Lecture 6. Wavelets and Multiresolution

Juno Kim

Department of Mathematics & Statistics
Seoul National University

Manifold Learning, Spring 2022

Table of Contents

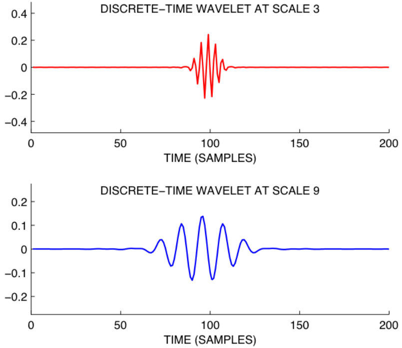


Figure: Wavelets at different scales.

- Wavelet theory is an analogue of Fourier theory using wavelet functions which are localized in time/space, generated from a single wavelet ψ by translations and dilations.

- Fix a *mother* (prototype) wavelet $\psi \in L^2(\mathbb{R})$
- For any signal $f \in L^2(\mathbb{R})$, define the wavelet coefficients

$$W[f](s,a) := \langle f, \psi_{s,a} \rangle, \quad \psi_{s,a}(t) := \frac{1}{s}\psi\left(\frac{t-a}{s}\right)$$

- If ψ satisfies the admissibility condition

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

  we can retrieve f from its transform via the inversion law:

$$f(t) = \frac{1}{c_\psi} \int_0^\infty \int_{-\infty}^\infty W[f](s,a)\psi_{s,a}(t) \frac{da\, ds}{s}$$

- In this case $\hat{\psi}(0) = \int \psi = 0$ and $\int f = 0$

- Assume ψ is real & even
- Define the *continuous* wavelet transform (CWT) operator:

$$T_s f(a) := W[f](s,a)$$

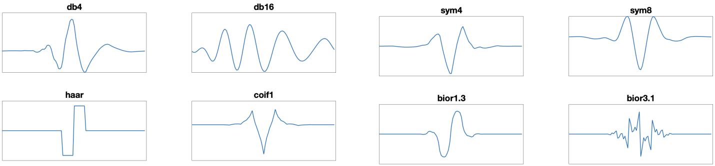
- Then $T_s f = \psi_{s,0} * f$, $\psi_{s,a} = T_s \delta_a$ and

$$T_s f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \hat{\psi}(s\omega) \hat{f}(\omega) d\omega$$

- T_s acts on the Fourier transform by multiplying the scaled band-pass filter $\hat{\psi}(s\omega)$. In this expression, the scale factor only appears in the frequency domain.
- The *discrete* WT is given by convolution with the family $\psi_{j,k}(t) = \alpha^{-j/2}\psi\left(\alpha^{-j}t - \beta k\right)$, $j,k \in \mathbb{Z}$

Exercise. Show the formal inversion law and verify the above properties.

Examples



- Haar wavelet: $\psi(t) = 1_{[0,1/2)}(t) - 1_{[1/2,1)}(t)$
- sinc wavelet: $\psi(t) = (\sin(2\pi t) - \sin(\pi t))/\pi t$
- 2D Mexican hat: $\psi(t) = \dfrac{1}{\pi\sigma^4}\left(1 - \dfrac{x^2+y^2}{2\sigma^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$
- Daubechies family: continuous, orthogonal, compact support

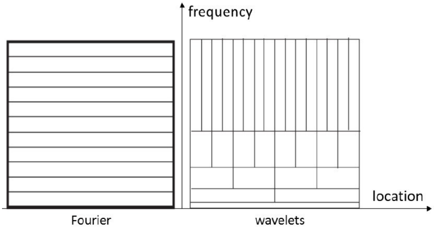


Figure: Comparison of Fourier and wavelet bases.

- While Fourier modes are globally distributed, wavelets are localized (translation a) in time/space (decay of ψ) and frequency (decay of $\hat{\psi}$).
- Wavelets allow more efficient representations of signals whose primary information lies in localized singularities, such as edges in images or step discontinuities in time series.

Table of Contents

FFT

Recall the 2-point Fast Fourier Transform. The discrete transform of size N input x_n is computed in $O(N \log N)$ time as:

$$
\begin{aligned}
DFT[N; x_n](k) &:= \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} nk} \\
&= \sum_{n=0}^{N/2-1} x_{2n} e^{-\frac{2\pi i}{N/2} nk} + e^{-\frac{2\pi i}{N} k} \sum_{n=0}^{N/2-1} x_{2n+1} e^{-\frac{2\pi i}{N/2} nk} \\
&= DFT[N/2; x_{2n}]_k + e^{-\frac{2\pi i}{N} k} DFT[N/2; x_{2n+1}]_k \quad k < N/2 \\
&= (\qquad " \qquad) - e^{-\frac{2\pi i}{N} k} (\qquad " \qquad) \quad k \geq N/2
\end{aligned}
$$

The DFT is recursively derived from the even/odd DFTs of half length – the information at half resolution (sampling) is reused.

MRA

- The analogous Fast Wavelet Transform is intimately related to multiresolution analysis – a decomposition of the signal space into wavelet spaces of discretized scale
- An MRA of $L^2(\mathbb{R})$ is an increasing sequence of closed linear subspaces V_j ($j \in \mathbb{Z}$) such that:

1. $v(x) \in V_j \Leftrightarrow v(2x) \in V_{j+1}$
2. $v(x) \in V_j \Leftrightarrow v(x+1) \in V_j$
3. $\overline{\cup V_j} = L^2(\mathbb{R})$, $\cap V_j = 0$
4. $\exists \phi \in V_0$ (*father wavelet*) s.t. $\{\phi(x-k) : k \in \mathbb{Z}\}$ is a frame for V_0 and $\hat{\phi}(0) = 1$

A *frame/Riesz basis* is a generalization of orthonormal basis: a spanning set e_k of V satisfying $m\|x\|^2 \leq \sum |\langle x, e_k \rangle|^2 \leq M\|x\|^2 \ \forall x \in V$.

- The projection operator $\Pi[V_j]$ provides approximations at resolution 2^j so that $\lim_{j\to\infty} \Pi[V_j]f = f$
- Let:

$$V_{j+1} = V_j \bigoplus^{\perp} W_j$$

  The detail space W_j contains the information required to move to higher resolutions, $\bigoplus_j^{\perp} W_j = L^2(\mathbb{R})$
- ψ is a mother wavelet corresponding to ϕ if $\{\psi(x-k) : k \in \mathbb{Z}\}$ is a frame for W_0
- Wavelet frames: $V_j = \langle \phi_{j,k}(t) := 2^{j/2}\phi(2^j t - k) : k \in \mathbb{Z}\rangle$ and $W_j = \langle \psi_{j,k} : k \in \mathbb{Z}\rangle$
- The Haar wavelet has father $1_{[0,1]}(t)$

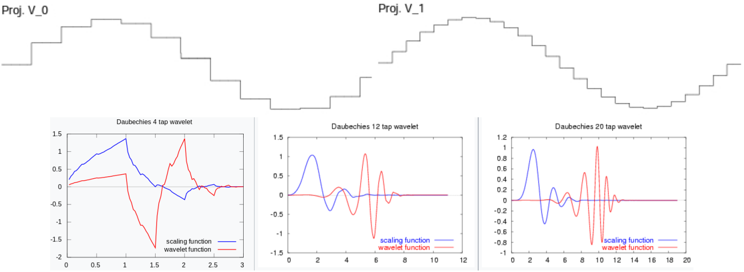


Figure: (1) The sine function at different resolutions of the Haar basis. (2) Mother-father pairs for the Daubechies family.

- Some wavelets do not admit a $\phi - \psi$ pair and has no MRA
- Orthonormality of $\phi_{j,k}$ and $\psi_{j,k}$ is generally too strict and *biorthogonal* wavelets in dual MRA sequences are used

Prop.

- refinement equation: $\phi(x) = \sum_k 2h_k\phi(2x-k)$ for a normalized 'scaling filter' h
- In the spectral domain, $\hat{\phi}(\omega) = DtFT[h](\omega/2) \cdot \hat{\phi}(\omega/2)$
- $V_j = \langle \phi_{j,k}(t) := 2^{j/2}\phi(2^j t - k) | k \in \mathbb{Z} \rangle$
- $\phi_{j,\ell} = \sum_k \sqrt{2} h_k \phi_{j+1,k+2\ell}$

Similarly, $\psi(x) = \sum_k 2g_k\phi(2x-k)$ and $\psi_{j,\ell} = \sum_k \sqrt{2} g_k \phi_{j+1,k+2\ell}$.

These structural equations form the basis of fast information extraction (halving resolution) and reconstruction (doubling resolution) without the need to evaluate integrals at every step.

- Denote the wavelet decomposition coefficients:

$$\Pi[V_{j+1}]f = \sum_k \lambda_{j+1,k}\phi_{j+1,k}$$
$$= \Pi[V_j]f + \Pi[W_j]f = \sum_k \lambda_{j,k}\phi_{j,k} + \sum_k \gamma_{j,k}\psi_{j,k}$$

- Using the structure equations, we derive:

Theorem (Fast Wavelet Transform)

The DWT $\gamma_{j,k}$ at resolution j can be retrieved from $j+1$ via the recursive relations

$$\lambda_{j,\ell} = \langle \Pi[V_j]f, \phi_{j,\ell}\rangle = \sum_k \sqrt{2}h_k\lambda_{j+1,k+2\ell} \quad \textit{and}$$
$$\gamma_{j,\ell} = \langle \Pi[W_j]f, \psi_{j,\ell}\rangle = \sum_k \sqrt{2}g_k\lambda_{j+1,k+2\ell}$$

Exercise. Verity the Proposition and FWT. Derive the Inverse FWT.

Table of Contents

- Let $\mathcal{M}^t$ be a compact oriented Riemannian manifold with Laplace-Beltrami operator $\Delta_{\mathcal{M}}$
- Suppose $K_s(\cdot,\cdot)$ is a smooth kernel in $C^\infty(\mathbb{R}^+ \times \mathcal{M} \times \mathcal{M})$ with associated $L^2(\mathcal{M})$ operator

$$T_s f(x) = \int_{\mathcal{M}} K_s(x,y) f(y) \, \mathrm{vol}_{\mathcal{M}}(y)$$

- K_s is a manifold wavelet if $T_s 1 = T_s^* 1 = 0$ and

$$\int_0^\infty \| T_s f \|^2 \frac{ds}{s} = c_K \, \| (I - \Pi_1) f \|^2 \quad \forall f \in L^2(\mathcal{M})$$

where Π_1 is the projection onto the constant functions and $c_K > 0$.

- Define the Schwartz space (algebra)

$$\mathcal{S}(\mathbb{R}^n) := \{g \in C^\infty(\mathbb{R}^n, \mathbb{C}) : \| |x|^\alpha D^\beta g \|_\infty < \infty \ \forall \alpha, \beta \in \mathbb{N}^n\}$$

- $\mathcal{S}(\mathbb{R}^n)$ consists of functions of rapid decay e.g. $|x|^\alpha e^{-c\|x\|^2}$
- $C_0^\infty \subset \mathcal{S} \overset{dense}{\subset} L^p \ \ (p < \infty)$

Lemma

Let T be a positive self-adjoint operator on Hilbert space H. For $g \in s \cdot \mathcal{S}(\mathbb{R}^+)$ and $c = \int_0^\infty |g(s)|^2 ds/s$,

$$\lim_{\epsilon \to 0^+,\, N \to \infty} \int_\epsilon^N |g(sT)|^2 \frac{ds}{s} = c(I_H - \Pi)$$

in the strong operator topology, where Π is projection onto $\ker T$.

Theorem

For $g \in s \cdot \mathcal{S}(\mathbb{R}^+)$, let K_s be the kernel (spectral wavelet) associated to $T_s = g(-s^2\Delta_{\mathcal{M}})$. Then K_s is a manifold wavelet.

Proof

- Denote the harmonics by (λ_k, u_k), $\lambda_0 = 0$, $u_0 = \mathrm{vol}(\mathcal{M})^{-1/2}$
- u_k are orthogonal w.r.t. $\langle f, g\rangle_{\mathcal{M}} = \int fg \,\mathrm{vol}_{\mathcal{M}}$
- Writing $K_{\sqrt{s}}(x, y) = \sum_{k,\ell} a_{k\ell} u_k(x) u_\ell(y)$,

$$g(-s\Delta_{\mathcal{M}})u_j = \int_{\mathcal{M}} K_{\sqrt{s}}(\cdot, y) u_j(y) \,\mathrm{vol}_{\mathcal{M}}(y) = \sum_{k,\ell} a_{k\ell} u_k \delta_{j\ell}$$

yields $a_{k\ell} = \delta_{k\ell} g(s\lambda_k)$

- Thus $K_{\sqrt{s}}(x,y) = \sum_k g(s\lambda_k) u_k(x) u_k(y)$
- $\int K_{\sqrt{s}}(x,y)\,\mathrm{vol}_{\mathcal{M}}(x) = \int K_{\sqrt{s}}(x,y)\,\mathrm{vol}_{\mathcal{M}}(y) = 0$ follows from $\langle u_k, 1\rangle = 0$ $(k \geq 1)$ and $g(0) = 0$
- The 2nd condition follows from

$$\int_0^\infty \|g(-s\Delta_{\mathcal{M}})f\|^2 \frac{ds}{s} = \left\langle \int_0^\infty |g|^2(-s\Delta_{\mathcal{M}}) f \frac{ds}{s}, f \right\rangle_{\mathcal{M}},$$

applying the Lemma and substituting s^2. $\blacksquare$

Remark

- In the language of distributions, $K_s(x,y) = T_s\delta_y(x) = g(-s^2\Delta_{\mathcal{M}})\delta_y(x)$
- Analogously on $\mathbb{R}$, the wavelet $\psi \stackrel{f}{=} g(-d^2/dx^2)\delta$ defined as the inverse CFT of $g(\omega^2)$ is admissible
- Spectral wavelets implement scaling in the spectral domain, same as $\widehat{T_s f}(\omega) = \hat{\psi}(s\omega)\hat{f}(\omega)$ on $\mathbb{R}$:

Proposition

Let the manifold FT of $f \in L^2(\mathbb{R})$ be $\hat{f}(k) = \langle f, u_k\rangle$. Then:

$$\widehat{T_s f}(k) = g(s^2\lambda_k)\hat{f}(k) \quad \forall s > 0,\ k = 0, 1, \cdots$$

Theorem (Inverse manifold WT)

Any $f \in (I - P_1)L^2(\mathcal{M})$ may be reconstructed from $T_s f$ via

$$\int_0^\infty T_s^* T_s f \frac{ds}{s} = c_K f$$

Proof

- Define $\mathcal{H} = (I - P_1)L^2(\mathcal{M})$ and $\mathcal{K} = L^2(\mathbb{R}^+, \mathcal{H}, dt/t)$
- Let the bounded operator $U : \mathcal{H} \to \mathcal{K}$ be $Uf := (T_s f)_{s>0}$
- Check $U^* : \mathcal{K} \to \mathcal{H}$ is $U^*(h_s)(x) = \int_0^\infty T_s^* h_s(x) dt/t$
- $\|Uf\|_{\mathcal{K}} = c_K \|f\|_{\mathcal{H}}^2$ and U is a scaled isometry
- Thus $\langle U^* Uf, h\rangle_{\mathcal{H}} = \langle Uf, Uh\rangle_{\mathcal{K}} = c_K \langle f, h\rangle_{\mathcal{H}}$ by polarization and $U^* U = c_K$. ■

Schwartz wavelets

g only needs moderate decay to yield the L^2 theory. However, imposing Schwartz-type conditions gives wavelets adapted to the study of other function spaces.

A manifold wavelet K_s on $\mathcal{M}^t$ is a Schwartz ($\mathcal{S}$-) wavelet if for any $X, Y \in \mathrm{PDO}(\mathcal{M}, \mathbb{R})$ with degree j, k resp. and $N \in \mathbb{N}_{\geq 0}$ there exists $C_{N,X,Y}$ such that:

$$s^{t+j+k} \left| \left(\frac{d^{\mathcal{M}}(x,y)}{s} \right)^N XY \left[K_s(x,y) \right] \right| \leq C_{N,X,Y}$$

for all $s > 0$, $x, y \in \mathcal{M}$.

Theorem

1. *The classical wavelet* $K_s(x,y) = \frac{1}{s^n}\psi(\frac{x-y}{s})$ *associated to* $g(-s^2\Delta)$ *on* $\mathbb{R}^n$ *where* $\hat{\psi}(\xi) = g(\|\xi\|^2)$ *is an* $\mathcal{S}$*-wavelet.*
2. *Spectral wavelets on compact manifolds are* $\mathcal{S}$*-wavelets.*

The following facilitates wavelet analysis of Hölder spaces.

Proposition

Let K_s *be a* $\mathcal{S}$*-wavelet on* $\mathcal{M}$ *and let* $f \in L^2(\mathcal{M})$. *Then* f *is* α*-Hölder continuous iff:*

$$\sup_{\mathcal{M}} |T_s f| \leq C s^\alpha \quad \forall s > 0$$

Exercise. Prove (1) of the Theorem.

Table of Contents

Diffusion Operators

- Let $(\mathcal{M}, d^{\mathcal{M}})$ be a complete Riemannian manifold
- Let L be a nonnegative essentially self-adjoint C^∞ (sub)elliptic order 2 PDO on $\mathcal{M}$ with spectral decomposition $\int \lambda dE_\lambda$
- Its diffusion semigroup $(P_s)_{s\geq 0}$ on $L^2(\mathcal{M})$ is:

$$P_s := e^{-sL} = \int_{-\infty}^{\infty} e^{-s\lambda} dE_\lambda$$

- Diffusion equation: $\frac{d}{ds} \circ P_s = LP_s$ (in the strong sense)
- Assume P_s is compact for $s > 0$, so $P = P_1$ has eigenfunctions $(\lambda_k, \xi_k)_{k\geq 0}$

Exercise. Show $\sigma(P) \subset [0,1]$ and $\sigma(P_s) = \{\lambda_k^s\}_{k\geq 0}$.

Lemma

- *(semigroup)* $P_0 = I$ *and* $P_s P_t = P_{s+t}$
- *(contraction)* $\|P_s\|_{L^2(\mathcal{M})} \leq 1$
- *(strong continuity) For* $f \in L^2(\mathcal{M})$, $s \mapsto P_s f$ *is continuous*
- *(self-adjointness)* $\int_{\mathcal{M}} (P_s f) g \operatorname{vol}_{\mathcal{M}} = \int_{\mathcal{M}} f (P_s g) \operatorname{vol}_{\mathcal{M}}$

Theorem

The diffusion semigroup $(P_s)_{s\geq 0}$ *has an associated symmetric heat kernel* $(H_s)_{s\geq 0}$ *which satisfies the Chapman-Kolmogorov equation:*

$$H_{t+s}(x,y) = \int_{\mathcal{M}} H_t(x,w) H_s(w,y) \operatorname{vol}_{\mathcal{M}}(w)$$

Exercise. Show the Lemma. Verify that $K_s(x,y) = \sum_k \lambda_k^s \xi_k(x) \xi_k(y)$.

Examples

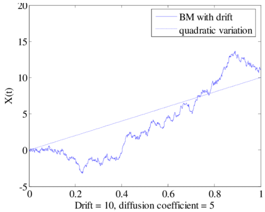


Figure: Brownian motion with positive drift and diffusion coefficient.

- Heat diffusion governed by the Laplacian
- Random walks induced by symmetric Markov chains
- p.d.f. of Brownian motion B_t, solutions of the SDE $dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t$
- Time evolution of the quantum wavefunction via the Schrödinger equation

Eigenmap

- The diffusion metric in Lec. 3 has the more general form

$$D_s(x,y)^2 = \sum_k \lambda_k^s(\xi_k(x)-\xi_k(y))^2 = \left\| P_{s/2}\delta_x - P_{s/2}\delta_y \right\|_{L^2(\mathcal{M})}^2$$

  i.e. it measures the L^2-embedded distance between diffused point sources.

- By defining a high-pass filter $H^\epsilon = \{k : \lambda_k \geq \epsilon\}$, the Eigenmap algorithm produces the approximate isometry

$$\Gamma_s^\epsilon : (\mathcal{M}, D_s) \to (\mathbb{R}^{|H^\epsilon|}, \|\cdot\|_{euc})$$
$$x \mapsto (\lambda_k^{s/2}\xi_k(x) : k \in H^\epsilon)$$

Diffusion MRA

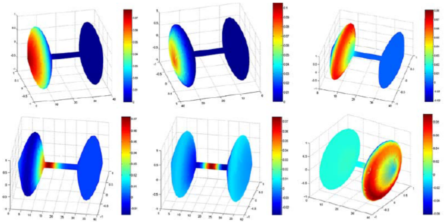


Figure: Localized wavelets on the dumbell manifold.

The classical MRA structure of wavelet pairs does not exist on manifolds. Coifman, Maggioni (2006) construct an MRA of $L^2(\mathcal{M})$ based on the time scaling inherent in diffusion processes.

Idea

- In MRA, $V_j = \langle \phi_{j,k} \rangle_{k \in \mathbb{Z}}$ was constructed from the base frame $V_0 = \langle \phi(t-k) \rangle_{k \in \mathbb{Z}}$ by applying the scaling semigroup R_s : $\nu(x) \mapsto s^{-1/2} \cdot \nu(s^{-1}x)$ at discrete $s_j = 2^{-j}$
- Similarly, we start with a frame Φ of local bump functions on 'dyadic cubes' at the finest level V_0
- The diffused family $P_{s_j}\Phi$ should form an approximate basis for a coarser (downsampled) space V_j
- Construct V_j by discarding eigenfunctions $\left\| P_{s_j} \xi_k \right\| = \lambda_k^{s_j} < \epsilon$ and taking only harmonics with high frequency content

Components

- precision $0 < \epsilon < 1$, discretized scales $(s_j)_{j \geq 0} \nearrow \infty$
- bump function family Φ constructed on dyadic cubes
- compact diffusion semigroup $(P_s)_{s \geq 0}$, eigenstates $(\lambda_k, \xi_k)_{k \geq 0}$
- high-pass band filters $H_j^\epsilon = \{k : \lambda_k^{s_j} \geq \epsilon\}$

Algorithm

- DMRA: $L^2(\mathcal{M}) = V_{-1} \supset V_0^\epsilon \supset V_1^\epsilon \supset \cdots$
- approximation spaces $V_j^\epsilon = \langle \xi_k : k \in H_j^\epsilon \rangle$
- detail spaces $V_j^\epsilon = V_{j+1}^\epsilon \bigoplus^\perp W_j^\epsilon$
- V_j is a compressed ϵ-approximation of $P_{s_j}\Phi$ or $\operatorname{im} P_{s_j}$
- apply ϵ-variants of Gram-Schmidt Orthogonalization to $P_{s_j}\Phi$ and $(P_{s_{j+1}} - P_{s_j})\Phi$ to obtain localized wavelets for V_j^ϵ, W_j^ϵ