BGSM/CRM
AL&DNN

Harmonic analysis. Wavelets. Graph spectral transforms

S. Xambó

UPC & IMTech

9/11/2021

Index

References

[1] (bronstein-bruna-cohen-velickovic-2021)
[2] (hammond-et2-2019)
[3] (hammond-et2-2011)
[4] (mallat-2009)
[5] (mohlenkamp-pereyra-2008)
[6] (katznelson-2004)

N

Background material

Basis of a Banach space
Basis of a Hilbert space
Riesz basis and dual basis
Frames and dual frames
Orthogonal decompositions and projections

Let $(\mathcal{B}, \| \, \|)$ be a Banach space. N

A *basis* for $\mathcal{B}$ is a sequence $u_1, \cdots, u_k, \cdots \in \mathcal{B}$ such that for each $x \in \mathcal{B}$ there is a *unique sequence* $\lambda_1, \cdots, \lambda_k, \cdots \in \mathbf{R}$ such that $x = \sum_{k\geqslant 1} \lambda_k u_k$, where the convergence of the series is in the sense of the norm: $x = \lim_{n\to\infty} \sum_{k=1}^{n} \lambda_k u_k$. Thus the vectors u_k are linearly independent and span a space (their finite linear combinations) that is dense in $\mathcal{B}$.

Example. Let $u_k \in \ell^p$ be $\{\delta_{k,j}\}_{j\geqslant 1}$ (has 1 in the position k and 0 otherwise). Then $\{u_k\}_{k\geqslant 1}$ is a basis of ℓ^p, for all $p \in [1, \infty)$. □

Given a basis, a subtle result is that the map $\lambda_k : \mathcal{B} \to \mathbf{R}$, $x \mapsto \lambda_k(x)$ is continuous for all k (see [7, Theorem 1.6]).

If for any $x \in \mathcal{B}$ the convergence of $\sum_{k\geqslant 1} \lambda_k u_k$ is unconditional, $\{u_k\}$ is said to be an *unconditional basis*.

- Let $\mathcal{H}$ be a Hilbert space. N

$\mathcal{H}$ is said to be *separable* if it contains a countable dense subset.

Examples. ℓ^2. We will also see that the (complex-valued) $L^2([0,1])$, $L^2(\mathbf{R})$ are separable.

Orthogonal and orthonormal sets. A set $\{\psi_n\}$ in $\mathcal{H}$ is an *orthogonal set* if $\langle\psi_n,\psi_m\rangle = 0$ whenever $n \neq m$. If in addition $\langle\psi_n,\psi_n\rangle = 1$ for all n, the set is said to be *orthonormal*.

Orthonormal basis (or *complete orthonormal systems*). An orthonormal set $\{\psi_n\}$ is a basis of $\mathcal{H}$ if and only if

$$h = \sum_{n\geqslant 1} \langle h,\psi_n\rangle\psi_n, \qquad [*]$$

where the convergence is with respect to the $\mathcal{H}$-norm. N

The point is that if $h = \sum_{n\geqslant 1}\lambda_n\psi_n$, then $\lambda_n = \langle f,\psi_n\rangle$. □

Lemma. A set $\{\psi_n\}_{n\geqslant 1}$ is an orthonormal basis if and only if $\|\psi_n\| = 1$ for all n and $\sum_{n\geqslant 1} |\langle h, \psi_n\rangle|^2 = \|h\|^2$ for all $h \in \mathcal{H}$.

If it is an orthonormal basis, then

$$\|h\|^2 = \|\sum_{n\geqslant 1} \langle h, \psi_n\rangle \psi_n\|^2 = \sum_{n\geqslant 1} |\langle h, \psi_n\rangle|^2.$$

Conversely, the condition for ψ_k states that

$1 = \|\psi_k\|^2 = \sum_{n\geqslant 1} |\langle \psi_k, \psi_n\rangle|^2 = 1 + \sum_{n\neq k} |\langle \psi_k, \psi_n\rangle|$, which implies that $\langle \psi_k, \psi_n\rangle = 0$ for all n, k.

On the other hand, the condition implies that sum $\sum_{n\geqslant 1} \langle h, \psi_n\rangle \psi_n$ converges (cf. [8, Th. 4.11]) and then $h - \sum_{n\geqslant 1} \langle h, \psi_n\rangle \psi_n$ must vanish because this difference is orthogonal to all ψ_n. □

Corollary. Orthonormal basis are unconditional basis.

Riesz basis: A basis ψ_n of $\mathcal{H}$ for which there are constants $0 < c \leqslant C < \infty$ such that

$$c\|h\|^2 \leqslant \sum_{n\geqslant 1} |\langle h, \psi_n\rangle|^2 \leqslant C\|h\|^2. \qquad [*]$$

- Riesz bases are *unconditional*.

Applying [*] to ψ_n, we get that $\|\psi_n\| \leqslant \sqrt{C}$:

$$\|\psi_n\|^4 = \langle \psi_n, \psi_n\rangle^2 = \sum_{k\geqslant 1} |\langle \psi_n, \psi_k\rangle|^2 \leqslant C\|\psi_n\|^2. \qquad \square$$

Let $\{\psi_n\}$ be a Riesz basis. A set $\{\psi_n^*\}$ is called a *dual Riesz basis* if $\langle \psi_n, \psi_m^*\rangle = \delta_{n,m}$ (*biorthogonality*) and

$$h = \sum_n \langle h, \psi_n^*\rangle \psi_n = \sum_n \langle h, \psi_n\rangle \psi_n^*. \qquad [**]$$

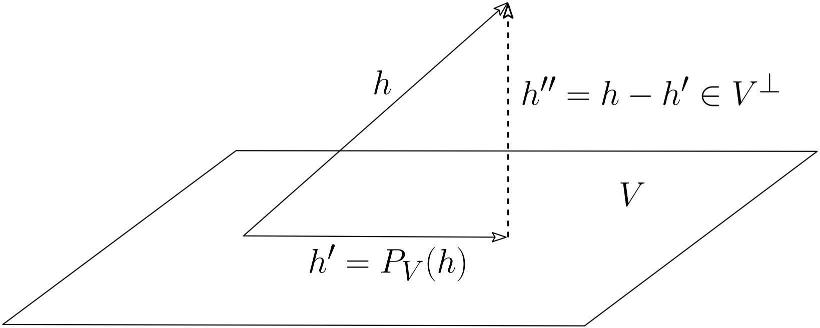
The pair $(\{\psi_n\}, \{\psi_n^*\})$ of dual Riesz basis is called a *biorthogonal basis*.

Remark. A system $\{\psi_n\}$ satisfying [$*$] in the previous page, not necessarily a basis of $\mathcal{H}$, but spanning a dense subspace, is called a *frame* of $\mathcal{H}$.

The bound $\|\psi_n\| \leqslant \sqrt{C}$ seen for a Riesz basis, and its proof, is valid for frames.

A frame for which $c = C$ is said to be *tight*.

- Given a frame $\{\psi_n\}$, there is a *dual frame* $\{\psi_n^*\}$ that satisfies the relations [$**$] in the previous page.



V is a closed subspace of $\mathcal{H}$. Then $\mathcal{H} = V \oplus V^\perp$. This defines a linear map $P_V : \mathcal{H} \to V$, $h \mapsto h'$, where $h = h' + h''$ is the unique decomposition of h with $h' \in V$ and $h'' \in V^\perp$.

This map is called the *orthogonal projection* of $\mathcal{H}$ on V.

Note that $h'' = P_{V^\perp}(h)$.

- If $\{\varphi_j\}_{j\in J}$ is an orthonormal basis of V, $P_V(h) = \sum_{j\in J}\langle h, \varphi_j\rangle\varphi_j$.

Notions of Fourier analysis

Real Fourier series

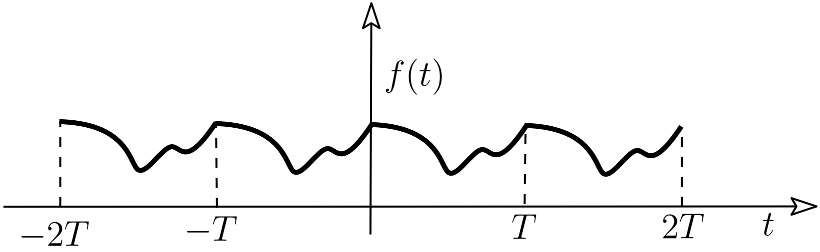
Trigonometric polynomials

Carleson's theorem

General intervals $[a, b]$

Time-frequency dictionary for Fourier series

The continuous Fourier transform



- $f(t)$ periodic function of *period* T.
- $\omega = 2\pi/T$ *angular frequency*.

Orthogonal relations

$$\int_0^T \cos n\omega t \cos n'\omega t = \begin{cases} 0 & \text{if } n \neq n' \\ T & \text{if } n = n' = 0 \\ T/2 & \text{if } n = n' > 0 \end{cases}$$

$$\int_0^T \sin n\omega t \sin n'\omega t = \begin{cases} 0 & \text{if } n \neq n' \\ T/2 & \text{if } n = n' > 0 \end{cases}$$

$$\int_0^T \sin n\omega t \cos n'\omega t = 0$$

Synthesis

$$f(t) = \frac{a_0}{2} + \sum_{n\geqslant 1}(a_n \cos n\omega t + b_n \sin n\omega t)$$

For bounded and piece-wise differentiable functions, the equality holds at points t where $f(t)$ is continuous. At jumps, the Fourier series is equal to $(f(t+) + f(t-))/2$.

Analysis

$$a_n = \frac{2}{T}\int_s^{s+T} f(t)\cos n\omega t dt \;(n \geqslant 0)$$

$$b_n = \frac{2}{T}\int_s^{s+T} f(t)\sin n\omega t dt \;(n \geqslant 1)$$

Even and odd functions

$$f(t) = f(-t) \Rightarrow a_n = \frac{4}{T}\int_0^{T/2} f(t)\cos n\omega t dt,\; b_n = 0.$$

$$f(t) = -f(-t) \Rightarrow a_n = 0,\; b_n = \frac{4}{T}\int_0^{T/2} f(t)\sin n\omega t dt.$$

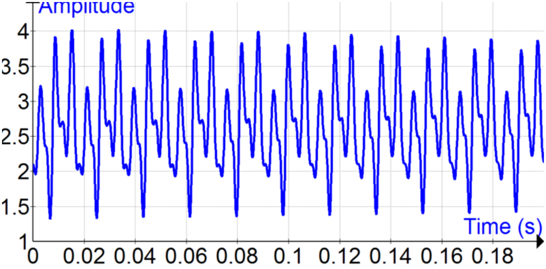
Case $T = 2\pi$

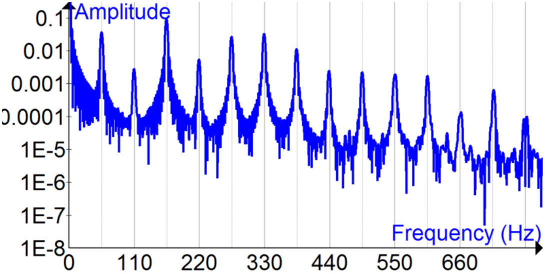
$$f(t) = \frac{a_0}{2} + \sum_{n\geqslant 1}(a_n \cos nt + b_n \sin nt)$$

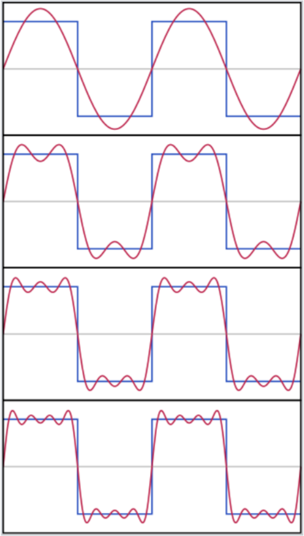
$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt,\; b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt$$

Amplitude-Phase form

$$f(t) = A_0 + \sum_{k\geqslant 1} A_n \cos(n\omega t + \alpha_n)$$


Amplitude
4
3.5
3
2.5
2
1.5
1
Time (s)
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18


Amplitude
0.1
0.01
0.001
0.0001
1E-5
1E-6
1E-7
1E-8
Frequency (Hz)
0
110
220
330
440
550
660



Going round the circle. The function $[0,1] \to \mathbf{C}$, $t \mapsto e^{2\pi it}$, goes once round the circle $S^1 \subset \mathbf{C}$.

So the function $(e^{2\pi it})^n = e^{2\pi int}$ goes n times round S^1.

Lemma $\int_0^1 e^{2\pi int} = 0$ if $n \neq 0$, and $= 1$ for $n = 0$. Consequently, for $n, n' \in \mathbf{Z}$, $\int_0^1 e^{2\pi int} e^{-2\pi in't} = 0$ if $n' \neq n$, and $= 1$ if $n' = n$.

Trigonometric polynomials (TPs). These are expressions of the form $p(t) = \sum_{n\in F} a_n e^{2\pi int}$, where $F \subset \mathbf{Z}$ is finite and $a_n \in \mathbf{C}$. It is a *superposition of pure harmonics* (*synthesis*).

Lemma. If $p(t)$ is a TP, then

$$a_n = \int_0^1 p(t)e^{-2\pi int}.$$

So $p(t)$ determines its coefficients (*analysis*). If we look at the a_n as a function $\hat{f} : \mathbf{Z} \in \mathbf{C}$, then we have

$$\hat{f}(n) = \int_0^1 f(t)e^{-2\pi int} \text{ and } f(t) = \sum_{n\in F} \hat{f}(n)e^{2\pi int}.$$

- What functions can be approximated by trigonometric polynomials?

The TPs form a subalgebra $\mathcal{P}$ of the algebra $\mathcal{C} = \mathcal{C}([0,1])$ of continuous functions $f : [0,1] \to \mathbf{C}$.

The algebra $\mathcal{P}$ is closed under complex cojugation and contains de constants. Under these conditions, the Stone-Weierstrass theorem (cf. [9, Th 8.1]) applies and hence $\mathcal{P}$ is dense in $\mathcal{C}$.

This means that for any $f \in \mathcal{C}$ and any $\epsilon > 0$, there is a $p \in \mathcal{P}$ such that $|f(t) - p(t)| < \epsilon$ for all $t \in [0,1]$.

This was anticipated by J. Fourier in his *Thérorie analytique de la chaleur* (1822), who claimed that *any* $f \in \mathcal{C}$ could be expressed as *trigonometric expansion*

$f(t) = \sum_{n \in \mathbf{Z}} a_n e^{2\pi int}$ (*synthesis*), with

$a_n = \hat{f}(n) = \int_0^1 f(t) e^{-2\pi int}$ (*analysis*).

N

It was not until 1966 that *L. Carleson* proved, in his paper *On convergence and growth of partial sums of Fourier series*, that (for a continuous function f) the Fourier partial sums converge pointwise almost everywhere to f.

■ That the convergence could not be 'everywhere' for all functions was known since the example, provided by du Bois-Raymond in 1873, of a continuous function whose Fourier series diverges in one point (cf. [10, Ch. 18]).

Remark. Carleson's stated and proved his theorem for functions that are in $L^2([0,1])$ (*square-integrable functions*).

Plancherel identity. $\|f\|_2^2 = \int_0^1 |f(t)|^2 dt = \sum_{n\in\mathbf{Z}} |\hat{f}(n)|^2$.

Consequently, by the lemma on page 7, the functions

$$\mathfrak{e}_n(t) = e^{2\pi int}$$

form an orthonormal basis of $L^2([0,1])$. Hence *this space is separable*.

- Let $[a, b]$, $a < b$, be an arbitrary interval, and set $L = b - a$. Then the following functions

$$\frac{1}{\sqrt{L}} e^{2\pi int/L}$$

form an orthonormal basis of $L^2([a, b]$.

In the case of the inteval $[-\pi, \pi]$, used by many authors, the basis is (redefining the symbol $\mathfrak{e}_n$)

$$\mathfrak{e}_n(t) = \frac{1}{2\pi} e^{int}.$$

Time/Space $[0,1]$	**Frequency** $\mathbb{Z}$
derivative $f'(t)$	polynomial $\widehat{f'}(n) = 2\pi in\,\hat{f}(n)$
circular convolution $(f*g)(t) = \int_0^1 f(t-s)g(s)ds$	product $\widehat{f*g}(n) = \hat{f}(n)\hat{g}(n)$
translation/shift $(\tau_s f)(t) = f(t-s)$	modulation $\widehat{\tau_s f}(n) = e^{-2\pi isn}\hat{f}(n)$

If $f \in L^2(\mathbf{R})$, its *Fourier transform* is the function $\hat{f}(\xi)$, $\xi \in \mathbf{R}$, defined by

$$\hat{f}(\xi) = \int_{\mathbf{R}} f(t)e^{-2\pi i\xi t}dt \text{ (\textit{analysis})}.$$

The *inverse Fourier transform* is the relation

$$f(t) = \int_{\mathbf{R}} \hat{f}(\xi)e^{2\pi i\xi t}d\xi \text{ (\textit{synthesis})}.$$

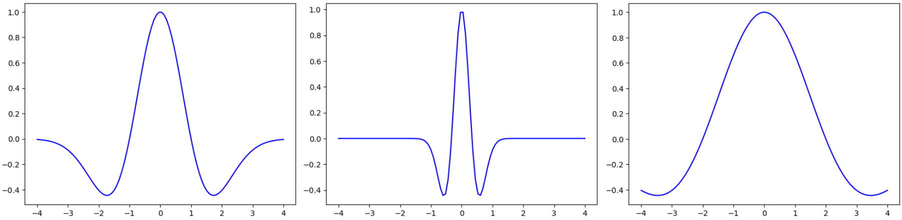
Plancherel's identity:

$$\|f\|_2^2 = \int_{\mathbf{R}} |f(t)|^2 dt = \int_{\mathbf{R}} |\hat{f}(\xi)|^2 d\xi = \|\hat{f}\|_2^2.$$

Remark. The trigonometric functions $\mathfrak{e}_\xi(t) = e^{2\pi i\xi t}$ do not belong to $L^2(\mathbf{R})$, but the inverse Fourier transform shows that they can be "superposed", with the coefficients $\hat{f}(\xi)$, to get $f(t)$.

Alternative formalism. $\hat{f}(\omega) = \frac{1}{2\pi}\int f(t)e^{-i\omega t}dt$.

Time/Space R	Frequency R (N)
derivative $d_t f(t) = f'(t)$	polynomial $\widehat{f'}(\xi) = 2\pi i\xi\, \hat{f}(\xi)$
convolution $(f * g)(t) = \int_{\mathbf{R}} f(t-s)g(s)ds$	product $\widehat{f * g}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$
translation/delay $(\tau_s f)(t) = f(t-s)$	modulation $\widehat{\tau_s f}(\xi) = e^{-2\pi i s\xi}\hat{f}(\xi)$
rescaling/dilation $f_s(t) = (1/s)\, f(t/s)$	rescaling $\widehat{f_s}(\xi) = \hat{f}(s\xi)$
conjugate flip $\tilde{f}(t) = \overline{f(-t)}$	conjugate $\widehat{\tilde{f}}(\xi) = \overline{\hat{f}(\xi)}$

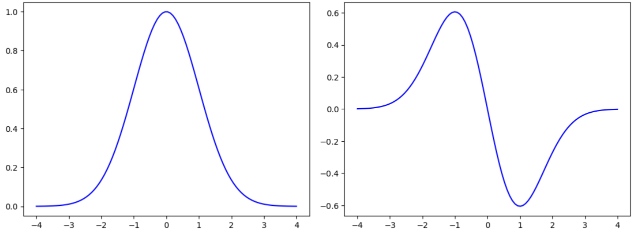


Left: Graph of the *Mexican hat*, $(1-t^2)e^{-t^2/2}$ (it is the negative of the second derivative of $e^{-t^2/2}$). Center and Right: same, but rescaled by 3 and $1/2$, respectively.

Gauss' function

$$e^{-t^2/2}$$

and its derivative.



It can be seen that if an operator $A : L^2(\mathbf{R}) \to L^2(\mathbf{R})$ commutes with translations ($A(\tau_s f) = \tau_s(Af)$), then there is a function $\hat{A}(\xi) \in L^2(\mathbf{R})$, called the *symbol* of A, such that

$$\widehat{Af}(\xi) = \hat{A}(\xi)\hat{f}(\xi).$$

For example, $\widehat{d_t}(\xi) = 2\pi i\xi$.

Smoothness and decay at infinity. Since $\widehat{f'}(\xi) = 2\pi i\xi\, \hat{f}(\xi)$, by Plancherel's formula we have, provided $f' \in L^2(\mathbf{R})$,

$$\int 4\pi^2\xi^2|\hat{f}(\xi)|^2 d\xi = \|f'\|_2^2 < \infty.$$

This shows, because of the factor ξ^2 in the integrand, that $|\hat{f}(\xi)|^2$ must decay fast enough to insure that the integral is finite.

In general, it can be seen that *the smoother the function $f(t)$, the faster has to be the decay of $|\hat{f}(\xi)|$* (see [4, Theorem 2.5]).

For each $n \in \mathbf{Z}$, the function $\mathfrak{e}_n(t) = e^{2\pi i n t}$ is a homomorphism $\mathfrak{e}_n : \mathbf{R} \to \mathrm{U}_1$ (the group $\mathrm{U}_1 = S^1$ is also denoted $\mathbb{T}$ by many authors writing on hamonic analysis). And the set $\{\mathfrak{e}_n\}_{n\in\mathbf{Z}}$ is a multiplicative group isomorphic to $\mathbf{Z}$ ($n \leftrightarrow \mathfrak{e}_n$).

Similarly, for each $\xi \in \mathbf{R}$, the function $\mathfrak{e}_\xi(t) = e^{2\pi i \xi t}$ is a homomorphism $\mathfrak{e}_\xi : \mathbf{R} \to \mathrm{U}_1$ And the set $\{\mathfrak{e}_\xi\}_{\xi\in\mathbf{R}}$ is a multiplicative group isomorphic to $\mathbf{R}$ ($\xi \leftrightarrow \mathfrak{e}_\xi$).

There is a similar formalism for $\mathbf{Z}^d$ and $\mathbf{R}^d$. In the latter case, for example, *analysis* and *synthesis* of a function $f \in L^2(\mathbf{R}^d)$ is given by:

$$\hat{f}(\boldsymbol{\xi}) = \int_{\mathbf{R}^d} f(\boldsymbol{t}) e^{-2\pi i \boldsymbol{\xi}\cdot\boldsymbol{t}} d\boldsymbol{t}$$

and (*inverse Fourier transform*)

$$f(\boldsymbol{t}) = \int_{\mathbf{R}^d} \hat{f}(\boldsymbol{\xi}) e^{2\pi i \boldsymbol{\xi}\boldsymbol{t}} d\boldsymbol{\xi}.$$

Plancherel: $\|f\|_2^2 = \int_{\mathbf{R}^d} |f(\boldsymbol{t})|^2 d\boldsymbol{t} = \int_{\mathbf{R}^d} |\hat{f}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} = \|\hat{f}\|_2^2$.

https://www.math.ucla.edu/ tao/preprints/fourier.pdf (Tao)

Wavelets

Definitions
Graphics
Calderrón's theorem
Discretization
Scalogram
Spectrogram

A *wavelet* is a function $\psi \in L^2(\mathbf{R})$ such that the functions

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k),\ j,k \in \mathbf{Z}$$

for an orthonomal basis of $L^2(\mathbf{R})$.

Orthogonal wavelet transform

$$Wf(j,k) = \langle f, \psi_{j,k}\rangle = \int_{\mathbf{R}} f(t)\psi_{j,k}(t)dt$$

Wavelet synthesis

For any $f \in L^2(\mathbf{R})$,

$$f = \sum_{j,k}\langle f, \psi_{j,k}\rangle\psi_{j,k}.$$

In the sequel we will further assume that $\int \psi dt = 0$ and $\|\psi\| = 1$.

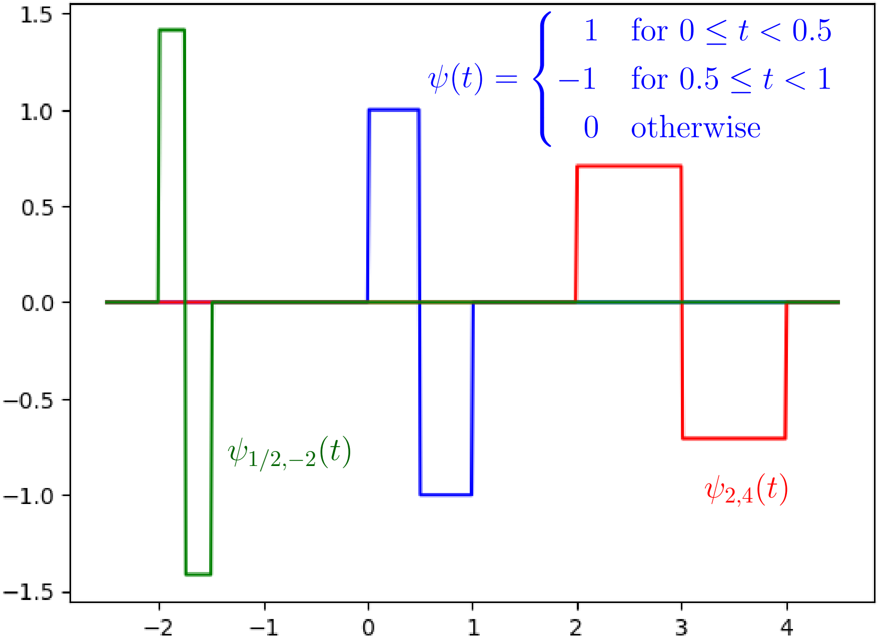
If ψ is a wavelet, for $u, s \in \mathbf{R}$, $s > 0$, define

$$\psi_{s,u}(t) = \frac{1}{\sqrt{s}}\psi(\frac{t-u}{s})$$

The factor $1/\sqrt{s}$ insures that $\|\psi_{s,u}\| = 1$.

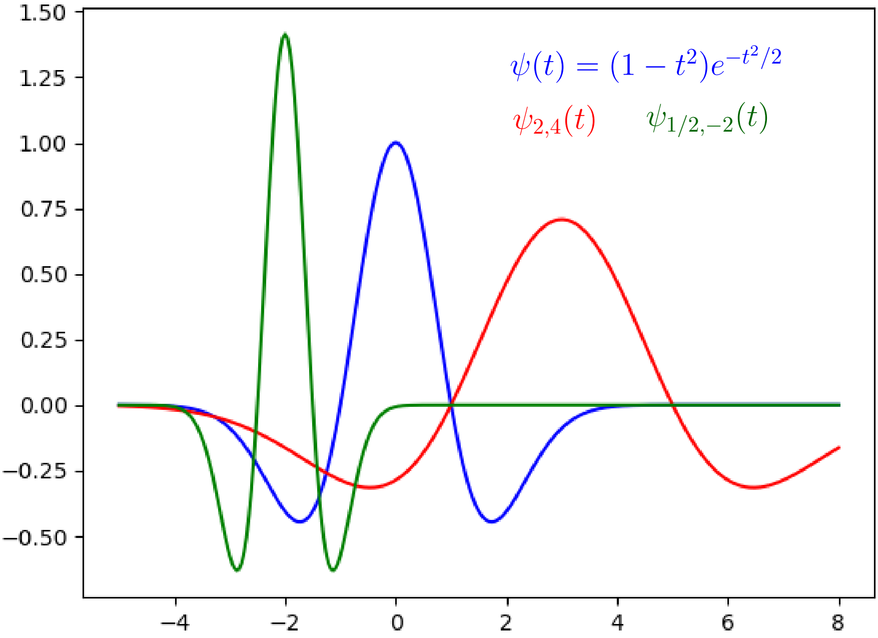
```
def atom(f,s,u):
    return lambda t: f((t-u)/s)/sqrt(s)

def haar(t):
    if 0 <= t < 0.5: return 1
    elif 0.5 <= t < 1: return -1
    else: return 0

def mex(t): return (1-t**2)*exp(-t**2/2)

def morlet(t): return exp(-t**2/2)*cos(5*t)
```

```
a = atom(haar,1,0)
b = atom(haar,2,2)
c = atom(haar,0.5,-2)

x = np.linspace(-2.5,4.5,500)

ya = [a(t) for t in x]
yb = [b(t) for t in x]
yc = [c(t) for t in x]

plt.plot(x,ya,'-',color='b')
plt.plot(x,yb,'-',color='r')
plt.plot(x,yc,'-',color='g')

plt.show()
```

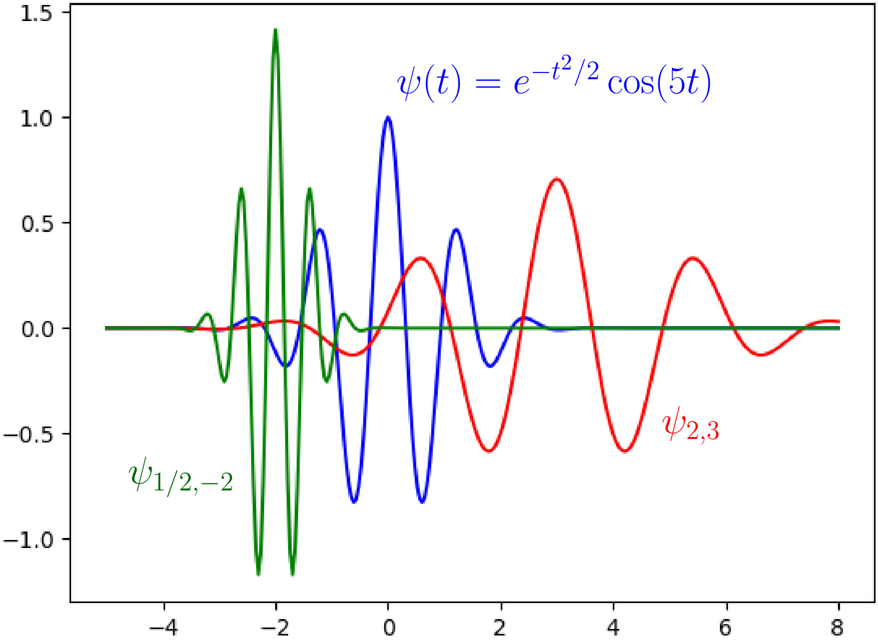


$\psi(t) = \begin{cases} 1 & \text{for } 0 \le t < 0.5 \\ -1 & \text{for } 0.5 \le t < 1 \\ 0 & \text{otherwise} \end{cases}$
$\psi_{1/2,-2}(t)$
$\psi_{2,4}(t)$
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2
−1
0
1
2
3
4

```
a = atom(mex,1,0)
b = atom(mex,2,3)
c = atom(mex,0.5,-2)

x = np.linspace(-5,8,300)

ya = [a(t) for t in x]
yb = [b(t) for t in x]
yc = [c(t) for t in x]


plt.plot(x,ya,'-',color='b')
plt.plot(x,yb,'-',color='r')
plt.plot(x,yc,'-',color='g')
plt.show()
```


$\psi(t) = (1 - t^2)e^{-t^2/2}$
$\psi_{2,4}(t)$
$\psi_{1/2,-2}(t)$
1.50
1.25
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−4
−2
0
2
4
6
8

```
a = atom(morlet,1,0)
b = atom(morlet,2,3)
c = atom(morlet,0.5,-2)

x = np.linspace(-5,8,300)
ya = [a(t) for t in x]
yb = [b(t) for t in x]
yc = [c(t) for t in x]


plt.plot(x,ya,'-',color='b')
plt.plot(x,yb,'-',color='r')
plt.plot(x,yc,'-',color='g')
plt.show()
```



$\psi(t) = e^{-t^2/2}\cos(5t)$
$\psi_{2,3}$
$\psi_{1/2,-2}$
1.5
1.0
0.5
0.0
−0.5
−1.0
−4
−2
0
2
4
6
8

Continuous wavelet transform

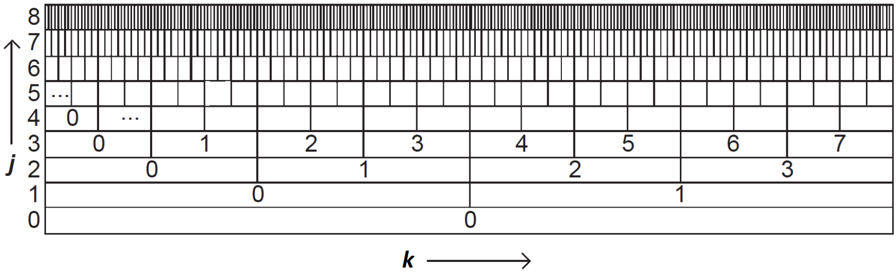
$$Wf(s,u) = \langle f, \psi_{s,u} \rangle = \int_{\mathbf{R}} f(t)\psi_{s,u}(t)dt$$

Calderon's admissible condition

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\xi)|^2}{\xi} d\xi < \infty \qquad [*]$$

Theorem (Calderón, 1966) If [∗] is sastified, then

$$f(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty Wf(s,u)\psi_{s,u}(t)\frac{duds}{s^2}.$$



$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$,

where j indicates the *scale* (in *octaves*) and k the *position*. For each j, the position is sampled at 2^j points.

Synthesis or reconstruction formula

$$f(t) = \sum_{j\in\mathbf{Z}}\sum_{k\in\mathbf{Z}}\langle f, \psi_{j,k}\rangle\psi_{j,k}(t) + R(t),$$

where $R(t)$ is a *residual*, and assuming the $\psi_{j,k}$ form an orthonormal basis.

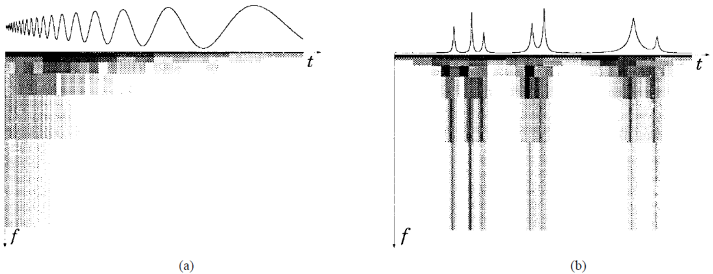
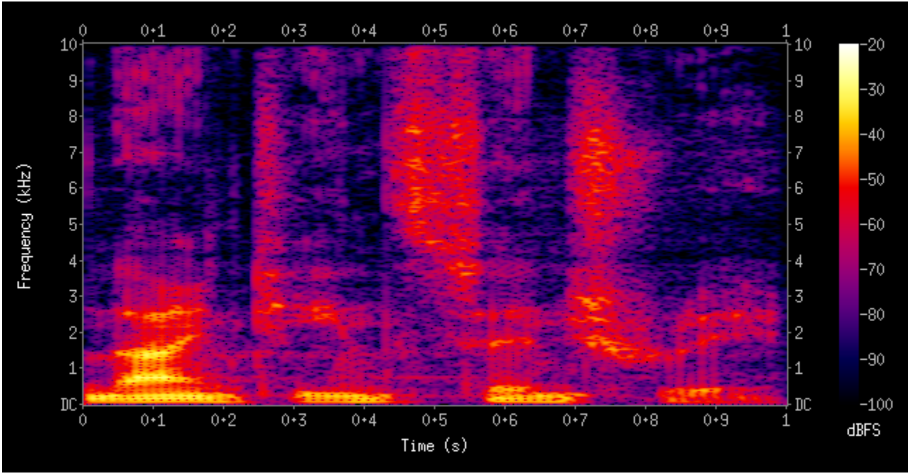


Fig. 2. Clustering and persistence illustrated, respectively, in Donoho and Johnstone's (a) Doppler and (b) Bumps test signals [1]. The signals lie atop the time–frequency tiling (Fig. 1) provided by a seven-scale wavelet transform. Each tile is colored as a monotonic function of the wavelet coefficient energy w_i^2, with darker tiles indicating greater energy.

From [11, Fig. 2]: Clustering and persistence illustrated, respectively, in Donoho and Johnstone's [see [12]] (a) Doppler and (b) Bumps test signals [1]. The signals lie atop the time–frequency tiling provided by a seven-scale wavelet transform. Each tile is colored as a monotonic function of the wavelet coefficient energy $w_{j,k}^2$ with darker tiles indicating greater energy.



From Wikepedia/Spectrogram.

Multiresolution analysis

Definitions and notations

The MRA wavelet

The Haar MRA

In practice

The fast wavelet transform

A *multiresolution analysis* (MRA; Mallat, 1989) of $L^2(\mathbf{R})$ is a sequence of closed subspaces V_j of $L^2(\mathbf{R})$, $j \in \mathbf{Z}$,

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset V_3 \subset \cdots \subset L^2(\mathbf{R})$$

with the following properties:

(1) $\cap_j V_j = \{0\}$ (*trivial intersection*) and $\cup_j V_j$ is *dense* in $L^2(\mathbf{R})$.

(2) $f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}$ (*scaling property*).

(3) $f(t) \in V_0 \Leftrightarrow f(t-k) \in V_0$ for any $k \in \mathbf{Z}$ (*translational invariance*).

(4) There exists a *scaling function* $\varphi \in V_0$ such that $\{\varphi(t-k)\}_{k\in\mathbf{Z}}$ is an orthonomal basis of V_0.

Notation: $\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k) = \frac{1}{2^{-j/2}}\varphi(\frac{t-2^{-j}k}{2^{-j}})$

- $\{\varphi_{j,k}\}_{k\in\mathbf{Z}}$ is an orthonormal basis of V_j, for all $j \in \mathbf{Z}$.
- $P_j f(t) = \sum_k \langle f, \varphi_{j,k}\rangle \varphi_{j,k}$ ($P_j = P_{V_j}$).
- $V_{j+1} = V_j \perp W_j$, hence $P_{j+1} = P_j + Q_j$, where $Q_j = P_{W_j}$.

Remark. $\oplus_j W_j$ is dense in $L^2(\mathbf{R})$.

Theorem (Mallat). The scaling function φ determines a wavelet ψ such that $\{\psi(t-k)\}_{k\in\mathbf{Z}}$ is an orthonormal basis of W_0.

- The functions $\{\psi_{j,k}\}_{k\in\mathbf{Z}}$ form an orthonormal basis of W_j, and hence $Q_j f(t) = \sum_{k\in\mathbf{Z}}\langle f,\psi_{j,k}\rangle\psi_{j,k}$.
- The set $\{\psi_{j,k}\}_{j,k\in\mathbf{Z}}$ is an orthonormal basis of $L^2(\mathbf{R})$.

Scaling function

$$\varphi(t) = \begin{cases} 1 & \text{if } 0 \leqslant t < 1 \\ 0 & \text{otherwise} \end{cases}$$

- $\varphi(t-k)$ is 1 for $k \leqslant t < k+1$ and 0 otherwise.

V_0: closure of the span of the functions $\varphi(t-k)$: functions in $L^2(\mathbf{R})$ that are *locally constant and with jumps at the integers*.

V_j: is the closure of the span of $\{\varphi_{j,k}\}_{k\in\mathbf{Z}}$: functions in $L^2(\mathbf{R})$ that are locally constant with jumps only at the integer multiples of 2^{-j}.

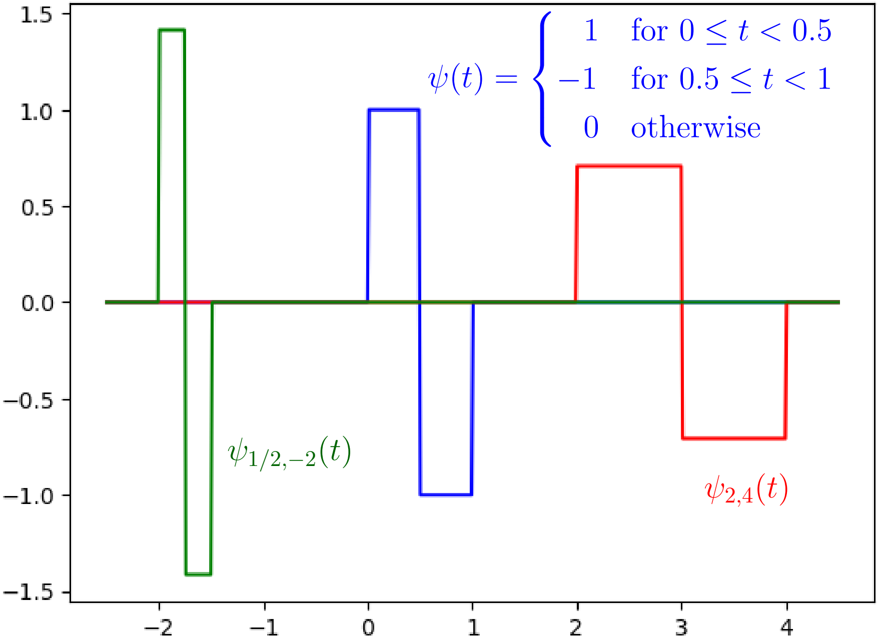
Haar wavelet

We have met it before: $\psi(t)$ is 1 in $0 \leqslant t < 1/2$, -1 in $1/2 \leqslant t < 1$, and 0 elswhere.

W_0: closure of $\{\psi(t-k)\}_{k\in\mathbf{Z}}$: locally constant functions in $L^2(\mathbf{R})$ with jumps only at half-integers and average 0 between any two integers.

$$\psi_{j,k}(t) = \begin{cases} 2^j & \text{for } 2^{-j}k \leqslant t < 2^{-j}(k+1/2) \\ -2^j & \text{for } 2^{-j}(k+1/2) \leqslant t < 2^{-j}(k+1) \\ 0 & \text{otherwise} \end{cases}$$

W_j: closure of $\{\psi_{j,k}\}_{k\in\mathbf{Z}}$: locally constant functions in $L^2(\mathbf{R})$ with jumps only at integer multiples 2^{j+1} and with average 0 between any two integer multiples of 2^j.


1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2
−1
0
1
2
3
4
$\psi(t) = \begin{cases} 1 & \text{for } 0 \leq t < 0.5 \\ -1 & \text{for } 0.5 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$
$\psi_{1/2,-2}(t)$
$\psi_{2,4}(t)$

- The Haar functions form an orthonormal basis of $L^2(\mathbf{R})$.
- The functions $\{\varphi, \psi_{j,k} : j \leqslant 0, 0 \leqslant k < 2^j\}$ form a basis of $L^2([0,1])$.

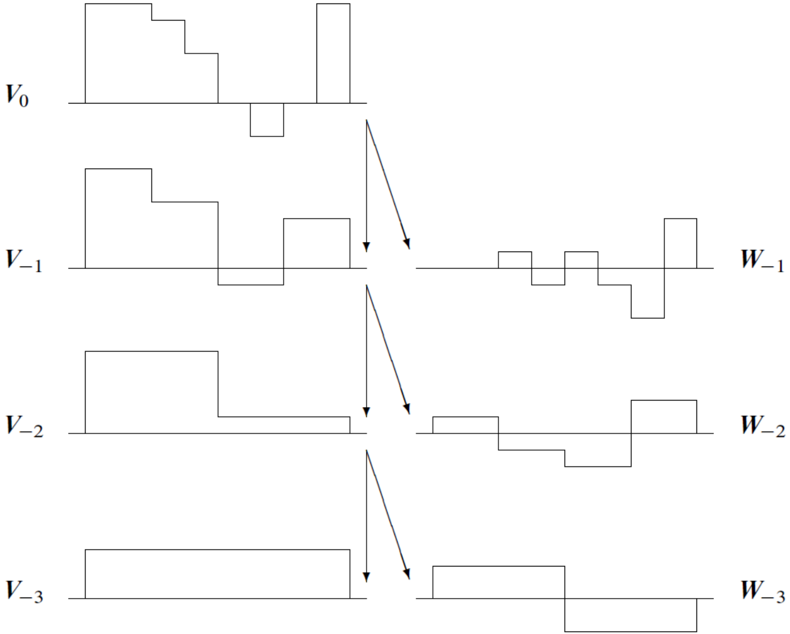
Let us look at the segment

$$V_{-n} \subset V_{-n+1} \subset \cdots \subset V_{-1} \subset V_{-1} \subset V_0$$

The V_{-n} contains the coarser representations and V_0 the finer. Then

$$\begin{aligned} V_0 &= V_{-1} \oplus W_{-1} \\ &= V_{-2} \oplus W_{-1} \oplus W_{-2} \\ &\dots \\ &= V_{-n} \oplus W_{-1} \oplus W_{-2} \oplus \cdots \oplus W_{-n} \end{aligned}$$

From [5, Fig. 14]



- The scaling function φ satisfies the *scaling property*:

$$\varphi(t) = \varphi(2t) + \varphi(2t-1).$$

This implies that

$$\varphi_{j,k} = (\varphi_{j+1,2k} + \varphi_{j+1,2k+1})/\sqrt{2}$$

and so we have the recurrence relations for the *approximation coefficients*

$$\langle f, \varphi_{j,k}\rangle = \tfrac{1}{\sqrt{2}}(\langle f, \varphi_{j+1,2k}\rangle + \langle f, \varphi_{j+1,2k+1}\rangle)$$

The wavelet ψ satisfies the scaling relation

$$\psi(t) = \psi(2t) - \psi(2t-1)$$

which implies a recurrent relation for the *detail coefficients*

$$\langle f, \psi_{j,k}\rangle = \tfrac{1}{\sqrt{2}}(\langle f, \varphi_{j+1,2k}\rangle - \langle f, \varphi_{j+1,2k+1}\rangle)$$

We have $\varphi \in V_0 \subset V_1$, and the $\varphi_{1,k}(t) = \sqrt{2}\varphi(2t-k)$ is an orthonormal basis of V_1, so we have a *scaling equation*

$$\varphi(t) = \sum_k h_k \varphi_{1,k}(t) = \sqrt{2}\sum_k h_k \varphi(2t-k),$$

for some coefficients h_k.

Similarly, there are coefficients g_k such that

$$\psi(t) = \sum_k g_k \varphi_{1,k}(t) = \sqrt{2}\sum_k g_k \varphi(2t-k).$$

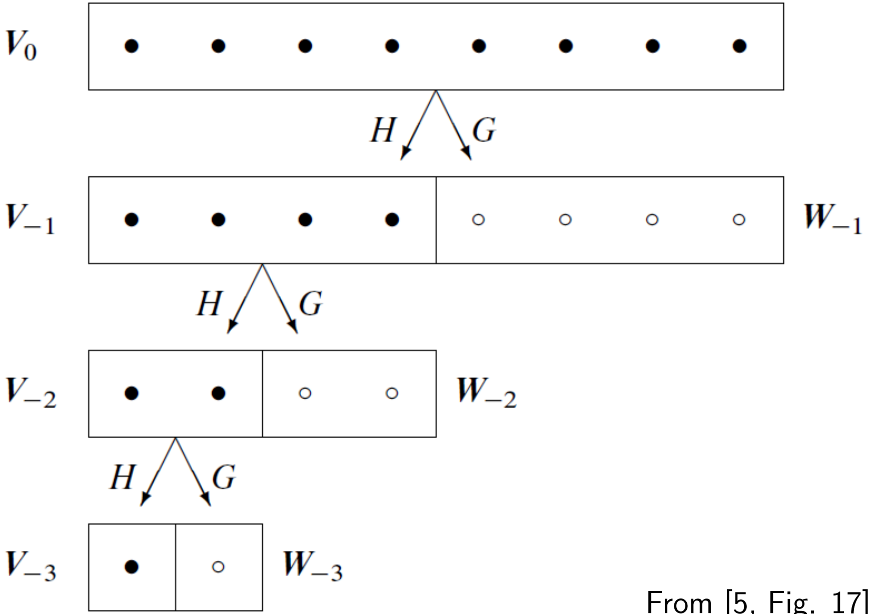
Example. In the Haar case, the only non-zero coefficients are $h_0 = h_1 = 1/\sqrt{2}$ and $g_0 = -g_1 = 1/\sqrt{2}$. In particular, the sums in the scaling equations are finite.

We will assume, as we can for many wavelets, that in the scaling equations there are only L non-zero terms.

Start with the data $\{a_{J,k} = \langle f, \varphi_{J,k}\rangle\}_{0\leqslant k<2^J}$, and define $a_{j,k} = \langle f, \varphi_{j,k}\rangle$ and $d_{j,k} = \langle f, \psi_{j,k}\rangle$ for each scale $j < J$. Then

$$a_{j,k} = \sum_{n=0}^{2^{J-j}-1} \bar{h}_n a_{j+1,n+2k},$$

$$d_{j,k} = \sum_{n=0}^{2^{J-j}-1} \bar{g}_n a_{j+1,n+2k}.$$



From [5, Fig. 17].

For the relation with *filter banks*, see [5, §3.3.2].

For the *Daubechies* and similar families of wavelets, see [5, §3.4].

Spectral Techniques on Graphs

[13] (bauer-2012)
[14] (dong-2017)
[15] (pan-chen-ortega-2020)
[16] (wu-pan-chen-long-zhang-philip-2021)
[17] (gama-ribeiro-bruna-2018)

References I

[1] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," *arXiv preprint arXiv:2104.13478*, 2021.

[2] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "The spectral graph wavelet transform: Fundamental theory and fast computation," in *Vertex-Frequency Analysis of Graph Signals*, pp. 141–175, Springer, 2019.

[3] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[4] M. Stephane, "A wavelet tour of signal processing: The sparse way *(third edition)*," 2009.

References II

[5] M. J. Mohlenkamp and M. C. Pereyra, *Wavelets, their friends, and what they can do for you*, vol. 8.

European Mathematical Society, 2008.

[6] Y. Katznelson, *An introduction to harmonic analysis*.

Cambridge University Press, 2004.

Third corrected edition. The first appeared in 1968 and the second in 2002.

[7] E. Hernández and G. Weiss, *A first course on wavelets*.

CRC press, 1996.

Foreword by Yves Meyer.

References III

[8] N. Young, *An introduction to Hilbert space*.

Cambridge university press, 1988.

[9] J. B. Conway, *A course in functional analysis (2nd edition)*, vol. 96 of *Graduate Texts in Mathematics*.

Springer, 2007.

[10] T. W. Körner, *Fourier analysis*.

Cambridge university press, 1988.

[11] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Transactions on signal processing*, vol. 46, no. 4, pp. 886–902, 1998.

References IV

[12] D. L. Donoho and I. M. Johnstone, “Adapting to unknown smoothness via wavelet shrinkage,” *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, 1995.

[13] F. Bauer, “Normalized graph Laplacians for directed graphs,” *Linear Algebra and its Applications*, vol. 436, no. 11, pp. 4193–4222, 2012.

[14] B. Dong, “Sparse representation on graphs by tight wavelet frames and applications,” *Applied and Computational Harmonic Analysis*, vol. 42, no. 3, pp. 452–479, 2017.

[15] C. Pan, S. Chen, and A. Ortega, “Spatio-temporal graph scattering transform,” 2020.

References V

[16] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 1, 2021.

https://arxiv.org/pdf/1901.00596.pdf.

[17] F. Gama, A. Ribeiro, and J. Bruna, "Diffusion scattering transforms on graphs," 2018.

https://arxiv.org/pdf/1806.08829.pdf.

Harmonic analysis is the study of *objects* (functions, measures, etc.), defined on *topological groups*. The group structure enters into the study by allowing the consideration of the *translates* of the object under study, that is, by placing the object in a translation-invariant space. The study consists of two steps. First: finding the "elementary components" of the object, that is, objects of the same or similar class, which exhibit the simplest behavior under translation and which "belong" to the object under study (*harmonic or spectral analysis*); and
Second: finding a way in which the object can be construed as a combination of its elementary components (*harmonic or spectral synthesis*).

From the Preface of [6].

P

In this session, we consider real and complex spaces (mainly complex spaces).

If $\mathcal{B}$ is finitely-generated, the notion of basis coincides with the notion introduced in elementary linear algebra: a finite set of linearly independent vectors that span the space. Since the results we will state turn out to be obvious in this case, we will assume that $\mathcal{B}$ is not finitely generated.

P

In the complex case, an inner-product is subject to the following properties:

(1) It is *linear with respect to the first variable*, and

(2) it is *conjugate-symmetric*, $\langle x, y\rangle = \langle y, x\rangle^{-}$, which implies that it is *conjugate-linear* with respect to the second variable: $\langle x, \lambda y\rangle = \bar{\lambda}\langle x, y\rangle$.

In particular, $\langle x, x\rangle$ is real for any x, and the condition for being *semi-positive* (*positive*) is the same as for the real case: $\langle x, x\rangle \geqslant 0$ for all x ($\langle x, x\rangle > 0$ for all $x \neq 0$).

The formal series [*], $\sum_{n\geqslant 1}\langle h, \psi_n\rangle\psi_n$, is called the *Fourier series* of h with respect to $\{\psi_n\}$. So the statement says that the Fourier series of h converges in norm to h.

P

In the *Théorie analytique de la chaleur*, the synthesis and analysis was phrased in terms of *periodic functions*.

The relation with $\mathcal{C}$ is that any $f \in \mathcal{C}$ can be prolongued to a function $\bar{f}$ defined on $\mathbf{R}$ that is *periodic* of *period* 1:

$\bar{f}(x) = f(x - [x])$, where $[x]$ is the *integer part* of x, that is, the greatest integer that does not exceed x.

In the formula $\hat{f}(n) = \int_0^1 f(t)e^{-2\pi int}dt$, the factor $e^{-2\pi int}$ never vanishes, so for each n all values $f(t)$ contribute to $\hat{f}(n)$. In particular, a local change in $f(t)$ afects all $\hat{f}(n)$.

P

$$\begin{aligned}\widehat{f'}(\xi) &= \int_{\mathbf{R}} f'(t)e^{-2\pi i\xi t}dt \\ &= e^{-2\pi i\xi t}\, d(f(t)) \\ &= e^{-2\pi i\xi t}f(t)]_{-\infty}^{+\infty} - \int_{\mathbf{R}} f(t)d(e^{-2\pi i\xi t}) \\ &= 2\pi i\xi\hat{f}(\xi).\end{aligned}$$

P