Lecture 6 – Orthonormal Wavelet Bases

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

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

Outline

- Wavelet systems
- Example 1: The Haar Wavelet basis
- Characterizations of wavelet bases
- Example 2: The Shannon Wavelet basis
- Example 3: The Meyer Wavelet basis

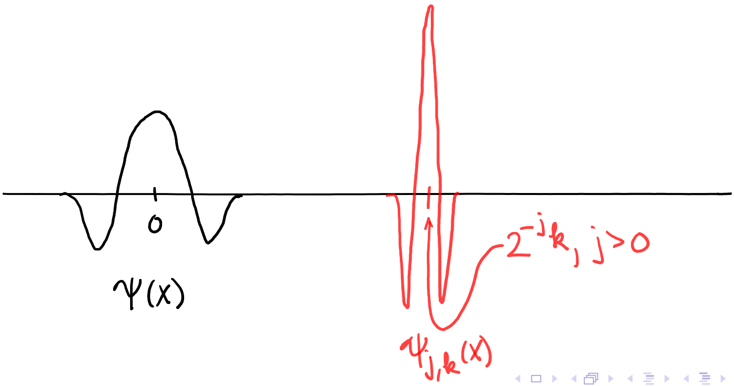
Definition

A *wavelet system* in $L^2(\mathbb{R})$ is a collection of functions of the form

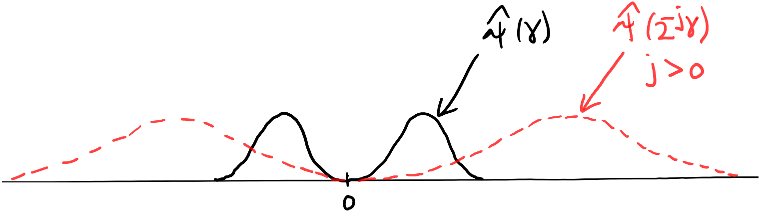
$$\{D_{2^j} T_k \psi\}_{j,k\in\mathbb{Z}} = \{2^{j/2}\psi(2^j x - k)\}_{j,k\in\mathbb{Z}} = \{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$$

where $\psi \in L^2(\mathbb{R})$ is a fixed function sometimes called the *mother wavelet*.
A wavelet system that forms an orthonormal basis for $L^2(\mathbb{R})$ is called a *wavelet orthonormal basis* for $L^2(\mathbb{R})$.

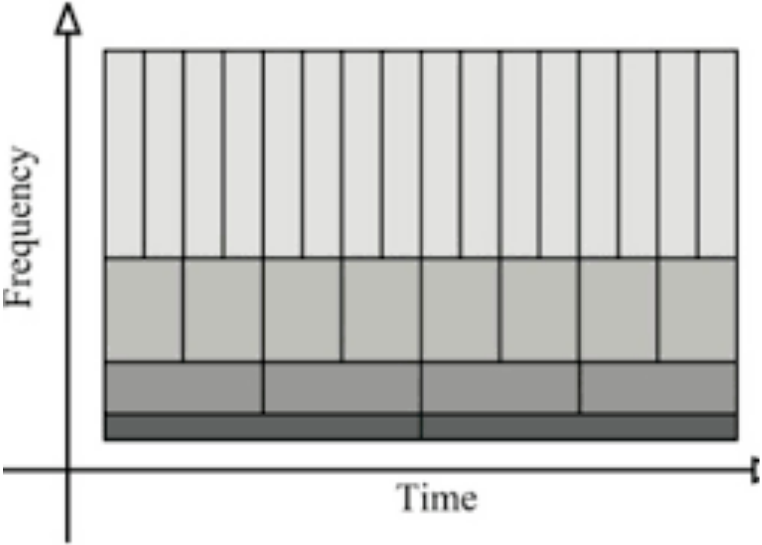
If the mother wavelet $\psi(x)$ is concentrated around 0 then $\psi_{j,k}(x)$ is concentrated around $2^{-j}k$. If $\psi(x)$ is essentially supported on an interval of length L, then $\psi_{j,k}(x)$ is essentially supported on an interval of length $2^{-j}L$.



- Since $(D_{2^j} T_k \psi)^\wedge(\gamma) = D_{2^{-j}} M_k \widehat{\psi}(\gamma)$ it follows that if $\widehat{\psi}$ is concentrated on the interval I then $\widehat{\psi_{j,k}}$ is concentrated on the interval $2^j I$.



Dyadic time-frequency tiling



Example 1: The Haar System

- This is historically the first orthonormal wavelet basis, described by A. Haar (1910) as a basis for $L^2[0, 1]$.
- The Haar basis is an alternative to the traditional Fourier basis but has the property that the partial sums of the series expansion of a continuous function f converges *uniformly* to f.
- We will prove "by hand" that the Haar basis is an orthonormal basis using a technique that will be generalized in the next lecture.

The dyadic intervals

Definition

Define the *dyadic intervals* on $\mathbb{R}$ by

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

We write $I_{j,k} = [2^j k, 2^j(k+1)]$, and refer to the collection

$$\{I_{j,k} \colon k \in \mathbb{Z}\}$$

as the *dyadic intervals at scale j*. $I_{j,k} = I^{\ell}_{j,k} \cup I^{r}_{j,k}$, where $I^{\ell}_{j,k}$ and $I^{r}_{j,k}$ are dyadic intervals at scale $j+1$, to denote the left half and right half of the interval $I_{j,k}$. In fact, $I^{\ell}_{j,k} = I_{j+1,2k}$ and $I^{r}_{j,k} = I_{j+1,2k+1}$.

- Note that given $I_{j,k}$,

$$I_{j,k} = I_{j+1,2k} \cup I_{j+1,2k+1}$$

  where $I_{j+1,2k}$ is the left half of $I_{j,k}$ and $I_{j+1,2k+1}$ the right half of $I_{j,k}$.
- Given $(j_0, k_0) \neq (j_1, k_1)$, either I_{j_1,k_1} and I_{j_0,k_0} are disjoint or one is contained in the other.
- In the latter case, the smaller interval is contained in either the right half or left half of the larger.

The Dyadic step functions

Definition

A *dyadic step function (at scale j)* is a step function in $L^2(\mathbb{R})$ constant on the dyadic intervals $I_{j,k}$, $k \in \mathbb{Z}$. We denote the collection of all such step functions by V_j.

- For each $j \in \mathbb{Z}$, V_j is a linear space.
- If $f \in V_j$ then $f \in V_{j'}$ for all $j' \geq j$, that is, $V_j \subseteq V_{j'}$ if $j' > j$.

The Haar scaling function

Definition

Let $p(x) = \mathbf{1}_{[0,1)}(x)$, and for each $j,\ k \in \mathbb{Z}$, let

$$p_{j,k}(x) = 2^{j/2}\, p(2^j x - k) = D_{2^j} T_k p(x).$$

The collection

$$\{p_{j,k}(x) \colon j, k \in \mathbb{Z}\}$$

is the system of *Haar scaling functions*. For fixed j,

$$\{p_{j,k}(x)\}_{k\in\mathbb{Z}}$$

is the system of *scale j Haar scaling functions*.

- Note that $p_{j,k} = 2^{j/2}\mathbf{1}_{I_{j,k}}$ and that for each j, $\{p_{j,k} \colon k \in \mathbb{Z}\}$ is an orthonormal system.
- In fact $p_{j,k} = 2^{j/2}\mathbf{1}_{I_{j,k}}$ is an orthonormal basis for V_j.
- Since $V_j \subseteq V_{j+1}$, we can write $p_{j,k}$ in terms of the basis $\{p_{j+1,k} \colon k \in \mathbb{Z}\}$, and in fact

$$p_{j,k} = 2^{1/2} p_{j+1,2k} + 2^{1/2} p_{j+1,2k+1}.$$

Definition

Let $h(x) = \mathbf{1}_{[0,1/2)}(x) - \mathbf{1}_{[1/2,1)}(x)$, and for each $j, k \in \mathbb{Z}$, let

$$h_{j,k}(x) = 2^{j/2}\, h(2^j x - k) = D_{2^j} T_k h(x).$$

The collection

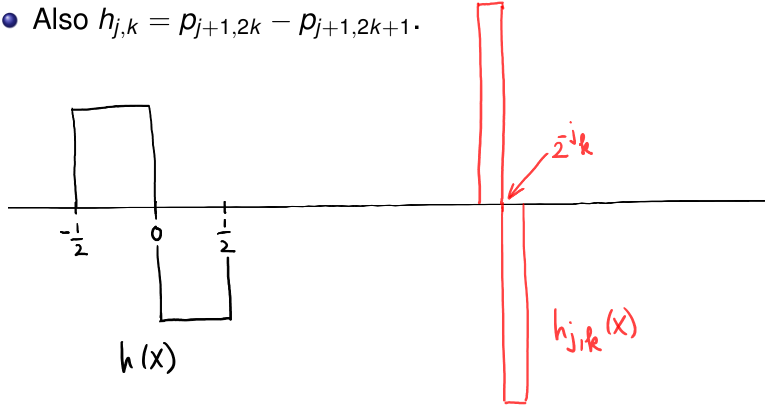
$$\{h_{j,k}(x)\}_{j,k\in\mathbb{Z}}$$

is the *Haar system on* $\mathbb{R}$. For fixed j,

$$\{h_{j,k}(x)\}_{k\in\mathbb{Z}}$$

is the system of *scale j Haar functions*.

- Note that supp $h_{j,k} = I_{j,k}$ and $h_{j,k} \in V_{j+1}$.
- Also $h_{j,k} = p_{j+1,2k} - p_{j+1,2k+1}$.



Theorem

The Haar system is an orthonormal system in $L^2(\mathbb{R})$.

- If $k \neq k' \in \mathbb{Z}$ then $h_{j,k}$ and $h_{j,k'}$ have disjoint supports.
- If $j < j'$ then either $h_{j,k}$ and $h_{j',k'}$ have disjoint supports, or supp $h_{j',k'}$ is contained in an interval on which $h_{j,k}$ is constant.
- In either case, $\langle h_{j',k'}, h_{j,k} \rangle = 0$.

Completeness of the Haar system

Lemma

For $j \in \mathbb{Z}$, define the space W_j by

$$W_j = \overline{\text{span}}\{h_{j,k} \colon k \in \mathbb{Z}\}.$$

Then $V_{j+1} = V_j \oplus W_j$.

Proof:

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} a_k\, p_{j+1,k} &= \sum_{n\in\mathbb{Z}} (a_{2n}\, p_{j+1,2n} + a_{2n+1}\, p_{j+1,2n+1}) \\
&= \sum_{n\in\mathbb{Z}} \frac{a_{2n} + a_{2n+1}}{2} (p_{j+1,2n} + p_{j+1,2n+1}) \\
&\quad + \frac{a_{2n} - a_{2n+1}}{2} (p_{j+1,2n} - p_{j+1,2n+1}) \\
&= \sum_{n\in\mathbb{Z}} \frac{a_{2n} + a_{2n+1}}{2^{1/2}}\, p_{j,n} + \sum_{n\in\mathbb{Z}} \frac{a_{2n} - a_{2n+1}}{2^{1/2}}\, h_{j,n}.
\end{aligned}$$

Completeness of the Haar system

Theorem

The Haar system is an orthonormal basis for $L^2(\mathbb{R})$.

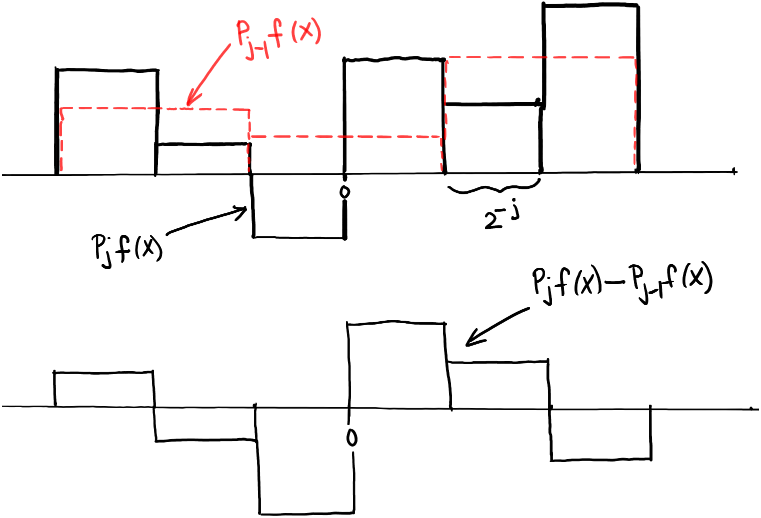
Proof:

- Define the orthogonal projection operators

$$P_j f = \sum_{k \in \mathbb{Z}} \langle f, p_{j,k} \rangle \, p_{j,k}.$$

- Because any L^2 function can be approximated by a dyadic step function, $P_j f \to f$ as $j \to \infty$, and also $P_{-j} f \to 0$ as $j \to \infty$.
- Note that $V_J = \bigoplus_{j=-J+1}^{J-1} W_j \oplus V_{-J}$ which implies that

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$


$P_{j-1}f(x)$
$P_jf(x)$
0
2^{-j}
$P_jf(x) - P_{j-1}f(x)$
0

Smooth wavelet bases

- The Haar basis is the paradigmatic wavelet basis with good time localization. Since $h(x)$ is discontinuous, it has very poor frequency localization.
- One of Meyer's early contributions to the theory was to create an orthonormal wavelet basis with good decay in both time and frequency.
- Before getting to that proof, we will describe the paradigmatic orthonormal wavelet basis with optimal frequency localization, known as the Shannon basis.
- This will again be a proof "by hand" and will give the flavor of Meyer's original proof.

Characterization of orthonormal bases

Lemma

A sequence $\{x_n\}$ in a Hilbert space H is an orthonormal basis for H if and only if

1. *for all $x \in H$, $\sum_n |\langle x, x_n \rangle|^2 = \|x\|^2$, and*
2. *$\|x_n\| = 1$ for all n.*

- In other words, $\{x_n\}$ is a normalized tight frame for H with frame bound 1.
- The proof is trivial: Given m, note that

$$1 = \|x_m\|^2 = \sum_n |\langle x_m, x_n \rangle|^2 = \sum_{n \neq m} |\langle x_m, x_n \rangle|^2 + 1.$$

Hence $\langle x_m, x_n \rangle = 0$ if $n \neq m$ and $\{x_n\}$ is an orthonormal system.

Characterization of wavelet ONB

Theorem (A.)

Suppose that $\psi \in L^2(\mathbb{R})$ *satisfies* $\|\psi\|_2 = 1$*. Then*

$$\{2^{j/2}\psi(2^j x - k) \colon j, k \in \mathbb{Z}\} = \{\psi_{j,k} \colon j, k \in \mathbb{Z}\}$$

is an orthonormal wavelet basis for $L^2(\mathbb{R})$ *if and only if*

1. $\sum_{j \in \mathbb{Z}} |\widehat{\psi}(2^j\gamma)|^2 \equiv 1$, *and*
2. $\sum_{j=0}^{\infty} \widehat{\psi}(2^j\gamma)\, \overline{\widehat{\psi}(2^j(\gamma + k))} = 0$ *for all odd integers* k.

Characterization of wavelet ON systems

Theorem (B.)

Given $\psi \in L^2(\mathbb{R})$*, the wavelet system* $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ *is an orthonormal system in* $L^2(\mathbb{R})$ *if and only if*

1. $\displaystyle\sum_{k\in\mathbb{Z}} |\widehat{\psi}(\gamma+k)|^2 \equiv 1$*, and*
2. $\displaystyle\sum_{k\in\mathbb{Z}} \widehat{\psi}(2^j(\gamma+k))\,\overline{\widehat{\psi}(\gamma+k)} = 0$ *for all* $j \geq 1$*.*

Corollary

Given $\psi \in L^2(\mathbb{R})$, *if* $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ *is an orthonormal basis for* $L^2(\mathbb{R})$ *then*

1. $\displaystyle\sum_{j\in\mathbb{Z}} |\widehat{\psi}(2^j\gamma)|^2 \equiv 1$, *and*
2. $\displaystyle\sum_{k\in\mathbb{Z}} |\widehat{\psi}(\gamma + k)|^2 \equiv 1.$

- This says that a wavelet orthonormal basis must form a partition of unity in frequency both by translation and dilation.
- This implies that, for example, any wavelet $\psi \in L^1 \cap L^2$ must satisfy $\widehat{\psi}(0) = 0$ and that the support of $\widehat{\psi}$ must intersect both halves of the real line.

The Shannon wavelet

Theorem

Let $\psi \in L^2(\mathbb{R})$ be defined by $\widehat{\psi} = \mathbf{1}_{[-1,-1/2]} + \mathbf{1}_{[1/2,1]}$. Then $\{\psi_{j,k}\}$ is a wavelet orthonormal basis for $L^2(\mathbb{R})$.

- The wavelet coefficients

$$\begin{aligned} \langle f, D_{2^j} T_k \psi \rangle &= \langle \widehat{f}, D_{2^{-j}} M_k \widehat{\psi} \rangle \\ &= \int_{-\infty}^{\infty} \widehat{f}(\gamma) \overline{\widehat{\psi}(2^{-j}\gamma)}\, 2^{-j/2} e^{-2\pi i k 2^{-j}\gamma}\, d\gamma \end{aligned}$$

  are the Fourier coefficients of $\widehat{f}$ restricted to the interval $[-2^j, -2^{j-1}] \cup [2^{j-1}, 2^j]$.
- For each fixed $j \in \mathbb{Z}$, the coefficients $\{\langle f, \psi_{j,k} \rangle\}_{k \in \mathbb{Z}}$ capture the features of f at "scale" 2^{-j}.

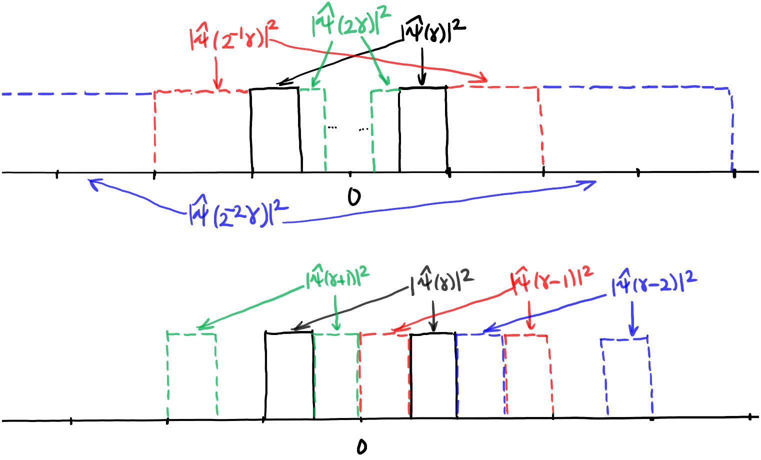
- Coefficients at different values of k identify the intensity of that range of frequencies at “time” $2^{-j}k$.
- While the Haar system have perfect time-localization but very poor frequency localization, the Shannon wavelet basis has perfect frequency localization and poor time localization.

Proof:

- The set of $\operatorname{supp} \widehat{\psi} = [-1, -1/2] \cup [1/2, 1]$ tiles the line by dilation by powers of 2 and integer shifts.
- Since Condition 2 of Theorem B holds, i.e.

$$\sum_{k \in \mathbb{Z}} \widehat{\psi}(2^j(\gamma + k)) \overline{\widehat{\psi}(\gamma + k)} = 0$$

for all $j \geq 1$, $\{\psi_{j,k}\}$ is an orthonormal system.



$|\hat\psi(2^{-1}\gamma)|^2$
$|\hat\psi(2\gamma)|^2$
$|\hat\psi(\gamma)|^2$
0
$|\hat\psi(2^{-2}\gamma)|^2$
$|\hat\psi(\gamma+1)|^2$
$|\hat\psi(\gamma)|^2$
$|\hat\psi(\gamma-1)|^2$
$|\hat\psi(\gamma-2)|^2$
0

- For completeness note that

$$\{2^{-j/2}\, e^{-2\pi i k 2^{-j}\gamma} : k \in \mathbb{Z}\}$$

is an orthonormal basis for $L^2([-2^j, -2^{j-1}] \cup [2^{j-1}, 2^j])$. This follows from the tiling property of $\operatorname{supp} \psi$.

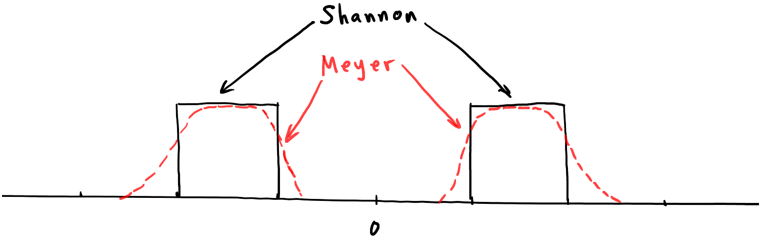
- The wavelet coefficients

$$\langle f, D_{2^j} T_k \psi\rangle = \langle \widehat{f}, D_{2^{-j}} M_k \widehat{\psi}\rangle = \int_{-\infty}^{\infty} \widehat{f}(\gamma)\overline{\widehat{\psi}(2^{-j}\gamma)}\, 2^{-j/2} e^{-2\pi i k 2^{-j}\gamma}\, d\gamma$$

are the Fourier coefficients of $\widehat{f}$ restricted to the interval $[-2^j, -2^{j-1}] \cup [2^{j-1}, 2^j]$.

Meyer wavelets

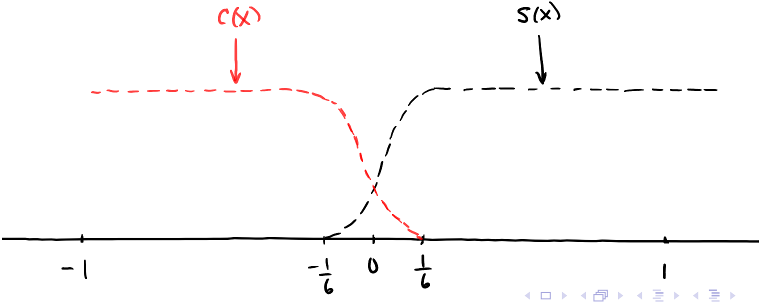
- Meyer constructed an orthonormal basis that is well-localized in both time and frequency.
- The idea is to "smooth-off" the Shannon wavelet in the frequency domain.
- This results in a bandlimited wavelet $\psi(x)$ with smoothness up to C^∞ and rapid decay.



The Meyer wavelet

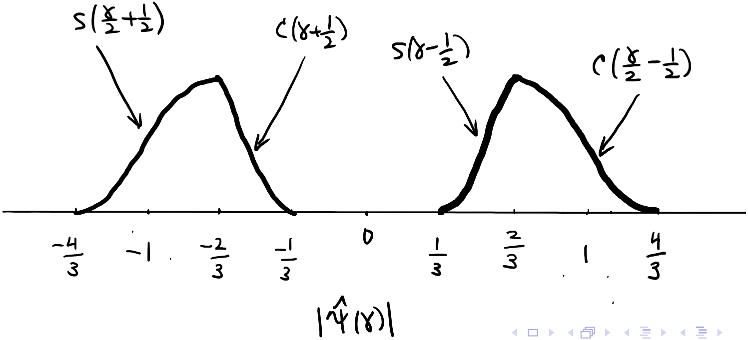
Define the functions $s(x)$ and $c(x)$ with the following properties:

- $s(x) = 0$ and $c(x) = 1$ if $x \leq -1/6$,
- $s(x) = 1$ and $c(x) = 0$ if $x \geq 1/6$,
- $0 \leq s(x), c(x) \leq 1$ for all x, and
- $s(x)^2 + c(x)^2 = 1$ for all x.

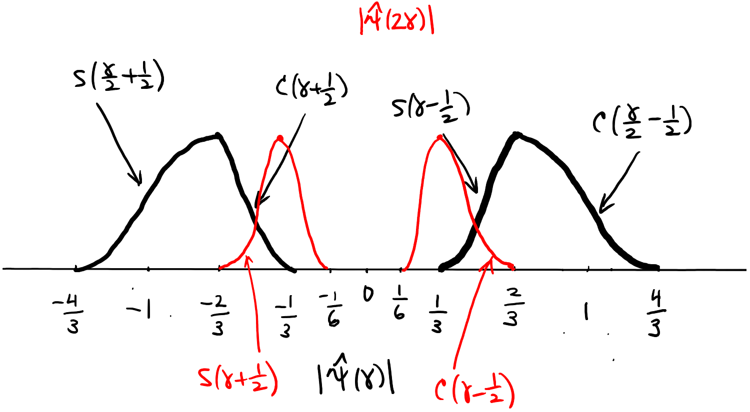


Define $\psi(x)$ by means of its Fourier transform by

$$\widehat{\psi}(\gamma) = -e^{-\pi i \gamma} \begin{cases} 0 & \text{if } |\gamma| \leq 1/3 \text{ or } |\gamma| \geq 4/3, \\ s(\gamma - 1/2) & \text{if } \gamma \in [1/3, 2/3], \\ c(\gamma/2 - 1/2) & \text{if } \gamma \in [2/3, 4/3], \\ s(\gamma/2 + 1/2) & \text{if } \gamma \in [-4/3, -2/3], \\ c(\gamma + 1/2) & \text{if } \gamma \in [-2/3, -1/3]. \end{cases}$$



We show first that $\sum_{j\in\mathbb{Z}} |\widehat{\psi}(2^j\gamma)|^2 \equiv 1.$



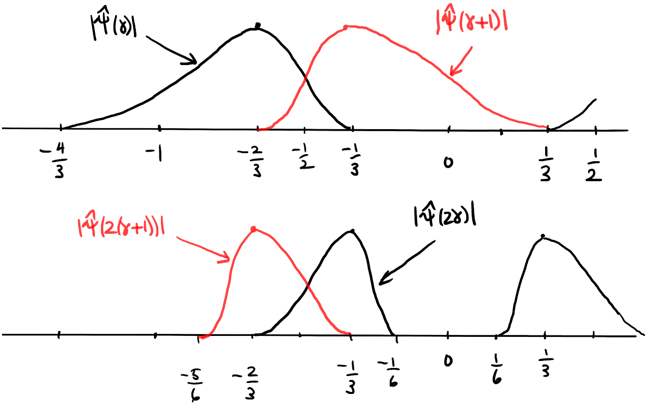
- Next we show that

$$\sum_{j=0}^{\infty} \widehat{\psi}(2^j\gamma)\,\overline{\widehat{\psi}(2^j(\gamma+k))} = 0$$

  for all odd integers k.
- The identity holds trivially when $k \neq -1, 1$. By symmetry it suffices to check it only for $k = 1$.

When $k = 1$, the product being summed is only nonzero if $j = 0, 1$. So we are verifying that

$$\widehat{\psi}(\gamma)\widehat{\psi}(\gamma+1) + \widehat{\psi}(2\gamma)\widehat{\psi}(2(\gamma+1)) = 0.$$



It suffices to check $\gamma \in [-2/3, -1/3]$,

$$\begin{aligned}
& \widehat{\psi}(\gamma)\widehat{\psi}(\gamma+1) + \widehat{\psi}(2\gamma)\widehat{\psi}(2(\gamma+1)) \\
= \; & [(-e^{-\pi i\gamma})(\overline{-e^{-\pi i(\gamma+1)}}) \\
& \quad +(-e^{-2\pi i\gamma})(\overline{-e^{-2\pi i(\gamma+1)}})]s(\gamma+1/2)c(\gamma+1/2) \\
= \; & (e^{\pi i}+1)s(\gamma+1/2)c(\gamma+1/2) = 0.
\end{aligned}$$

- The above cancellation has been referred to as “miraculous” and beg to be understood more systematically.
- This is done through the notion of a Multiresolution Analysis (Meyer, Mallat).

Local cosine bases

- Choose a sequence $\{\alpha_j\}_{j\in\mathbb{Z}}$ strictly increasing and going to infinity in both directions and numbers ϵ_j satisfying $\alpha_j + \epsilon_j \leq \alpha_{j+1} - \epsilon_{j+1}$.
- The goal is to construct orthonormal bases for $L^2(\mathbb{R})$ consisting of smooth functions supported on the intervals

$$I_j = [\alpha_j - \epsilon_j, \alpha_{j+1} + \epsilon_{j+1}].$$

- The bases involve multiplying a function by a smooth cut-off function supported on I_j, then applying a sine or cosine basis to the product.
- In this sense it is like a Gabor basis but avoids the Balian-Low phenomenon.

Recall the functions $s_\epsilon(x)$ and $c_\epsilon(x)$ from the construction of the Meyer wavelet:

(1) $s_\epsilon(x) = c_\epsilon(-x)$,

(2) $s_\epsilon^2(x) + c_\epsilon^2(x) = 1$,

(3) $0 \leq s_\epsilon(x),\ c_\epsilon(x) \leq 1$,

(4) $s_\epsilon(x) = 0$ for $x \leq -\epsilon$ and 1 for $x \geq \epsilon$, and

(5) $c_\epsilon(x) = 0$ for $x \geq \epsilon$ and 1 for $x \leq -\epsilon$.

Definition

Given $I = [\alpha, \beta]$ and $\epsilon,\ \epsilon' > 0$ such that $\alpha + \epsilon \leq \beta - \epsilon'$, define the function $b_I(x)$ by $b_I(x) = s_\epsilon(x-\alpha)\, c_{\epsilon'}(x-\beta)$. $b_I(x)$ is referred to as a *smooth bell function over I*.

Definition

Given adjacent intervals $I = [\alpha, \beta]$ and $J = [\beta, \gamma]$, the smooth bell functions b_I and b_J are *compatible* if
$b_I(x) = s_\epsilon(x-\alpha)\, c_{\epsilon'}(x-\beta)$ and $b_J(x) = s_{\epsilon'}(x-\alpha)\, c_{\epsilon''}(x-\beta)$.

Theorem

There exist projection operators P_I such that

(a) *If b_I and b_J are compatible bells on the intervals $I = [\alpha, \beta]$ and $J = [\beta, \gamma]$ then $P_I + P_J = P_{I \cup J}$.*

(b) $P_I P_J = 0$.

Theorem

If I is given as above and if b_I is a smooth bell function over I, then the collection $\{b_I(x)\, e_k(x)\}_{k=0}^{\infty}$ is an orthonormal basis for the subspace $P_I(L^2(\mathbb{R}))$, where

$$e_k(x) = \sqrt{\frac{2}{l}} \cos \frac{2k+1}{2l} \pi (x - a)$$

Then the collection $\{b_{I_j}(x)\, e_k(x) \colon j \in \mathbb{Z}, k \geq 0\}$ is an orthonormal basis for $L^2(\mathbb{R})$

- The bell functions can be chosen to be arbitrarily smooth.
- Note however that because a cosine basis replaces an exponential basis, the Fourier transform of a basis function will occupy a pair of separated intervals.
- This seems to be a requirement to have smooth localized orthonormal bases of Gabor or wavelet type.

Wavelet packets.

- Given any partition of the interval $[0, \infty)$ with dyadic intervals we can find an orthonormal basis for $L^2(\mathbb{R})$ of functions whose Fourier transforms are supported near the set $\{\gamma \colon |\gamma| \in I\}$.
- The functions in the basis are related to a multiresolution analysis with scaling function φ and wavelet ψ.
- Recall that there exist periodic functions $m_0(\gamma)$ and $m_1(\gamma)$ such that

$$\widehat{\varphi}(\gamma) = m_0(\gamma/2)\widehat{\varphi}(\gamma/2) \text{ and } \widehat{\psi}(\gamma) = m_1(\gamma/2)\widehat{\varphi}(\gamma/2).$$

- Given a dyadic interval I, we can define a finite sequence $\epsilon_n = 0$ of 1 and a function $\psi_I(x)$ by

$$\widehat{\psi}_I(\gamma) = \prod_{n=1}^{N} m_{\epsilon_n}(\gamma/2^n)\widehat{\varphi}(\gamma/2^N).$$

- Letting $\mathcal{P}$ denote the dyadic partition of $[0, \infty)$, then the collection of functions

$$\{\psi_I(x - k/|I|) \colon I \in \mathcal{P},\ k \in \mathbb{Z}\}$$

is an orthonormal basis for $L^2(\mathbb{R})$.

- These bases can consist of arbitrarily smooth functions with close to arbitrarily localization in frequency.

Wilson bases

Theorem

Let $\mathcal{G}(g, 1/2, 1)$ be a tight frame for $L^2(\mathbb{R})$ with $\|g\|_2 = 1$ and $g(x) = \overline{g(-x)}$. Then the collection

$$\psi_{k,n}(x) = \begin{cases} \sqrt{2}\cos(2\pi nx)\, g(x - k/2), & \text{if } k+n \text{ is even} \\ \sqrt{2}\sin(2\pi nx)\, g(x - k/2), & \text{if } k+n \text{ is odd} \end{cases}$$

is an orthonormal basis for $L^2(\mathbb{R})$.

These bases were described first by Wilson (1987), and were linked to tight Gabor frames by Daubechies, Jaffard and Journé (1991).

- Again note that in the Fourier domain, $\widehat{\psi_{k,n}}$ will typically look like a pair of symmetric bumps.
- Wilson bases are still poorly understood.