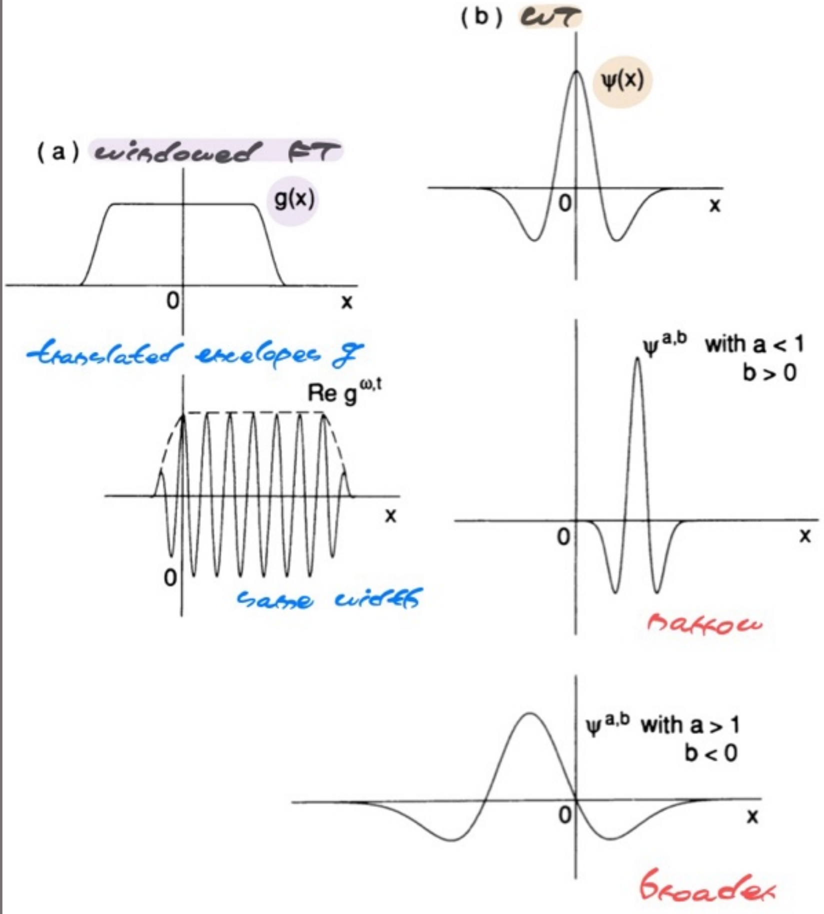
↪ $\psi^{a,b}(s)$ is localized around $s=b$

$\therefore$ ①,② : time − frequency description of f similarity!

2. FT : Analogies and differences with the windowed FT

— difference between WT and windowed FT

↪ analyzing $g^{\omega,t}_{FT}$ and $\psi^{a,b}_{WT}$



① $g^{\omega,t}$

— consist of the same envelop function g

⇒ time location

filled in with higher frequency oscillation

— Regardless of the ω, $g^{\omega,t}$ have the same width

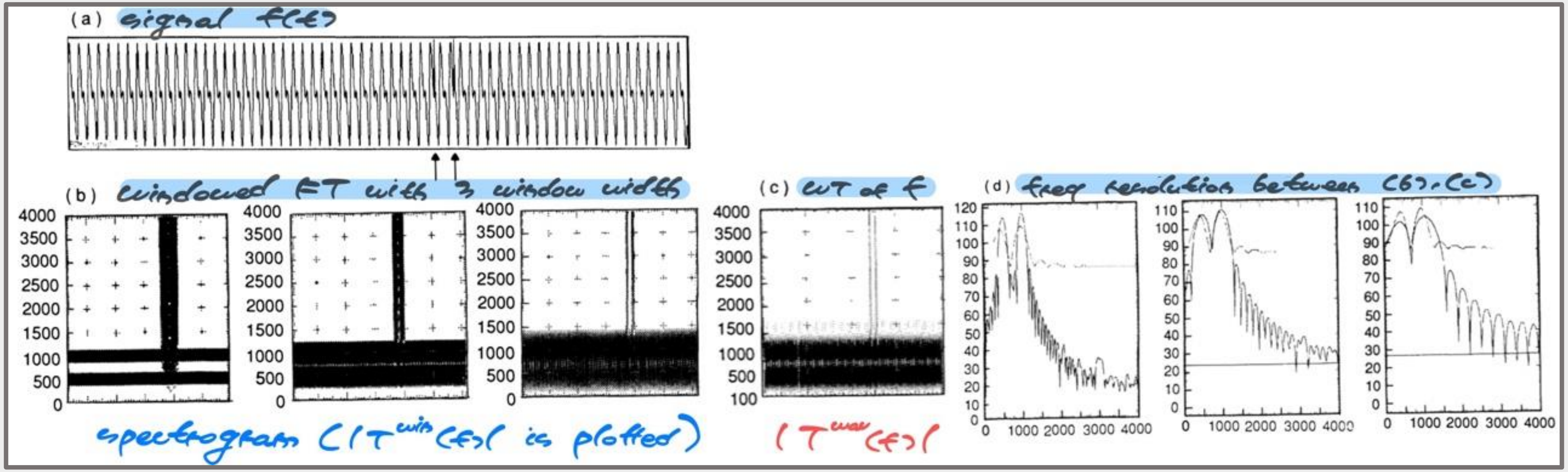
② $\psi^{a,b}$

— $\psi^{a,b}$ have time-widths adapted to their frequency

	freq		$\psi^{a,b}$
1	high	⟷	narrow
2	low	⟷	broader

WT is better able than the windowed FT to zoom in very short lived high frequency phenomena

2. FT : Analogies and differences with the windowed FT



As the windowed width increases, the resolution of two pure tones gets better, But it becomes harder to resolve two pulses.

The frequency resolution for the two pure tones is comparable

Dynamic range of WT is comparable

3. Different types of wavelet transform

The discrete but redundant wavelet transform - frames

- $a = a_0^m$ with $a_0 > 1$ ($m \longleftrightarrow$ width of wavelets)
- b depends on m : 1 narrow wavelets $\longleftrightarrow$ small steps   cover the whole time range
  2 wider $\longleftrightarrow$ larger steps

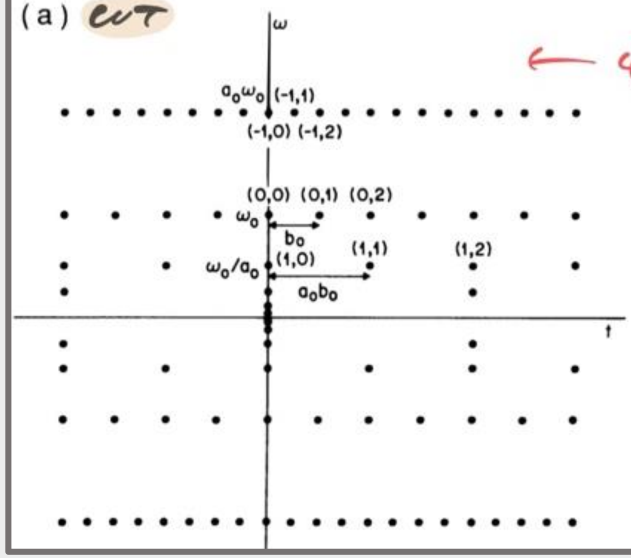
  ↪ $b = n b_0 a_0^m$ where $b_0 > 0$, $n \in \mathbb{Z}$

$\Rightarrow$ discretely labelled wavelets

$$\psi_{m,n}(x) = a_0^{-m/2}\,\psi\left(a_0^{-m}(x - n b_0 a_0^m)\right) = a_0^{-m/2}\,\psi\left(a_0^{-m}x - n b_0\right)$$

⊕ lattices of time-frequency localization

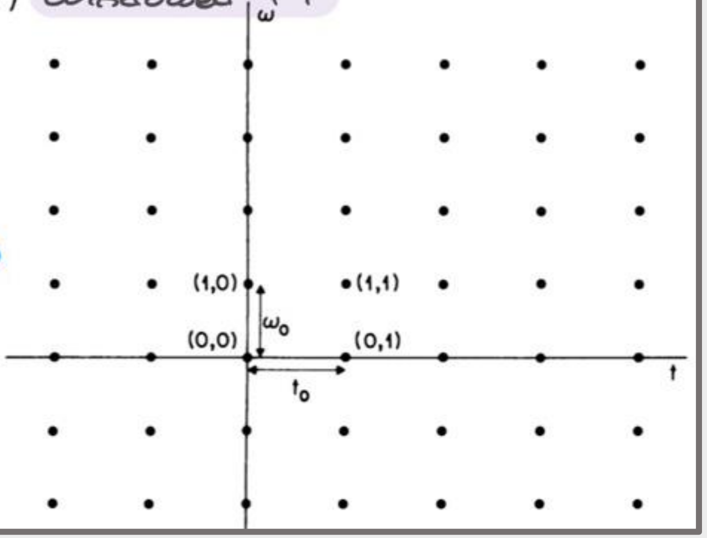
(a) CWT



← $\psi_{m,n}$ is localized around $a_0^m n b_0$ in time

$g_{m,n}$ is localized around $n t_0$ in time, around $m\omega_0$ in freq →

(b) Windowed FT



3. Different types of wavelet transform

Orthonormal wavelet bases : Multiresolution analysis

$\psi_{m,n}$ constitute an orthonormal basis for $L^2(\mathbb{R})$. choose $a_0 = 2,\ b_0 = 1$

$\longrightarrow \exists\, \psi_{m,n}(x) = 2^{-m/2}\,\psi(2^{-m}x - n)$ good time-freq localization

ex

$$\psi(x) = \begin{cases} 1 & 0 \le x < \frac{1}{2} \\ -1 & \frac{1}{2} \le x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Haar function

does not have good time-freq localization

($\because$) FT $\hat{\psi}(\xi)$ decays like $|\xi|^{-1}$ for $\xi \to \infty$

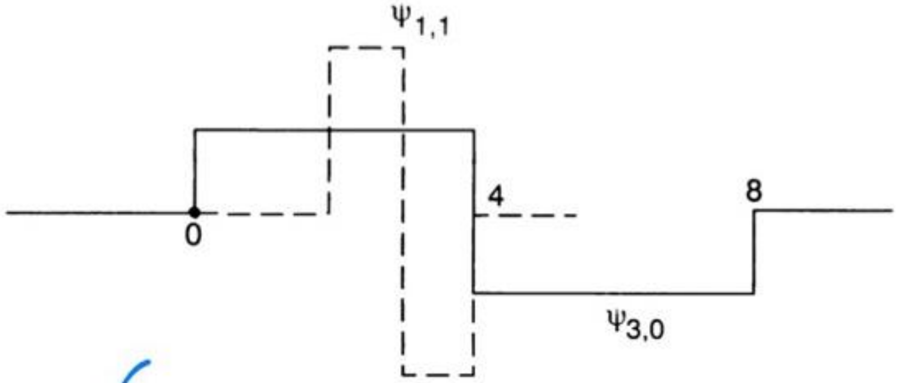
$\hookrightarrow$ proof $\psi_{m,n}(x)$ constitute an orthonormal basis

1. $\psi_{m,n}$ are orthonormal
2. any L^2-function f can be approximated, up to arbitrarily small precision by a finite linear combination of $\psi_{m,n}$

3. Different types of wavelet transform

$\psi_{m,n}$ are orthonormal

— pf 1) $(\psi_{m,n}) = [2^m n, 2^m(n+1)]$, $\langle \psi_{m,n}, \psi_{m,n'} \rangle = \delta_{n,n'}$



if two wavelets have different size

→ overlapping

if $m<m'$, $(\psi_{m,n})$ lies wholly within a region where $\psi_{m',n'}$ is constant

→ narrower wavelet is completely contained in an interval where wider wavelet is constant.

3. Different types of wavelet transform

2 any L^2 − function f can be approximated, up to arbitrarily small precision by a finite linear combination of $\psi_{m,n}$

pf 2) $\forall f$ in $L^2(\mathbb{R})$ can be arbitrarily well approximated by ft with compact support which is piecewise constant on $[\ell 2^{-j}, (\ell+1)2^{-j}]$

So restrict to piecewise constant functions :

assume f to be supported on $[-2^{J_1}, 2^{J_2}]$ and piecewise constant on $[\ell 2^{-J_0}, (\ell+1)2^{-J_0}]$

constant value of $f^0 = f$ on $[\ell 2^{-J_0}, (\ell+1)2^{-J_0}]$ by f^0_ℓ

$f^0 = f^1 + \delta^1$ where f^1 is an approximation to f^0

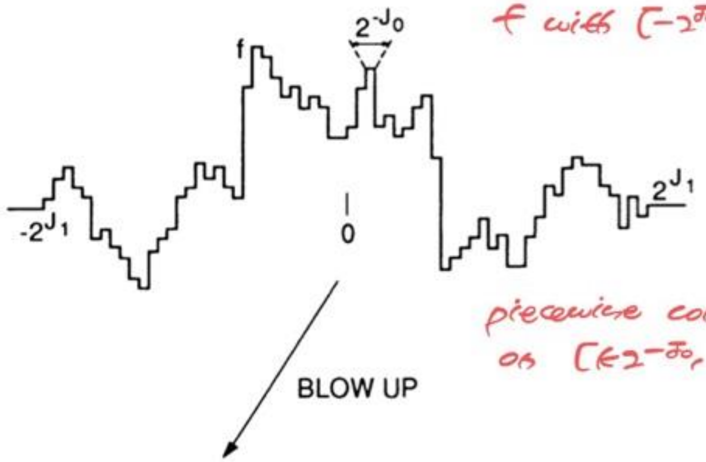
i.e $f^1|_{[k 2^{-J_0+1}, (k+1)2^{-J_0+1}]} = \text{constant} = f^1_k = \frac{1}{2}(f^0_{2k} + f^0_{2k+1})$

average

The function δ^1 is piecewise constant with same stepwidth as f^0

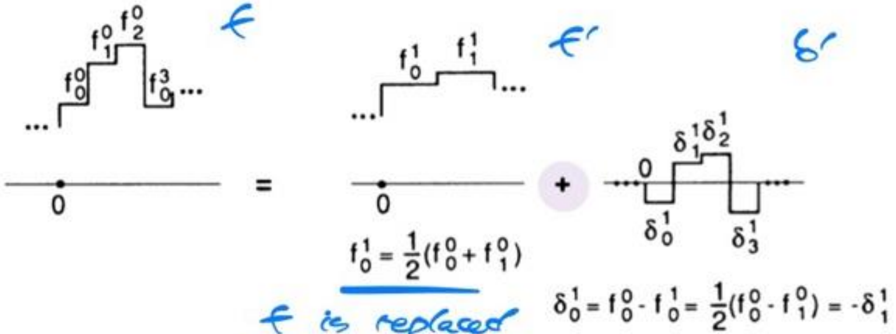
$\delta^1_{2\ell} = f^0_{2\ell} - f^1_\ell = \frac{1}{2}(f^0_{2\ell} - f^0_{2\ell+1})$

3. Different types of wavelet transform



f with $[-2^{J_1}, 2^{J_1}]$

piecewise constant on $[\ell 2^{-J_0}, (\ell+1)2^{-J_0}]$



f $\quad f^1 \quad \delta^1$

f is replaced by its average

$$\delta^0_{2\ell+1} = f^0_{2\ell+1} - f^1_\ell$$

$$= \frac{1}{2}(f^0_{2\ell+1} - f^0_{2\ell}) = -\delta^1_{2\ell}$$

$$\Rightarrow \delta^1 = \sum_{\ell=-2^{J_1+J_0-1}+1}^{2^{J_1+J_0-1}} \delta^1_{2\ell}\,\varphi(2^{J_0-1}x - \ell)$$

linear combination

therefore $f = f^0 = f^1 +$

$$\Rightarrow f^1 = f^2 + \sum_\ell c_{-J_0+2,\ell}\,\varphi_{-J_0+2,\ell}$$

with f^2 on $[-2^{J_1}, 2^{J_1}]$

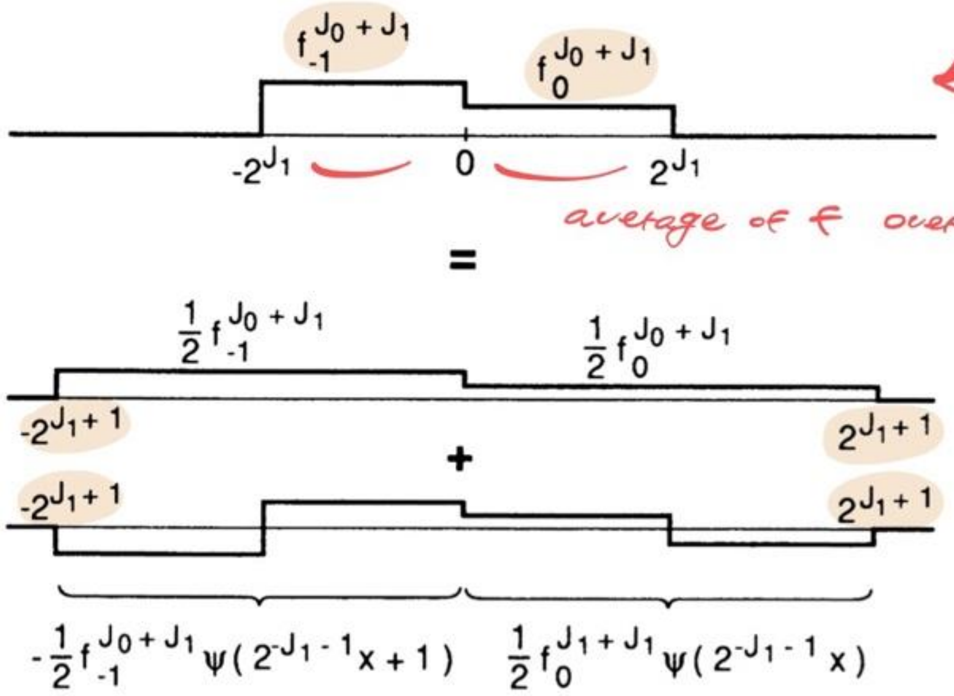
⋮ keep going

$$f = f^{J_0+J_1} + \sum_{m=-J_0+1}^{J_1} \sum c_{m,\ell}\,\varphi_{m,\ell}$$

3. Different types of wavelet transform

$$f = f^{J_0+J_1} + \sum_{m=-J_0+1}^{J_1} \sum c_{m,l}\,\psi_{m,l}$$

two constant pieces



average of f over $[0, 2^{J_1}]$

$$f^{J_0+J_1}\big|_{[0,2^{J_1}]} \equiv f_0^{J_0+J_1}$$

$$f^{J_0+J_1}\big|_{[-2^{J_1},0]} \equiv f_{-1}^{J_0+J_1}$$

∘ $f^{J_1+J_2} = f^{J_1+J_2+1} + \delta^{J_1+J_2+1}$

where $f^{J_1+J_2+1}\big|_{[0,2^{J_1+1}]} \equiv \frac{1}{2} f_0^{J_1+J_2}$

$f^{J_1+J_2+1}\big|_{[-2^{J_1+1},0]} \equiv \frac{1}{2} f_{-1}^{J_1+J_2}$

∘ $\delta^{J_1+J_2} = \frac{1}{2} f_0^{J_1+J_2}\,\psi(2^{-J_1-1}x) - \frac{1}{2} f_{-1}^{J_1+J_2}\,\psi(2^{-J_1-1}x+1)$

3. Different types of wavelet transform

Repeat $\rightarrow$ $$f = f^{J_0+J_1+k} + \sum_{m=-J_0+1}^{J_1+k} \sum_{l} c_{m,l}\, \psi_{m,l}$$

where $\text{support}\,(f^{J_0+J_1+k}) = [-2^{J_1+k}, 2^{J_1+k}]$.

$$f^{J_0+J_1+k}\big|_{[0,\,2^{J_1+k}]} = 2^{-k} f_0^{J_0+J_1}.$$

$$f^{J_0+J_1+k}\big|_{[-2^{J_1+k},\,0]} = 2^{-k} f_{-1}^{J_0+J_1}$$

$$\Rightarrow \left\| f - \sum_{m=-J_0+1}^{J_1+k} \sum_{l} c_{m,l}\, \psi_{m,l} \right\|_{L^2}^2 = \left\| f_0^{J_0+J_1+k} \right\|_{L^2}^2$$

$$= 2^{-k/2} \cdot 2^{J_1/2} \left[\, |f_0^{J_0+J_1}|^2 + |f_{-1}^{J_0+J_1}|^2 \,\right]^{\frac{1}{2}}$$

which can be made arbitrarily small by taking sufficiently large k

$\longrightarrow$ multiresolution : successive coarser approximation to f

↪ averaging f over larger intervals

↪ differences between the approximation with resolution 2^{j-1} the next coarser level, with 2^j, as a linear combination of $\psi_{j,k}$

3. Different types of wavelet transform

- A ladder of spaces $(V_j)_{j\in\mathbb{Z}}$ representing successive resolution levels.

$V_j = \{ f \in L^2(\mathbb{R}) \mid f : \text{piecewise constant on } [2^j k, 2^j (k+1)],\ k \in \mathbb{Z} \}$

these spaces have the following Properties :

① $\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots$

② $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$, $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$

③ $f \in V_j \leftrightarrow f(2^j \cdot) \in V_0$

"all spaces are scaled version of one space"

④ $f \in V_0 \leftrightarrow f(\cdot - n) \in V_0$ for $\forall n \in \mathbb{Z}$

⑤ $\exists \phi \in V_0$ so that the $\phi_{0,n}(x) = \phi(x-n)$ constitute an orthonormal basis for V_0

↪ ex $\exists \psi$ so that $\mathrm{Proj}_{V_{j-1}} f = \mathrm{Proj}_{V_j} f + \sum_{k\in\mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}$

($\psi_{j,k}$ constitute automatically an orthonormal basis)

3. Different types of wavelet transform

⑦ construction of ψ

Since $\phi \in V_0 \subset V_{-1}$ and $\phi_{-1,n}(x) = \sqrt{2}\,\phi(2x-n)$ constitute an orthonormal basis for V_{-1} by ③, ⑤.

$\exists\, \alpha_n = \sqrt{2}\langle \phi, \phi_{-1,n}\rangle$ so that $\phi(x) = \sum_n \alpha_n \phi(2x-n)$.

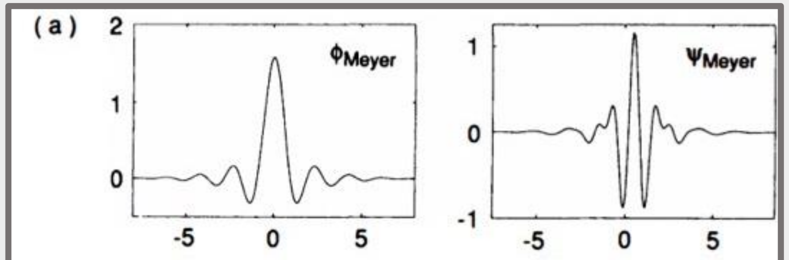
scaling function

It then suffices to take

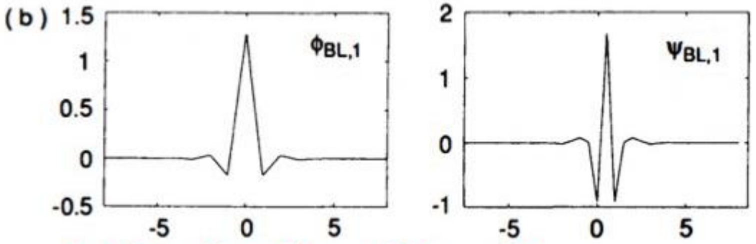
$$\psi(x) = \sum_n (-1)^n \alpha_{-n+1}\, \phi(2x-n)$$

3. Different types of wavelet transform

• pairs of ϕ, ψ corresponding to different multiresolution analysis

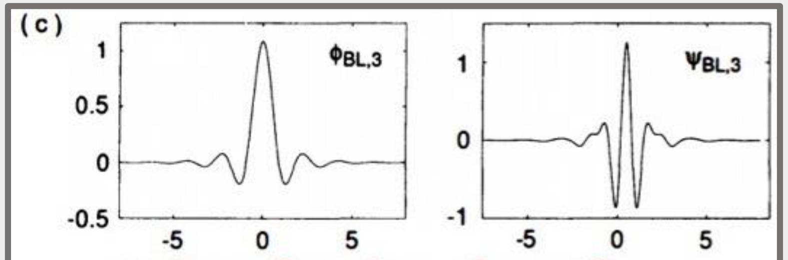


Both are infinitely supported



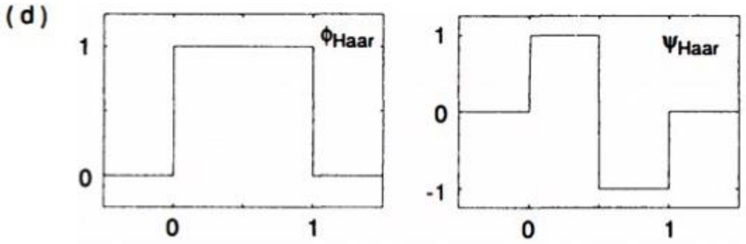
spline function (linear)

Both are infinitely supported
Exponential decay

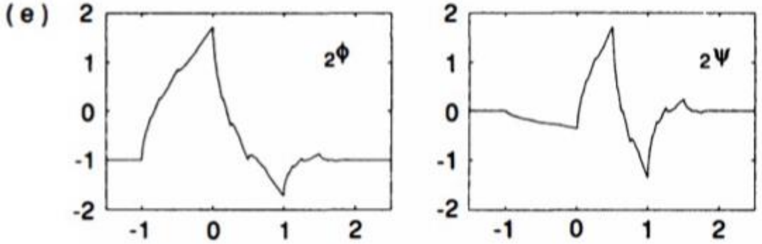


spline function (cubic)

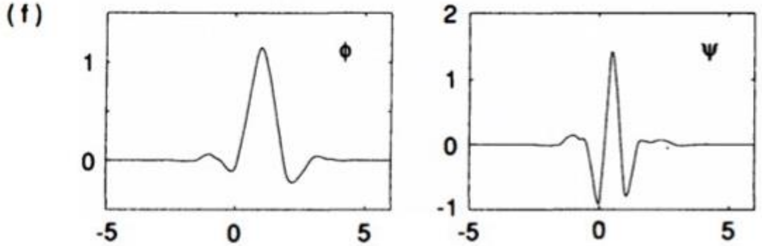
Both are infinitely supported
Exponential decay



the first of a family of compactly supported wavelets



next member of the family of compactly supported wavelets



Another compactly supported wavelets and less asymmetry