Lecture 5 – Wavelet Transform – Time and Scale

David Walnut
Department of Mathematical Sciences
George Mason University
Fairfax, VA USA

Chapman Lectures, Chapman University, Orange, CA
6-10 November 2017

Outline

- The Continuous Wavelet transform of Grossman and Morlet
- The CWT as a time-frequency (time-scale) transformation
- Relation to the Calderon Reproducing Formula
- Discrete Wavelet decompositions of Frazier and Jawerth
- Relation to Littlewood-Paley theory

Time shifts and scaling

Definition

Given $a > 0$, $b \in \mathbb{R}$, define the *dilation operator* on $L^2(\mathbb{R})$ by $D_a f(x) = a^{1/2} f(ax)$ and the *time-shift* operator T_b by $T_b f(x) = f(x - b)$. Note that $(D_a f)^\wedge(\gamma) = (D_{1/a}\widehat{f})(\gamma)$, and that $(T_b f)^\wedge(\gamma) = e^{-2\pi i b\gamma}\widehat{f}(\gamma)$.

- Suppose $g(x)$ is a bump function centered at 0.
  - If $a > 1$, then $D_a g$ is more concentrated near 0 than g is.
  - If $0 < a < 1$, then $D_a g$ is more spread out near 0 than g is.
  - The function $D_a T_b g$ will be a bump function centered at $a^{-1}b$.
- Note that $D_a T_b g(x) = a^{1/2} g(ax - b)$.

Continuous Wavelet Transform

- Grossmann and Morlet (1986) define a distribution analogous to the STFT but which draws out features of a function related to “time” and “scale.”
- The distribution is analogous to a “coherent state” decomposition as the analyzing functions are transforms of a single function.
- The idea was proposed earlier for the analysis of seismic traces related to oil exploration.
- The distribution has properties analogous to the type of time-frequency distribution discussed earlier.

Definition

Given a function $g \in L^2(\mathbb{R})$, the *continuous wavelet transform* of a function $f \in L^2$ is defined by

$$W_g(f)(a,b) = \int_{-\infty}^{\infty} f(t)\, a^{1/2}\overline{g(at-b)}\, dt = \langle f, D_a T_b g \rangle_{L^2(\mathbb{R})}$$

for $a > 0$ and $b \in \mathbb{R}$.

- $W_g(f)(a,b) = f * D_a g(b/a)$.
- By Plancherel,

$$\begin{aligned} W_g(f)(a,b) &= \langle f, D_a T_b g \rangle = \langle \widehat{f}, D_{1/a} M_b \widehat{g} \rangle \\ &= \langle D_a \widehat{f}, M_b \widehat{g} \rangle = \left[D_a \widehat{f}\, \overline{\widehat{g}} \right]^{\wedge}(b). \end{aligned}$$

Isometry

$$
\begin{aligned}
\int_0^\infty \int_{-\infty}^\infty |W_g f(a,b)|^2 db \, \frac{da}{a} &= \int_0^\infty \int_{-\infty}^\infty |(D_a \widehat{f}\, \overline{\widehat{g}})^\wedge(b)|^2 db \, \frac{da}{a} \\
&= \int_0^\infty \int_{-\infty}^\infty a|\widehat{f}(a\gamma)|^2 \, |\widehat{g}(\gamma)|^2 d\gamma \, \frac{da}{a} \\
&= \int_{-\infty}^\infty |\widehat{g}(\gamma)|^2 \left[\int_0^\infty |\widehat{f}(a\gamma)|^2 \, da \right] d\gamma \\
&= \left(\int_{-\infty}^\infty \frac{|\widehat{g}(\gamma)|^2}{|\gamma|} \, d\gamma \right) \|f\|_2^2.
\end{aligned}
$$

Admissibility

- We say that g is *admissible* if it satisfies
$$\int_{-\infty}^{\infty} \frac{|\widehat{g}(\gamma)|^2}{|\gamma|}\, d\gamma < \infty.$$
- The mapping $W_g \colon L^2(\mathbb{R}, dx) \to L^2((0,\infty) \times \mathbb{R}, db\, da/a)$ is an isometry if and only if g is admissible.
- Intuitively, admissibility means that $\widehat{g}(0) = 0$ or that $\int g(x)\, dx = 0$.
- If $\operatorname{supp} \widehat{g} \subseteq [0, \infty)$ then W_g naturally restricts to that subspace of L^2 consisting of functions whose Fourier transforms are supported on the half-line.

Hardy space H^2

Definition

Define the *Hardy space* on the upper half plane, $H^2(D_+)$ as the space of all holomorphic functions on D_+ such that

$$\|f\|_{H^2} = \sup_{y>0} \left(\int_{-\infty}^{\infty} |f(x+iy)|^2 \, dx \right)^{1/2} < \infty.$$

Theorem (Paley-Wiener, 1934)

A holomorphic function $f \in H^2(D_+)$ if and only if its Fourier transform vanishes on the negative half-line.

- This means that if $g \in H^2$ is admissible, then W_g can be thought of as an isometry on H^2.
- By symmetry, if g is admissible and $\widehat{g}$ vanishes on the positive half-line then W_g is an isometry on the subspace of L^2 consisting of functions whose reflection about the origin is in H^2.
- If g is admissible and $\widehat{g}$ is symmetric about the origin, W_g is an isometry on $L^2(\mathbb{R})$.

Invertibility

- If g is admissible and if $c_g = \int_{-\infty}^{\infty} |\widehat{g}(\gamma)|^2 \, |\gamma|^{-1} \, d\gamma$, then $(c_g)^{-1/2} W_g$ is an isometry.
- By the polarization identity we obtain the following inversion formula for $f \in L^2(\mathbb{R})$.

$$f(x) = \frac{1}{\sqrt{c_g}} \int_0^{\infty} \int_{-\infty}^{\infty} Wgf(a, b) \, a^{1/2} g(ax - b) \, db \, \frac{da}{a}$$

where the integral is interpreted weakly.

In the spring of 1985, Yves Meyer recognized that a recovery formula found by Morlet and Alex Grossmann was an identity previously discovered by Alberto Calderón. At that time, Yves Meyer was already a leading figure in the Calderón–Zygmund theory of singular integral operators. Thus began Meyer's study of wavelets, which in less than ten years would develop into a coherent and widely applicable theory.

– Yves Meyer's Abel Prize announcement, Norwegian Academy of Sciences

Calderón Reproducing Formula

- Let $\varphi \in C^\infty(\mathbb{R}^d)$ be radial such that
  - $\operatorname{supp}\varphi \subseteq \{x\colon \|x\| \le 1\}$,
  - $\displaystyle\int_{\mathbb{R}^d} x^\alpha \varphi(x)\,dx = 0$ for $|\alpha| \le N$ some $N \in \mathbf{N}$, and
  - $\displaystyle\int_0^\infty |\widehat{\varphi}(t\gamma)|^2 \,\frac{dt}{t} = 1$ for all $\gamma \in \mathbb{R}^d \setminus \{0\}$.
- Then for all $f \in L^2(\mathbb{R}^d)$,

$$f(x) = \int_0^\infty \varphi_t * \varphi_t * f(x)\,\frac{dt}{t}$$

where $\varphi_t(x) = t^{-d}\varphi(x/t)$, $t > 0$.

Proof:

To see formally why the Calderón Reproducing Formula holds, note that

$$\widehat{\varphi}_t(\gamma) = \widehat{\varphi}(t\gamma).$$

Hence

$$\begin{aligned} & \left(\int_0^\infty \varphi_t * \varphi_t * f(\cdot) \, \frac{dt}{t} \right)^\wedge (\gamma) \\ = & \int_0^\infty (\varphi_t * \varphi_t * f)^\wedge(\gamma) \, \frac{dt}{t} \\ = & \int_0^\infty |\widehat{\varphi}(t\gamma)|^2 \widehat{f}(\gamma) \, \frac{dt}{t} = \widehat{f}(\gamma) \end{aligned}$$

- This formal calculation can be rigorized by a limiting process.

Atomic Decompositions

- The Calderón formula allows for the characterization of certain function spaces in terms of so-called *atomic decompositions*.
- This work bears the names of Calderón (1977), Calderón and Torchinski (1975, 77), Coifman and Weiss (1977), Tabileson and Weiss (1980), and Uchiyama (1982).
- The idea is to use the atomic decompositions to prove that so-called Calderón-Zygmund operators are bounded on a large class of function spaces.
- A *CZO*, T, has the form

$$Tf(x) = \int_{\mathbb{R}^d} K(x, y)\, f(y)\, dy$$

  where the kernel K is continuous off the diagonal $x = y$ and is singular on the diagonal.

Atomic Decompositions

- Define (on $\mathbb{R}$ for convenience) the *dyadic intervals*

$$\mathcal{D} = \{[2^j k, 2^j (k+1)] \colon j,\, k \in \mathbb{Z}\},$$

and for each $I \in \mathcal{D}$, define the cube $Q_I \subseteq \mathbb{R} \times (0, \infty)$ by

$$Q_I = I \times [|I|/2, |I|].$$

- The decompositions take the form $f = \sum_{I \in \mathcal{D}} s_I \, a_I(x)$.
- Here the atoms a_I are compactly supported near I, and encode the oscillatory behavior of f near I.
- Certain function spaces can be characterized in terms of the *magnitude* of the coefficients $\{s_I\}$.

- More specifically, we define

$$a_I(x) = \frac{1}{s_I} \int_{Q_I} \varphi(x-y)(\varphi_t * f)(y)\, dy \frac{dt}{t}$$

where the coefficients s_I are chosen based on the function space.

- For example, for $L^2(\mathbb{R})$,

$$s_I = \left(\int_{Q_I} |(\varphi_t * f)(y)|^2\, dy \frac{dt}{t} \right)^{1/2}$$

and $\sum_I (s_I)^2 = \|f\|_2^2$.

- Note that $Q_I = [2^j k, 2^j(k+1)] \times [2^{j-1}, 2^j]$, $j,\ k \in \mathbb{Z}$.
- Hence by Plancherel,

$$
\begin{aligned}
\sum_I (s_I)^2 &= \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \int_{2^{j-1}}^{2^j} \int_{2^j k}^{2^j(k+1)} |(\varphi_t * f)(y)|^2\, dy \frac{dt}{t} \\
&= \sum_{j\in\mathbb{Z}} \int_{2^{j-1}}^{2^j} \sum_{k\in\mathbb{Z}} \int_{2^j k}^{2^j(k+1)} |(\varphi_t * f)(y)|^2\, dy \frac{dt}{t} \\
&= \sum_{j\in\mathbb{Z}} \int_{2^{j-1}}^{2^j} \int_{\mathbb{R}} |\widehat{\varphi}_t(\gamma)|^2 |\widehat{f}(\gamma)|^2\, d\gamma \frac{dt}{t} \\
&= \int_{\mathbb{R}} \left(\int_0^\infty |\widehat{\varphi}(t\gamma)|^2 \frac{dt}{t} \right) |\widehat{f}(\gamma)|^2\, d\gamma \\
&= \int_{\mathbb{R}} |\widehat{f}(\gamma)|^2\, d\gamma.
\end{aligned}
$$

Lipschitz Spaces

- These decompositions hold for spaces defined in terms of the local smoothness or oscillatory behavior of functions on $\mathbb{R}^d$.
- Example: Lipschitz Spaces. Given $0 < \alpha < 1$, define

$$\dot{\Lambda}_\alpha = \{f\colon |f(x) - f(y)| \leq C|x - y|^\alpha\}$$

with

$$\|f\|_{\dot{\Lambda}_\alpha} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}.$$

- It can be shown that $\|f\|_{\dot{\Lambda}_\alpha}$ is equivalent to

$$\sup_{y \in \mathbb{R}, t > 0} |t^{-\alpha}(\varphi_t * f)(y)|.$$

- This leads to an atomic decomposition of $\dot{\Lambda}_\alpha$ with

$$s_I = |I|^{-\alpha - 1/2} \|f\|_{\dot{\Lambda}_\alpha}$$

and a_I as above.

Besov and Triebel-Lizorkin Spaces

- These spaces are generalizations of Lipschitz spaces and characterized by local smoothness and global decay properties.
- For $\alpha \in \mathbb{R}$, $0 < p,\ q \leq \infty$, the *Besov space*, $\dot{B}_p^{\alpha,q}$ is characterized by

$$\|f\|_{\dot{B}_p^{\alpha,q}} = \left(\int_0^\infty t^{-\alpha q} \|\varphi_t * f\|_p^q \, \frac{dt}{t} \right)^{1/q} < \infty.$$

- For $\alpha \in \mathbb{R}$, $0 < p < \infty$, $0 < q \leq \infty$, the *Triebel-Lizorkin space*, $\dot{F}_p^{\alpha,q}$ is characterized by

$$\|f\|_{\dot{F}_p^{\alpha,q}} = \left\| \left(\int_0^\infty t^{-\alpha q} |(\varphi_t * f)(y)|^q \, \frac{dt}{t} \right)^{1/q} \right\|_p < \infty.$$

- In each case, we choose

$$s_I = |I|^{1/2} \sup_{y \in I} |(\varphi_t * f)(y)|$$

and obtain the proper atomic characterization of the spaces in terms of $|s_I|$.

- In the Besov case, $\|f\|_{\dot{B}_p^{\alpha,q}}$ is equivalent to

$$\left(\sum_{\nu \in \mathbb{Z}} \left(\sum_{|I| = 2^{-\nu}} (|I|^{-1/2-\alpha+1/p} |s_I|^p \right)^{q/p} \right)^{1/q}.$$

- In the Triebel-Lizorkin case, $\|f\|_{\dot{F}_p^{\alpha,q}}$ is equivalent to

$$\left\| \left(\sum_I (|I|^{-1/2-\alpha} |s_I| \mathbf{1}_I)^q \right)^{1/q} \right\|_p.$$

- Notice that by this construction, both the coefficients *and the atoms* depend on f.
- Here the atoms a_I are compactly supported near I, and encode the oscillatory behavior of f near I.
- Membership of f in certain function spaces can be characterized in terms of the *magnitude* of the coefficients $\{s_I\}$.

Littlewood-Paley Theory

- Littlewood-Paley theory addressed the question of characterizing $L^p(\mathbb{T})$ ($\mathbb{T}$ is the torus) in terms of Fourier coefficients.
- The difficulty takes the form of the following classical theorem.

Theorem

Let $1 \leq p < \infty$, $p \neq 2$, and suppose that $\sum_n c_n\, e^{2\pi inx}$ is the Fourier series of a function in $L^p(\mathbb{T}) \setminus L^2(\mathbb{T})$. Then for almost every choice of $\epsilon_n = \pm 1$, the series $\sum_n \epsilon_n c_n\, e^{2\pi inx}$ is not the Fourier series of a function in $L^p(\mathbb{T})$.

Littlewood-Paley Theory

- For $\sum_n c_n \, e^{2\pi i n t}$ the Fourier series of some $f \in L^p(\mathbb{T})$, define $\Delta_0 f = c_0$ and for $N \in \mathbf{N}$,

$$(\Delta_{N+1} f)(x) = \sum_{2^N \leq |n| < 2^{N+1}} c_n \, e^{2\pi i n x}.$$

  Let

$$d(f)(x) = \left(\sum_{N=0}^{\infty} |(\Delta_N f)(x)|^2 \right)^{1/2}.$$

  Then for $1 < p < \infty$, there exist constants A_p, B_p such that for all $f \in L^p(\mathbb{T})$,

$$A_p \|f\|_p \leq \|d(f)\|_p \leq B_p \|f\|_p.$$

- The key point here is the use of dyadic blocks of the Fourier transform.

Discrete Calderón Formula

- Given φ as in the Calderón reproducing formula, observe that for $\nu \in \mathbb{Z}$, the function $\varphi_{2^{-\nu}} * f$ will pick out the frequency content of f in the range $|\gamma| \approx 2^{\nu}$.
- Let $\varphi, \psi \in C^{\infty}(\mathbb{R}^d)$ be radial such that
  - $\operatorname{supp} \psi \subseteq \{x \colon \|x\| \leq 1\}$,
  - $\displaystyle\int_{\mathbb{R}^d} x^{\alpha} \psi(x)\, dx = 0$ for $|\alpha| \leq N$ some $N \in \mathbf{N}$, and
  - $\operatorname{supp} \widehat{\varphi} \subseteq \{\gamma \colon 1/2 \leq \|x\| \leq 2\}$,
  - $|\widehat{\varphi}(\gamma)| \geq c > 0$ if $\{\gamma \colon 3/5 \leq \|x\| \leq 5/3\}$,
  - $\displaystyle\sum_{\nu \in \mathbb{Z}} \widehat{\psi}(2^{-\nu}\gamma)\widehat{\varphi}(2^{-\nu}\gamma) = 1$ for all $\gamma \in \mathbb{R}^d \setminus \{0\}$.
- Then for all $f \in L^2(\mathbb{R}^d)$, $f(x) = \displaystyle\sum_{\nu \in \mathbb{Z}} \psi_{2^{-\nu}} * \varphi_{2^{-\nu}} * f(x)$.

Characterizations

We can characterize several classical function spaces in terms of the *g-function*:

$$g_d(f)(x) = \left(\sum_{\nu \in \mathbb{Z}} |\varphi_{2^{-\nu}} * f(x)|^2 \right)^{1/2}.$$

- L^p spaces, $1 < p < \infty$: $\|f\|_p \approx \|g_d(f)\|_p$.
- Hardy spaces H^p, $0 < p \leq 1$:
  $$\|f\|_{H^p} \approx \left(\int |g_d(f)(x)|^p \, dx \right)^{1/p}.$$
- Sobolev spaces L_k^p, $1 < p < \infty$, $k \in \mathbf{N}$:
  $$\|f\|_{L_k^p} \approx \left(\int \left(\sum_{\nu \in \mathbb{Z}} |2^{\nu k} |\varphi_{2^{-\nu}} * f(x)|^2| \right)^{p/2} dx \right)^{1/p}.$$

Besov and Triebel-Lizorkin spaces

- In addition we can characterize these spaces in a similar way using the discrete Calderón formula.
- $\dot{B}_p^{\alpha,q}$: $\|f\|_{\dot{B}_p^{\alpha,q}} \approx \left(\sum_{\nu\in\mathbb{Z}} 2^{\nu\alpha q} \|\varphi_{2^{-\nu}} * f\|_p^q \right)^{1/q}$.
- $\dot{F}_p^{\alpha,q}$: $\|f\|_{\dot{B}_p^{\alpha,q}} \approx \left\| \left(\sum_{\nu\in\mathbb{Z}} 2^{\nu\alpha q} |(\varphi_{2^{-\nu}} * f)(y)|^q \right)^{1/q} \right\|_p$.

The φ-transform

- Frazier and Jawerth (1988) unified atomic decomposition techniques by designing a “proto-wavelet basis” they dubbed the φ-transform.
- The idea was to obtain useful atomic decompositions in which the atoms were independent of the function being analyzed.
- The decompositions had the form

$$f = \sum_{I \in \mathcal{D}} \langle f, \varphi_I \rangle \, \psi_I$$

  where φ_I, ψ_I are concentrated near I in the spatial variable and near $[|I|^{-1}, 2|I|^{-1}] \cup [-|I|^{-1}, -2|I|^{-1}]$ in the frequency variable.
- In fact, the collections $\{\varphi_I \colon I \in \mathcal{D}\}$ and $\{\psi_I \colon I \in \mathcal{D}\}$ are frames for $L^2(\mathbb{R})$.

The φ-transform construction

- Let $\varphi,\ \psi \in C^\infty(\mathbb{R}^d)$ be radial such that
  - $\operatorname{supp} \widehat{\varphi} \subseteq \{\gamma \colon 1/2 \leq \|\gamma\| \leq 2\}$,
  - $|\widehat{\varphi}(\gamma)| \geq c > 0$ if $\{\gamma \colon 3/5 \leq \|\gamma\| \leq 5/3\}$,
  - ψ satisfies the same conditions, and
  - $\displaystyle\sum_{\nu \in \mathbb{Z}} \overline{\widehat{\psi}(2^{-\nu}\gamma)}\widehat{\varphi}(2^{-\nu}\gamma) = 1$ for all $\gamma \in \mathbb{R}^d \setminus \{0\}$.
- Then for all $f \in L^2(\mathbb{R}^d)$,

$$f = \sum_{I \in \mathcal{D}} \langle f, \varphi_I \rangle\, \psi_I$$

where $h_I(x) = 2^{\nu/2} h(2^\nu x - k)$ when $I = [2^{-\nu}k, 2^{-\nu}(k+1)]$.

The φ-transform construction

- Starting with the functions φ and ψ as above, we have

$$f(x) = \sum_{\nu \in \mathbb{Z}} \psi_{2^{-\nu}} * \widetilde{\varphi}_{2^{-\nu}} * f(x)$$

where $\widetilde{\varphi}(x) = \overline{\varphi(-x)}$.

- Using the fact that both $\widehat{\varphi}$ and $\widehat{\psi}$ have the same compact support, we argue as in the Shannon Sampling Formula and obtain

$$f(x) = \sum_{\nu \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} 2^{-\nu/2} (\widetilde{\varphi}_{2^{-\nu}} * f)(2^{-\nu} k)\, 2^{-\nu/2} \psi_{2^{-\nu}}(x - 2^{-\nu} k).$$

- Finally we observe that

$$2^{-\nu/2} \psi_{2^{-\nu}}(x - 2^{-\nu} k) = 2^{\nu/2} \psi(2^{\nu} x - k) = \psi_I$$

and

$$2^{-\nu/2} (\widetilde{\varphi}_{2^{-\nu}} * f)(2^{-\nu} k) = \langle f, \varphi_I \rangle.$$

Orthonormal wavelets

- The coefficients $\{\langle f, \varphi_I \rangle\}$ of the φ-transform play the same role as s_I in the atomic characterization of function spaces that arise in Littlewood-Paley theory.
- In particular, membership in such spaces can be characterized by the magnitude of the coefficients $\{|\langle f, \varphi_I \rangle\}$.
- At the same time, Meyer took this further and constructed a smooth, orthonormal basis consisting of functions of the form

$$\{\psi(2^\nu x - k) \colon \nu,\, k \in \mathbb{Z}\}.$$