Wavelets and Applications

Kévin Polisano
kevin.polisano@univ-grenoble-alpes.fr

M2 MSIAM & Ensimag 3A MMIS

January 28, 2022
Linear time-invariant filtering in classical signal processing

- **Time-invariant operator** \( L \). If the input \( f(t) \) is delayed/shifted by \( \tau \), \( f_\tau(t) = f(t - \tau) \), then the output is also delayed/shifted by \( \tau \):

\[
g(t) = Lf(t) \Rightarrow g(t - \tau) = Lf_\tau(t)
\]

- **Impulse response** \( h \) of \( L \):

\[
h(t) = L\delta(t) \Rightarrow h(t - \tau) = L\delta_\tau
\]

**Proposition**

A time-invariant linear filtering \( L \) is equivalent to a convolution with the impulse response \( h \).

**Proof.** Assume that \( f \) is continuous so that \( f(t) = \int_{-\infty}^{\infty} f(\tau)\delta_\tau(t) \, d\tau \) and \( L \) is linear and (weak) continuous hence

\[
Lf(t) = \int_{-\infty}^{\infty} f(\tau)L\delta_\tau(t) \, d\tau
\]

\[
= \int_{-\infty}^{\infty} f(\tau)h(t - \tau) \, d\tau = (f * h)(t) \quad \square
\]
Linear time-invariant filtering in classical signal processing

- With \( f[k] = f(kT) \), the sampled signal is

\[
f_s(t) = \sum_{k=-\infty}^{\infty} f[k] \delta(t - kT)
\]

- Let \( g(t) = f(t - kT) \) then

\[
g[n] = g(nT) = f(nT - kT) = f((n - k)T) = f[n - k] = (T_k f)[n]
\]

**Proposition**

A time-invariant linear filtering \( L \) is equivalent to a convolution with the impulse response \( h \).

**Proof.**

\[
(Lf_s)(nT) = \sum_{k=-\infty}^{\infty} f[k] L(\delta[n - k]) = \sum_{k=-\infty}^{\infty} f[k] h[n - k] = (f \ast h)[n] \quad \square
\]
Diagonalization of time-invariant operators

Proposition

Complex exponentials are eigenvectors of convolution operators

Proof. Consider \( f(t) = e^{i2\pi \xi t} \) and the sampled signal \( f_s(t) \)

- **Continuous case:**
  \[
  L e^{i2\pi \xi t} = \int_{-\infty}^{\infty} h(\tau) e^{i2\pi \xi (t-\tau)} \, d\tau = e^{i2\pi \xi t} \int_{-\infty}^{\infty} h(\tau) e^{-i2\pi \xi \tau} \, d\tau = \hat{h}(\xi) e^{i2\pi \xi t}
  \]

- **Discrete case:**
  \[
  L e^{i2\pi \xi nT} = \sum_{k=-\infty}^{\infty} h[k] e^{i2\pi \xi (n-k)T} = e^{i2\pi \xi nT} \sum_{k=-\infty}^{\infty} h[k] e^{-i2\pi \xi kT} = H(\xi) e^{i2\pi \xi nT}
  \]

NB. The Fourier transform of \( \delta(t - kT) \) is \( e^{-i2\pi \xi kT} \), which lead to the Discrete Time Fourier Transform (DTFT):

\[
H(\xi) = \sum_{k=-\infty}^{\infty} h[k] e^{-i2\pi \xi kT} = \mathcal{F} \left( \sum_{k=-\infty}^{\infty} h[k] \delta(t - kT) \right) = \hat{f}_s(\xi)
\]
The $\mathcal{Z}$-transform

More generally notice that for a complex $z$:

$$L \sum_{k=-\infty}^{\infty} h[k]z^{n-k} = z^n \sum_{k=-\infty}^{\infty} h[k]z^{-k} = H(z)z^n$$

which involves the linear $\mathcal{Z}$-transform of $h$ defined as:

$$\mathcal{Z}([h_k]_k) : z \in \mathbb{C} \mapsto H(z) = \sum_{k=-\infty}^{\infty} h[k]z^{-k}$$

- $H(z)$ is the transfer function of $L$
- DTFT corresponds to the $z$-transform evaluated in $z = e^{i2\pi \xi T}$

Properties

- **Translation**: $\mathcal{Z}(T_l h)(z) = \mathcal{Z}([h_{k-l}]_k)(z) = z^{-l}H(z)$
- **Scaling**: $\mathcal{Z}(D_a h)(z) = \mathcal{Z}(a^k [h_k]_k)(z) = H\left(\frac{z}{a}\right)$
- **Convolution**: $\mathcal{Z}(h_1 \ast h_2)(z) = \mathcal{Z}(\{\sum_i h_{1,i}h_{2,k-l}\}_k)(z) = H_1(z)H_2(z)$
FIR filters

Definition

Let a filter with an impulse response $h$. The filter is said to be with a Finite Impulse Response (FIR) if $h$ is finite that is $h = \{h_n\}_{n=0}^N$ and

$$H(z) = \sum_{n=0}^{N} h_n z^{-n} \quad \text{(polynomial in } z^{-1})$$

The FIR filter difference equation for a discrete time signal $f$, output $g$ and filter coefficients $h$ at sample $k$ is:

$$g[k] = (f \ast h)[k] = h[0]f[k] + h[1]f[k - 1] + \cdots + h[N]f[k - N]$$
The role of shift operator in classical signal processing

Let consider the (formal) polynomial representation of the signal $f$ via its $\mathcal{Z}$-transform:

$$F(z) = \sum_{n=0}^{N-1} f[n]z^{-n}$$

as well for the output $g = Lf$ represented by $G(z)$. Then we have:

$$G(z) = F(z)H(z) \iff H(z) = \frac{G(z)}{F(z)} \quad \text{(transfer function)}$$

**Shift operator**

Let consider a periodic extension of $f$ where the real line is folded around the circle $f_n = f_{n \mod N}$ and let define the *shift operator* which perform a simple delay

$$f = [f_0, f_1, \ldots, f_{N-1}] \mapsto g = \text{shift } f = [f_{N-1}, f_0, \ldots, f_{N-1}]$$

It is clear that $H_{\text{shift}}(z) = z^{-1}$. Observe the *shift invariance* with any other operator $L$ due to the commutativity $z^{-1} \cdot H_L(z) = H_L(z) \cdot z^{-1}$. 
Shift operator on graphs

Analogy between 1-D periodic signal and the ring graph

The (directed) **ring graph**, associated to a periodic time-serie
\[ f = [f_0, f_2, \ldots, f_{n-1}]^\top \in \mathbb{R}^n \] with \( f[k + n] = f[k] \), has the following adjacency matrix:

\[
A = \begin{pmatrix}
0 & 0 & \cdots & 1 \\
1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & 0
\end{pmatrix}
\]

Notice that \( g = Af = [f_{n-1}, f_0, \ldots, f_{n-2}]^\top \) is the signal \( f \) shifted by one.

**Figure:** A (directed) ring graph
Filtering operator on graphs

- Using the matricial notation, a filter $h$ on a graph can be in general represented by the matrix $H$:

$$g = Hf$$

- Let $A$ be an arbitrary adjacency matrix, which play the role of the *shift operator* on neighbors. Following the analogy with classical signal processing, a filter represented by $H$ is said to be *shift-invariant* if it commutes with the shift, that is:

$$AH = HA$$

- If the characteristic and minimal polynomial of $A$ are equals then every filter commuting with $A$ is a polynomial in $A$ i.e

$$H = H(A) = \sum_{k=0}^{K} h_k A^k$$
Proposition

The eigenvectors of the shift operator $A$ are the eigenfunctions of the polynomial filter $H$.

Proof. Let consider $A = U\Lambda U^{-1}$ where $U = (u_1|\cdots|u_n)$ are the eigenvectors and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ is the matrix of eigenvalues of $A$. Then, it is straightforward to verify that:

$$H = H(A) = \sum_{m=0}^{M} h_m (U\Lambda U^{-1})^m = UH(\Lambda)U^{-1}$$

where $H(\Lambda) = \text{diag}(H(\lambda_1), \ldots, H(\lambda_n))$. Finally one has

$$Hu_k = UH(\Lambda)U^{-1}u_k = UH(\Lambda)e_k = H(\lambda_k)u_k$$

$\Rightarrow$ invariance of the eigenvectors of the shift operator $A$ with respect to graph filters.
Frequency analysis of graph signals
Analogy with the 1D Fourier transform

- In the 1-D continuous time setting one can decompose a signal on the Fourier basis of complex exponentials:

\[ f(t) = \int_{-\infty}^{+\infty} \hat{f}(\xi) e^{2\pi it \xi} d\xi \]

where the argument \(2\pi \xi\) determines the frequency of oscillation of such functions.

- Observe that this basis also corresponds to the eigenfunctions of the 1-D Laplace operator \(\Delta = -\frac{\partial^2}{\partial t^2}\):

\[-\frac{\partial^2}{\partial t^2} e^{2\pi it \xi} = (2\pi \xi)^2 e^{2\pi it \xi}\]

⇒ One can consider the Laplacian matrix \(L\) as the shift operator instead of the adjacency matrix \(A\).
Frequency analysis of graph signals
Analogy between 1-D periodic signal and the ring graph

The **undirected ring graph** has the following adjacency matrix:

$$\begin{align*}
A &= \begin{pmatrix}
0 & 1 & 0 & 0 & \cdots & 0 & 1 \\
1 & 0 & 1 & 0 & \cdots & 0 & 0 \\
0 & 1 & 0 & 1 & \cdots & 0 & 0 \\
& \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots \\
1 & 0 & 0 & \ddots & \ddots & \ddots & 1 & 0
\end{pmatrix} \\
\end{align*}$$

$$\begin{align*}
L &= D - A = 2I - A = \\
\begin{pmatrix}
2 & -1 & 0 & 0 & \cdots & 0 & -1 \\
-1 & 2 & -1 & 0 & \cdots & 0 & 0 \\
0 & -1 & 2 & -1 & \cdots & 0 & 0 \\
& \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ddots & \ddots & \ddots & 2 & -1 \\
-1 & 0 & 0 & \ddots & \ddots & \ddots & -1 & 2
\end{pmatrix}
\end{align*}$$
Circulant matrices and Fourier basis

- Any circulant matrix $C$ is **diagonalizable** in the Fourier basis $C = U \Lambda U^\top$ where

$$U = \frac{1}{\sqrt{n}} \begin{pmatrix}
1 & 1 & \ldots & 1 \\
1 & \omega & \ldots & \omega^{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \omega^{n-1} & \ldots & \omega^{(n-1)^2}
\end{pmatrix}$$

and $\omega = \exp\left(\frac{2\pi i}{n}\right)$ is a primitive $n$-th root of unity.

- The columns $u_k$ (**Fourier modes**) of the matrix $U$ are the eigenvectors of any circulant matrix.

- Multiplying a vector $f \in \mathbb{R}^n$ by $U$ performs a discrete Fourier transform (DFT)

$$\hat{f} = U^\top f = \begin{pmatrix}
\langle u_1, f \rangle \\
\vdots \\
\langle u_n, f \rangle
\end{pmatrix}$$
Graph Fourier basis

Equivalence with classical Fourier basis for the ring graph

- The **undirected** ring graph has this *circulant* Laplacian matrix:

\[
L = \begin{pmatrix}
2 & -1 & 0 & 0 & \cdots & 0 & -1 \\
-1 & 2 & -1 & 0 & \cdots & 0 & 0 \\
0 & -1 & 2 & -1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 2 & -1 \\
-1 & 0 & 0 & 0 & \cdots & -1 & 2 \\
\end{pmatrix}
\]

- The eigendecomposition of the graph Laplacian is

\[L = U\Lambda U^\top\]

where the eigenvectors \(u_k\) are the Fourier modes.

- The eigenvalues are given by

\[\lambda_k = 2 - 2\cos\frac{\pi k}{n}\]
Graph reference matrix

Definition
Suppose we have a graph reference matrix $R$ associated to the graph $G = (V, E, w)$ as:

$$\forall i \neq j, \quad R_{ij} \neq 0 \iff (i \to j) \in E$$

and suppose it is diagonalizable in $\mathbb{C}$ as:

$$R = U \Lambda U^{-1}$$

with $U = (u_1 | \cdots | u_n)$ and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$
Graph Fourier modes

Definition
The eigenvectors \( \{u_k\} \) of \( R \) are considered to be graph Fourier modes and \( \{\lambda_k\} \) their associated graph frequency if:

1. (consistency) \( R \) is circulant for \( G \) reduced to the ring graph
2. (variational interpretation) \( \text{Re}(\lambda_k) \) or \( |\lambda_k| \) is a measure of variation of \( u_k \)

Laplacian matrix
For undirected graphs \( R \leftarrow L = D - A \) satisfies the two properties.

Many others possible definitions!

- \( R \leftarrow L_n = I - D^{-1/2}AD^{-1/2} \) (normalized Laplacian)
- \( R \leftarrow L_d = I - \frac{A}{\lambda_{\text{max}}(A)} \) (deformed Laplacian)
- \( R \leftarrow L_{\text{rw}} = I - D^{-1}A \) (random walk Laplacian)
Graph Fourier transform

The case of the normalized Laplacian of undirected graphs

Let consider the graph reference matrix:

\[ R \leftarrow L_n = I - D^{-1/2}A D^{-1/2} = U \Lambda U^\top \]

- \( U = (u_1 | \cdots | u_n) \) is a graph Fourier basis, each \( u_k \) is a generalized (co)sine
- \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) the spectrum of \( L_n \), each \( \lambda_k \) is a generalized (squared) frequency and

\[ 0 = \lambda_1 \leq \lambda_2 \leq \lambda_n \leq 2 \]

- Variational interpretation:

\[ \lambda_k = u_k^\top L_n u_k = \frac{1}{2} \sum_{i \sim j} A_{ij} \left[ \frac{u_k(i)}{\sqrt{d_i}} - \frac{u_k(j)}{\sqrt{d_j}} \right]^2 \]

- The Graph Fourier transform of a graph signal \( f \in \mathbb{R}^n \) reads

\[ \hat{f} = U^\top f \]
The graph Fourier transform encodes the graph structure

Credits: D. Shuman
Graph filtering

Given a filter $H$ defined in the Fourier space, the signal $f$ filtered by $h$ is

$$g = UH(\Lambda)U^Tf$$
Fast Graph filtering

Problem: Computing $g = U H(\Lambda) U^\top f$ costs $O(n^3)$

Solution: To use a polynomial approximation of order $p$ of $h$:

$$\tilde{H}(\lambda) = \sum_{l=1}^{p} \alpha_l \lambda^l \approx H(\lambda)$$

Indeed, in this case one has:

$$g = U H(\Lambda) U^\top f \approx U \tilde{H}(\Lambda) U^\top f = U \sum_{l=1}^{p} \alpha_l \Lambda^l U^\top f = \sum_{l=1}^{p} \alpha_l L^l f$$

$\Rightarrow$ only involves matrix-vector multiplication of cost $O(p|E|)$

Credits: N. Tremblay
Generalized convolution on graphs

In the vertex domain, the \( n \)-th element of the output signal

\[
g = U H(\Lambda) U^T f
\]

is given by this kind of \textbf{generalized convolution}\(^1\) on the graph:

\[
g(i) = (f \ast h)(i) = \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} f(j) u_k(j) H(\lambda_k) u_k(i) = \sum_{j=0}^{n-1} f(j) h_i(j)
\]

where the \textit{transfert function} is defined by:

\[
H(\lambda_k) = h_0 + h_1 \lambda_k + \cdots + h_M \lambda_k^M
\]

and the \textit{graph impulse response} is:

\[
h_i(j) = \sum_{k=0}^{n-1} H(\lambda_k) u_k(i) u_k(j) = (h \ast \delta_i)(j) = \frac{1}{\sqrt{n}} (\mathcal{T}_i h)(j)
\]

\(^1\)By replacing the complex exponentials \( \psi_k(t) = e^{2\pi i k t} \) by Laplacian eigenvectors \( u_k(n) \) in the classical relationship \( g(t) = (f \ast h)(t) = \int_\mathbb{R} \hat{g}(k) \psi_k(t) \, dk = \int_\mathbb{R} \hat{f}(k) \hat{g}(k) \psi_k(t) \, dk \).
Tikhonov regularization

We observe a noisy graph signal $y = f_0 + \epsilon$ where $\epsilon$ is uncorrelated additive Gaussian noise, and we want to recover $f_0$ which is a smooth with respect to the underlying graph.

To enforce this a priori information we penalize the optimization problem with a regularization term of the form $f^T L f$ measuring the smoothness and a fixed $\gamma$:

$$J(f) = \frac{1}{2} \| f - y \|_2^2 + \gamma f^T L f$$

The optimal reconstruction is given by

$$f^* = H(L)y, \quad H(\lambda) = \frac{1}{1 + 2\gamma \lambda}$$
**Proof.** We want to minimize the objective function

\[ J(f) = \frac{1}{2} \| f - y \|^2 + \gamma f^\top L f \]

By differentiating

\[ \frac{\partial J}{\partial f} = f - y + 2\gamma L f = 0 \]

which results in

\[ f = (I + 2\gamma L)^{-1} y \]

In the spectral domain, from \( L = U\Lambda U^\top \) and by noting \( Y = U^\top y \) and \( F = U^\top f \) one finally has

\[ F = (I + 2\gamma \Lambda)^{-1} Y \]

hence the expression

\[ H(\lambda) = \frac{1}{1 + 2\gamma \lambda} \]
Windowed Graph Fourier Transform

- **Modulation operator** for a function $f \in L^2(\mathbb{R})$ is defined by
  \[(M_\xi f)(t) = e^{2\pi i \xi t} f(t)\]

- Let $g \in L^2(\mathbb{R})$ a window, the windowed Fourier atom is given by
  \[g_{u,\xi}(t) = (M_\xi T_ug)(t) = g(t - u)e^{2\pi i \xi t}\]

- By analogy one can define the generalized modulation operator by:
  \[(M_k f)(i) = \sqrt{n} f(i) u_k(i)\]

- Then a windowed graph Fourier atom by:
  \[g_{i,k}(j) = (M_k T_i g)(j) = N u_k(j) \sum_{\ell=0}^{n-1} G(\lambda_\ell) u_\ell(i) u_\ell(j)\]

- The windowed graph Fourier transform by:
  \[Sf(i, k) = \langle f, g_{i,k} \rangle\]
Wavelets on graph

- The 1D continuous wavelet transform of \( f \in L^2(\mathbb{R}) \) at scale \( a \) and position \( b \) is given by:
  \[
  Wf(a, b) = \frac{1}{a} \int_{\mathbb{R}} \psi^* \left( \frac{x - b}{a} \right) f(x) \, dx = \int_{\mathbb{R}} \hat{\psi}^* (a \xi) \hat{f} (\xi) e^{i2\pi \xi b} \, d\xi
  \]

- The wavelet at scale \( a \) centered around \( b \) reads:
  \[
  \psi_{a,b}(x) = \int_{\mathbb{R}} \hat{\psi}^* (a \xi) \delta_b(\xi) e^{i2\pi \xi x} \, d\xi
  \]

- By analogy one can define the wavelet transform of \( f \) at node \( i \) of the graph and scale \( a > 0 \) by:
  \[
  Wf(i, a) = \sum_{k=0}^{n-1} H(a \lambda_k) \hat{f}(\lambda_k) u_k(i)
  \]

- The wavelet on graph is defined as:
  \[
  \psi_{a,b} = U H(a \Lambda) U^\top \delta_a
  \]
Wavelets on graph

**A wavelet:**

**Translated:**

**Scaled:**
Some applications of wavelets on graph

- Tikhonov regularization for denoising: \( \text{argmin}_f \left\{ \| f - y \|_2^2 + \gamma f^T L f \right\} \)

- Wavelet denoising: \( \text{argmin}_a \left\{ \| y - W^* a \|_2^2 + \gamma \| a \|_1 \right\} \)

Figures courtesy of D. Shuman
Some applications of wavelets on graph

Figures from Hammond et al., *Wavelets on graphs via spectral graph theory*, 2011
Some applications of wavelets on graph

Figures from Hammond et al., *Wavelets on graphs via spectral graph theory*, 2011
Some applications of wavelets on graph

Figures from Hammond et al., *Wavelets on graphs via spectral graph theory*, 2011