Lecture 13

In this lecture we will extend our ideas from the construction of Haar Basis on $L_2[0,1]$ to ~~construct~~ ideas of more general compactly supported wavelet bases having desired approximation properties in suitable function spaces.

We saw earlier that the Holder smooth functions with smoothness/regularity $\alpha \in (0,1)$ has desired approximation properties regarding finite dimensional spaces created by first few basis elements of a Haar Basis. This indeed raises an interesting question regarding whether this ~~class~~ property can be reversed. In particular if a function has suitable decay of its fourier coefficients in certain orthonormal basis, can we say that function belongs to some smoothness/regularity. It turns out indeed to be the case while using suitable wavelet bases.

Historically function spaces have ~~a~~ been defined by looking at its local and global fluctuations. Here we will state certain equivalent characterization of such function spaces by looking at fourier coefficients in certain wavelet bases.

We begin with ideas leading to construction of wavelet bases in $L_2(\mathbb{R})$ and $L_2([0,1])$, followed by a modicum of function space theory.

Multiresolution Analysis of $L_2(\mathbb{R})$

We mimic ideas from how we proceeded for the Haar Basis. In particular, let $\varphi \in L_2(\mathbb{R})$ be such that the family of translates of φ i.e. $\{\varphi_{0k}, k \in \mathbb{Z}\}$ is an orthonormal system. Let as usual

$$\varphi_{jk}(x) = 2^{j/2}\varphi(2^j x - k), \quad j \in \mathbb{Z}, k \in \mathbb{Z}.$$

Define linear subspaces,

$$V_0 = \{f(x) = \sum_k c_k \varphi(x-k) : \sum_k |c_k|^2 < \infty\}$$

$$V_1 = \{h(x) = f(2x) : f \in V_0\}$$

$$\vdots$$

$$V_j = \{h(x) = f(2^j x) : f \in V_0\} \quad j \in \mathbb{Z}.$$

Definition: Let $\{\varphi_{0k}\}$ be an orthonormal system of $L_2(\mathbb{R})$. The sequence of spaces $\{V_j, j \in \mathbb{Z}\}$ is called a Multiresolution Analysis (MRA) of $L_2(\mathbb{R})$ if it satisfies

(*) $V_j \subset V_{j+1} \quad \forall j \in \mathbb{Z}$

(**) $\bigcup_{j \geq 0} V_j$ is dense in $L_2(\mathbb{R})$

Alternatively, a MRA of $L_2(\mathbb{R})$ is given by a sequence of subspaces $\{V_j, j \in \mathbb{Z}\}$ satisfying the following conditions

(i) $V_j \subset V_{j+1}$ for all j

(ii) $f(x) \in V_j$ iff $f(2x) \in V_{j+1} \quad \forall j$

(iii) $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad \overline{\bigcup_{j \in \mathbb{Z}} V_j} = \overline{\bigcup_{j \geq 0} V_j} = L_2(\mathbb{R})$

(iv) $\exists \varphi \in V_0$ s.t. $\{\varphi(x-k), k \in \mathbb{Z}\}$ is an o.n.b. of V_0.

or father wavelet

The function φ in (iv) above is called a scaling function, of the given (MRA). Setting $\varphi_{jk} = 2^{j/2}\varphi(2^j x - k)$ as before one says that φ_{jk} has scale 2^{-j} and location $k2^{-j}$

Note that (ii) + (iv) $\Rightarrow \{\varphi_{jk}, k \in \mathbb{Z}\}$ is an o.n.b. of V_j. with orthogonal projection onto V_j given by

$$\Pi_{V_j}(f) = \sum_k \langle f, \varphi_{jk}\rangle \varphi_{jk}$$

(iii) $\Rightarrow \Pi_{V_j}(f) \to f$ in $L_2(\mathbb{R})$ as $j \to \infty$.

Example: (Haar MRA) $V_j = \{ f \in L_2(\mathbb{R}) :$ f is constant on $\left[\frac{k}{2^j}, \frac{k+1}{2^j}\right]$ $\forall k\}$ generated by $\varphi = I[0,1]$.

Example: (Box spline MRA) Given $r \in \mathbb{N}$

$$V_j = \{ f \in L_2(\mathbb{R}) \cap C^{r-1}(\mathbb{R}) : f \text{ is a polynomial of degree } r \text{ on } \left[\tfrac{k}{2^j}, \tfrac{k+1}{2^j}\right] \ \forall k \}$$

$r = 0 \to$ Haar.

$r = 1 \to$ piecewise linear continuous functions.

$r = 3 \to$ Cubic splines.

Assume now that $\{V_j, j \in \mathbb{Z}\}$ is a MRA of $L_2[0,1]$ with scaling function or father wavelet φ.

Let $W_j = V_{j+1} \ominus V_j \qquad j \in \mathbb{Z}$

$$\Rightarrow V_j = V_0 \oplus \bigoplus_{l=0}^{j} W_l \Rightarrow \bigcup_{j=0}^{\infty} V_j = V_0 \oplus \bigoplus_{j=0}^{\infty} W_j$$

$\Rightarrow L_2(\mathbb{R}) = V_0 \oplus \bigoplus_{j=0}^{\infty} W_j$ i.e. any $f \in L_2(\mathbb{R})$ can be represented as a series (convergent in $L_2(\mathbb{R})$)

$$f(x) = \sum_k \alpha_k \varphi_{0k}(x) + \sum_{j=0}^{\infty} \sum_k \beta_{jk} \psi_{jk}(x) \qquad (***)$$

where $\beta_{jk} = \langle f, \psi_{jk} \rangle$ and $\{\psi_{jk}, k \in \mathbb{Z}\}$ ~~are~~ is some o.n.b. of W_j^o.

(***) is called an Multiresolution expansion of f.
(***) is called an wavelet expansion if $\exists\, \psi \in L_2(\mathbb{R})$ s.t

$$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k) \quad \forall j, k$$

and ψ is called a mother wavelet.

Indeed where $\{V_j, j \in \mathbb{Z}\}$ is the Haar MRA,
$\psi(x) = I(x \in [0, \tfrac{1}{2}]) - I(x \in (\tfrac{1}{2}, 1])$ is a mother wavelet and we do have a wavelet expansion. In general however there is no guarantee that given an MRA of $L_2(\mathbb{R})$ one can obtain a wavelet expansion for some mother wavelet ψ. Conditions for existence of such a ψ is a theory in its own and construction of such ψ is part of seminal works of Daubechies and Meyer in early 1990's. We are about to provide some snippets. In particular the general framework of wavelet system looks like the following:

1. Pick a function φ (father wavelet) such that $\{\varphi_{0k}\}$ is an o.n.s and (*) and (**) (on page ①) is satisfied i.e φ generates a MRA on $L_2(\mathbb{R})$.
2. Find $\psi \in W_0$ s.t. $\{\psi_{0k}\}$ is o.n.b. of W_0. (mother wavelet or simply wavelet) then consequently $\{\psi_{jk}\}$ is an onb of W_j. (~~Note~~ The mother wavelet is always orthogonal to the father wavelet).

③ Conclude that $f \in L_2(\mathbb{R})$ has unique representation in terms of an L_2-convergent series

$$f(x) = \sum_k \alpha_k \varphi_{0k}(x) + \sum_{j=0}^{\infty} \sum_k \beta_{jk} \psi_{jk}(x) \quad \cdots\cdots (****)$$
$$= \sum_k \alpha_{Jk} \varphi_{Jk}(x) = \sum_{j=J}^{\infty} \sum_k \beta_{jk} \psi_{jk}(x)$$

with wavelet coefficients

$$\alpha_{jk} = \langle \varphi_{jk}, f \rangle, \quad \beta_{jk} = \langle f, \psi_{jk} \rangle$$

(****) is called an inhomogeneous wavelet expansion.
One might often consider the homogeneous wavelet expansion

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_k \beta_{jk} \psi_{jk}(x)$$

where the "reference space" V_0 is eliminated.

α_k's $\leftrightarrow$ General form of the function
β_{jk}'s $\leftrightarrow$ local details inside this general from (detail coefficients)

Questions: (a) How can we check $\{\varphi_{0k}\}$ is an o.n.s?
(b) What are sufficient condition for $V_j \subset V_{j+1}$ for all j?
(c) What are sufficient conditions for $\bigcup_j V_j$ to be dense in $L_2(\mathbb{R})$?
(d) Can we find $\psi \in W_0$ s.t. $\{\psi_{0k}, k \in \mathbb{Z}\}$ is a o.n.b. of W_0?

We provide some answers to these questions first (proofs of these can be found in Giné & Nickl (2015) book).

<u>Answer to (a)</u>: Let $\varphi \in L_2(\mathbb{R})$. Then $\{\varphi_{0k}, k \in \mathbb{Z}\}$ is an o.n.s iff

$$\sum_k |\hat{\varphi}(\xi + 2\pi k)|^2 = 1 \quad (a.e.)$$

where $\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-ix\xi} f(x)\,dx$ is the fourier transform of f.

<u>Answer to (b)</u>: The spaces V_j are nested i.e. $V_j \subset V_{j+1}$ for all j iff $\exists$ a 2π-periodic possibly complex valued function $m_0(\xi)$, $m_0 \in L_2(0, 2\pi)$ such that

$$\hat{\varphi}(\xi) = m_0(\xi/2)\,\hat{\varphi}(\xi/2) \quad (a.e.)$$

<u>Answer to (c)</u>: Let φ be a ~~scaling~~ fn such that $\{\varphi_{0k}\}$ is an o.n.s. and $V_j \subset V_{j+1}$ for all j. Also assume that there exists a bounded non increasing fn Φ such that

$$\int \Phi(|u|)\,du < \infty \quad \text{and} \quad |\varphi(u)| \le \Phi(|u|) \quad (a.e.)$$

Then $\bigcup_{j \ge 0} V_j$ is dense in $L_2(\mathbb{R})$.

<u>Answer to (d)</u>: Let φ be a father that generates a MRA of $L_2(\mathbb{R})$ and let m_0 satisfy condition of (b) i.e. $\hat{\varphi}(\xi) = m_0(\xi/2)\hat{\varphi}(\xi/2)$ (a.e.) Then the inverse fourier ψ of $\hat{\psi}(\xi)$ satisfying the following is a mother wavelet.

$$\hat{\psi}(\xi) = m_1(\xi/2)\,\hat{\varphi}(\xi/2)$$

where $m_1(\xi) = \overline{m_0(\xi + \pi)}\, e^{-i\xi}$.

The above discussion yields the following 6 conclusions/ideas about how to develop a wavelet expansion.

(I) As soon as we know/get hold of the father wavelet $\varphi(x)$ (and hence $\hat{\varphi}(\xi)$) we can immedialily find the mother using answer to part (d).

(II) Although its not clear yet about how to find φ, answers to (a), (b), and (c) provide some useful guidelines to proceed.

<u>Remark</u>: It is also natural to ask the reverse question: "How to construct fathers from a mother?" To be more precise, let $\psi \in L_2(\mathbb{R})$ satisfy that $\{\psi_{jk}, j \in \mathbb{Z}, k \in \mathbb{Z}\}$ is an o.n.b. of $L_2(\mathbb{R})$. Then is ψ necessarily the mother wavelet of a MRA? The answer is in general no! But under mild regularity conditions this question can be answered positively.

<u>Why Study Wavelets</u>: The key reason we will be using the theory of wavelets is the following. The main ~~heuristic~~ idea behind wavelets is that they capture local behavior of functions in a desired way. Heuristically, for typical functions f, the wavelet coefficients $\beta_{jk} = \langle f, \psi_{jk} \rangle$ are large only at low frequencies or wavelets located close to singularities of f. A typical result is the following.

<u>Proposition</u>: Suppose ψ has a compact support $[-S, S]$ and $\int \psi = 0$. Suppose f is piecewise constant with d discontinuities. Then at level j, at most $(2S-1)d$ of the wavelet coefficients $\beta_{jk} = \langle f, \psi_{jk} \rangle$ are nonzero, and those ~~are~~ are bounded by $\|\psi\|_1 \|f\|_\infty 2^{-j/2}$. ∎

Desired Properties of Wavelet Basis

The construction of some celebrated pair of wavelets (φ, ψ) can be found in

- Ingrid Daubechies: Ten lectures on Wavelets (1992)
- Yves Meyer: Wavelets and Operator I & II (1990)
- W. Härdle, G. Kerkyacharian, D. Picard, A. Tsybakov: Wavelets, approximations, and statistical applications (1998)

Before providing a short list of the well known families of pairs (φ, ψ), we first discuss some properties, which are often desired for statistical and approximation theory purposes, that the pair (φ, ψ) might possess.

(I) Support size: suppose that the support of ψ is an interval of length S. Assume it is $[0, S]$ for the purpose of illustration. Then $\psi_{jk} = 2^{j/2}\psi(2^j x - k)$ is supported on $k \cdot 2^{-j} + 2^{-j}[0, S]$. Now if f has a singularity/discontinuity/ugly behavior at x_0, then the size of S determines the range of influence of the singularity on $\langle f, \psi_{jk} \rangle$. Obviously, at resolution level $j \geq 0$, the number of coefficients $\langle f, \psi_{jk} \rangle$ that "feel" the ill behavior at x_0 is simply the number of location indices k for which support(ψ_{jk}) contains x_0, which is S (if $x_0 \in$ Interior(support(ψ_{jk})) $\forall k$) and ($S-1$ o.w.) It is therefore desirable to have small support to make it easier to identify difficult points.

(II) Vanishing Moments: The wavelet ψ is said to have r vanishing moments if $\int x^l \psi(x)\,dx = 0 \quad l = 0, 1, \ldots, r-1$, i.e. ψ is orthogonal to all polynomials of degree $r-1$. It turns out (a fact of Taylor Series Expansion) the number of vanishing moments of ψ governs the decay of wavelet coefficients of a smooth function. For example, we can prove the following result.

Proposition: If f is C^α on $\mathbb{R}$ and ψ has $r \geq \lceil \alpha \rceil$ vanishing moments, then

$$|\langle f, \psi_{jk} \rangle| \leq C(\psi)\, C^* 2^{-j(\alpha + 1/2)}$$

(If $\alpha \in \mathbb{N}$, then $f \in C^\alpha$ means f has α continuous derivatives and $C^* = \|f^{(\alpha)}\|_\infty / \alpha!$. For $\alpha > 0$ but not an integer, $f \in C^\alpha$ implies $C^* = \|f\|_\infty + \sup_{x,y} |f^{(\lfloor \alpha \rfloor)}(x) - f^{(\lfloor \alpha \rfloor)}(y)| / |x-y|^{\alpha - \lfloor \alpha \rfloor} < \infty$)

Proof: We note that Hölder functions can be uniformly approximated by (Taylor) polynomials. Of course, $f \in C^\alpha$ implies that $\exists$ a constant C^* such that for each $x \in \mathbb{R}$, $\exists$ a polynomial $p_x(y)$ of degree $\lceil \alpha \rceil - 1$ such that

$$|f(x+y) - p_x(y)| \leq C^* |y|^\alpha$$

where C^* can be taken as $\sup_{x,y} \dfrac{|f(x) - f(y)|}{|x-y|^{\alpha}}$ if $0 < \alpha < 1$

and $\prod_{j=0}^{\lceil \alpha \rceil - 1} (\alpha - j)$ if $\alpha \geq 1$.

In the context of the proposition, let $p(y)$ be the approximating Taylor polynomial of degree $\lceil \alpha \rceil - 1$ at $x_k = k 2^{-j}$. Then,

$$\int f(x)\, \psi_{jk}(x)\,dx = \int f(x)\, 2^{j/2}\, \psi(2^j x - k)\,dx$$

$$= 2^{-j/2} \int \left\{ f(x_k + 2^{-j}v) - p(2^{-j}v) \right\} \psi(v)\,dv \quad \text{(change of variable \& vanishing moments)}$$

$$\leq 2^{-j/2}\, C^* \int 2^{-j\alpha}\, |v|^\alpha\, |\psi(v)|\,dv$$

$$= C^*\, 2^{-j(\alpha+\frac{1}{2})} \underbrace{\int |v|^\alpha\, |\psi(v)|\,dv}_{C(\psi)<\infty}$$

as it exists. ∎

(III) Regularity: We often use wavelets for projection estimates/approximations $\hat{f}(x) = \sum \hat{\beta}_{jk}\, \psi_{jk}(x)$, and as a result the smoothness of $x \mapsto \psi_{jk}(x)$ can affect the visual appearance of such estimates. However, as we saw, the number of vanishing moments affects the size of the wavelet coefficients at finer scales (large j), at least in regions where f is smooth. Therefore both properties are in general desired. As we shall list below, for commonly used wavelet families, regularity increases with the number of vanishing moments. For orthonormal wavelet bases, regularity of ψ implies that a corresponding number of moments vanish.

Proposition (Daubechies (1992), section 5.5) If $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ is an o.n.b. of $L_2(\mathbb{R})$, and if ψ is C^r with $\psi^{(k)}$ bounded for $k \leq r$ and $|\psi(x)| \leq C(1+|x|)^{-r-1-\varepsilon}$ for some $C>0$, $\varepsilon>0$, then $\int x^k \psi(x)\,dx = 0$ for $k = 0, \ldots, r$.

<u>Some Wavelet Families</u>: The usual construction of pairs of (φ, ψ) use the Fourier analytic techniques discussed earlier. Moreover, many of these constructions yield a family indexed by the number of vanishing moments p.

(i) <u>Haar</u>: $\varphi = I[0,1]$, $\psi = I_{[0,1/2]} - I_{[1/2,1]}$. It has a single vanishing moment & no smoothness.

(ii) <u>Meyer</u>: $\hat{\varphi}(\xi), \hat{\psi}(\xi)$ have compact support $\xi \Rightarrow \varphi, \psi$ are C^∞ but do not compact support (Heisenberg's Uncertainty Principle). The wavelet ψ has infinitely many vanishing moments.

(iii) <u>Battle-Lemarié Spline</u>: This family is derived from the box-spline wavelet described earlier. (φ, ψ) are both polynomial splines of degree $m \Rightarrow C^{m-1}$. ψ has $(m+1)$ vanishing moments and not compact support.

(iv) <u>Compactly Supported Wavelet</u>: These are most useful for us. Ingrid Daubechies constructed several sets of wavelets & scaling functions, indexed by the number of vanishing moments (p) for ψ.

(a) <u>Daubechies Family (D_{2p})</u>: ψ has minimum support length $2p-1$ on $[-p+1, p]$. These wavelets are quite asymmetric and are in $\approx C^{0.2p}$. ~~So~~ φ has support $[0, 2p-1]$ and also in similar regularity class. Actually $\varphi, \psi \in C^{\lambda p}$ for $p \geq 2$ where $0.1936 \leq \lambda \leq 0.2075$. (Note Haar $= D_2$ i.e. $p=1$)

(b) Symmlet Family: Also has support $[-p+1, p]$ for ψ but ~~it~~ more symmetric looking.

(c) Coiflet Family: Has $K = 2p$ vanishing moments for ψ and for φ, $\int \varphi = 1$, $\int t^k \varphi = 0 \quad 1 \le k < K$.

This constraint forces larger support length of $3K-1$.

Wavelets on $[0,1]$: So far we were dealing with MRA on $L_2(\mathbb{R})$. However, in many statistical applications one works with functions defined on compact intervals $I=[0,1]$ (say). A brute-force extension of f to $\mathbb{R}$ by setting it to 0 outside $[0,1]$, or ~~sop~~ even sophisticated ideas of extensions of f outside by reflection or folding, introduces a discontinuity in f or its derivatives at the boundary. Two approaches have typically been taken in literature:

(a) Periodization: (for periodic functions on $I=[0,1]$)

(b) Orthonormalization on $[0,1]$: The idea is done in detail in Cohen, Daubechies, Vial (1993b) where one starts with a Daubechies pair (φ, ψ) with p vanishing moments, keeps the φ_{jk} and ψ_{jk}'s whose support lie inside $[0,1]$ and change the rest carefully to produce a new o.n.b. of $L_2[0,1]$ which puts less stress for boundary effects. This we will refer to CDV wavelets with p vanishing moments.

Function Spaces and Wavelets Coefficients: A basic idea to measure smoothness/regularity of functions is to use relative magnitude of wavelet coefficients across scales. To avoid boundary issues we first work with function spaces as subspace/subset of $L_2(\mathbb{R})$ and corresponding orthonormal basis on $L_2(\mathbb{R})$. Corresponding definition of function spaces ~~as~~ for $L_2[0,1]$ can be obtained using easily extendable ideas/definitions with corresponding CDV family of orthonormal basis of $L_2[0,1]$.

We shall see that suitable smoothness/regularity properties of f can be translated into norms of wavelet coefficients at different scales. In particular we shall look at l_p-norms of $\{\beta_{jk}, k \in \mathbb{Z}\}$ for each $j \geq 0$. We start by explaining/exploring the connection for $p = 1, 2$, and ∞. Assume for all x,

$$f(x) = \sum_k \alpha_{Lk} \varphi_{Lk}(x) + \sum_{j \geq L} \sum_k \beta_{jk} \psi_{jk}(x), \quad L \geq 0$$

Hölder Smoothness ($p = \infty$) We shall only consider $0 < \alpha < 1$. for which $|f(x) - f(y)| \leq c\,|x-y|^\alpha \;\; \forall x, y$.

Proposition: suppose $0 < \alpha < 1$ and (φ, ψ) are C^1 with compact supports. Then f is C^α smooth iff $\exists\, c > 0$ such that

$$|\alpha_{Lk}| \leq c, \quad |\beta_{jk}| \leq c\, 2^{-(\alpha + 1/2)j}, \quad j \geq L, \; k \in \mathbb{Z}$$

$$\text{i.e. } 2^{(\alpha + \frac{1}{2})j} |\beta_{jk}| \leq c \quad \forall j \geq L, \; k \in \mathbb{Z}$$

(conditions are uniform on k)

$$\text{i.e. } 2^{(\alpha + \frac{1}{2})j} \|\beta_{j\cdot}\|_\infty \leq c \text{ and } \|\alpha_{L\cdot}\|_\infty \leq c \quad \forall j$$

Proof: Recall that by definition $f \in C^\alpha$ iff (for $0<\alpha<1$)

$$\|f\|_\infty + \underbrace{\sup_{x,y} \frac{|f(x)-f(y)|}{|x-y|^\alpha}}_{|f|_\alpha \text{ (this is not a norm)}} < \infty$$

The scale coefficients are easily verified to be bounded i.e.

$$|\alpha_{Lk}| \le \int |f(x)|\,|\varphi_{Lk}(x)|\,dx$$

$$\le 2^{-L/2} \|f\|_\infty \|\varphi\|_1 .$$

For the wavelet coefficients, first note that we can assume $\int \psi(x)\,dx = 0$ since $\psi \in C^1$ (choose a suitable Daubechies wavelet)

$$\Rightarrow \beta_{jk} = \langle f, \psi_{jk} \rangle = 2^{-j/2} \int f(k2^{-j} + 2^{-j}v)\, \psi(v)\, dv$$

$$= 2^{-j/2} \int \left\{ f(k2^{-j} + 2^{-j}v) - f(k2^{-j}) \right\} \psi(v)\, dv$$

(as $\int \psi = 0$)

By Hölder smoothness therefore,

$$|\beta_{jk}| \le 2^{-j/2} |f|_\alpha\, 2^{-j\alpha} \int |v|^\alpha |\psi(v)|\, dv = 2^{-j(\alpha + \frac{1}{2})} c_{\psi,\alpha} |f|_\alpha$$

with $c_{\psi,\alpha} = \int |v|^\alpha |\psi(v)|\, dv$.

$\Rightarrow$ Take $C = \max\left\{ 2^{-L/2} \|f\|_\infty \|\varphi\|_1 ,\ c_{\psi,\alpha} |f|_\alpha \right\}$

In the reverse direction we wish to show that the prescribed decay of wavelet coefficients implies $\|f\|_\infty + |f|_\alpha \le C^*$ for some constant C^* depending C, α, and properties of φ & ψ.

The bound on $\|f\|_\infty$ is easy. Using its wavelet expansion, for any x

$$|f(x)| \le \sum_k |\alpha_{Lk}|\, |\varphi_{Lk}(x)| + \sum_{j \ge L} \sum_k |\beta_{Lk}|\, |\psi_{\ell k}(x)|$$

Now if length of support of φ and ψ are at most S, then for any x, at most S terms in both the first sum over k and second sum over k covers x. Therefore

$$\|f\|_\infty \leq SC\|\varphi\|_\infty 2^{L/2} + \sum_{j\geq L} SC\, 2^{-j(\alpha+\frac{1}{2})} \cdot 2^{j/2}\|\psi\|_\infty$$

$$\leq SC 2^{L/2}\|\varphi\|_\infty + \sum_{j\geq L} SC\|\psi\|_\infty 2^{-\alpha j}$$

$$\leq SC\max(\|\varphi\|_\infty, \|\psi\|_\infty)\left\{2^{L/2} + 2^{-\alpha L}/1-2^{-\alpha}\right\}$$

Now we bound $|f|_\alpha$. Take any $x, x' \in \mathbb{R}$. Decompose the difference of $f(x) - f(x')$ into $\Delta_1(f)$ and $\Delta_2(f)$ where

$$\Delta_1(f) = \sum_k \alpha_{Lk}\left(\varphi_{Lk}(x) - \varphi_{Lk}(x')\right)$$

$$\Delta_2(f) = \sum_{j\geq L}\sum_k \beta_{Lk}\left(\psi_{jk}(x) - \psi_{jk}(x')\right)$$

We deal with $\Delta_2(f)$ here. The control of $\Delta_1(f)$ is similar and easier. Once again using compactness of support of ψ with support length at most S,

$$|\Delta_2(f)| \leq C\sum_{j\geq L} 2^{-(\alpha+\frac{1}{2})j}\sum_k 2^{j/2}\left|\psi(2^jx-k) - \psi(2^jx'-k)\right|$$

$$\leq C\sum_{j\geq L} 2^{-j\alpha}\, 2S\left(\|\psi'\|_\infty \overset{|x-x'|^\alpha}{\wedge} I\left(|2^jx-2^jx'| \leq 1\right) + 2\|\psi\|_\infty I\left(|2^jx - 2^jx'| > 1\right)\right)$$

$$\leq 2CS\left(\|\psi'\|_\infty + 2\|\psi\|_\infty\right)\sum_{j\geq L} 2^{-j\alpha}\left(2^j|x-x'| \wedge 1\right)$$

$$\leq \frac{2CS\left(\|\psi'\|_\infty + 2\|\psi\|_\infty\right)}{1-2^{-\alpha}}|x-x'|^\alpha$$

This completes the proof of the proposition. ∎

Mean Square Smoothness ($p=2$): Assume f is r-times weakly differentiable and let D^r stand for the weak differential operator. The mean square smoothness is measured by the norm

$$\|f\|^2_{W_2^r} = \int f^2 + \int (D^r f)^2 < \infty$$

It turns out that mean squared smoothness also has a very natural expression in terms of wavelet expansion. To demonstrate this, let (φ, ψ) be C^r. Then we may formally differentiate the homogeneous wavelet expansion $f = \sum_{j,k} \beta_{jk} \psi_{jk}$ to get $D^r f(x) = \sum_{jk} 2^{rj} \beta_{jk} \psi_{jk}^{(r)}(x)$

Although the system $\{\psi_{jk}^{(r)}\}$ is no longer orthonormal, it turns out that it is almost as good $\longleftrightarrow$ a thing called a tight frame. This means that $\exists\, c_1, c_2$ such that for all f with $\|f\|_{W_2^r} < \infty$,

$$c_1 \sum_{jk} 2^{2rj} \beta_{jk}^2 \le \Big\| \sum_{jk} 2^{rj} \beta_{jk} \psi_{jk}^{(r)}(x) \Big\|_2^2 \le c_2 \sum_{jk} 2^{2rj} \beta_{jk}^2$$

This in turn implies the following proposition, proof of which can be found in Johnstone: Gaussian Sequence Models (2015) Appendix (B.4).

Proposition: If (φ, ψ) are C^r with compact support and ψ has $r+1$ vanishing moments, then $\exists\, c_1$ and c_2 such that

$$c_1 \|f\|^2_{W_2^r} \le \underbrace{\sum_k \alpha_{Lk}^2}_{\|\alpha_{L\cdot}\|_2^2} + \sum_{j \ge L} \underbrace{\sum_k 2^{2rj} \beta_{jk}^2}_{2^{2rj} \|\beta_{j\cdot}\|_2^2} \le c_2 \|f\|^2_{W_2^r}$$

Average Smoothness ($p=1$): The average smoothness is measured in an L_1-sense of the weak derivative. One uses the norm

$$\|f\|_{W_1^1} = \|f\|_1 + \int |Df| .$$

We now show that membership in W_1^1 (i.e. $\|f\|_{W_1^1} < \infty$) implies near characterization by ℓ_1-type conditions on the wavelet coefficients.

Proposition: suppose (φ, ψ) are C^1 with compact support. Then $\exists$ constants C_1 and C_2 such that

$$C_1 \left[\|\alpha_{L\cdot}\|_1 + \sup_{j \geq L} 2^{j/2} \|\beta_{j\cdot}\|_1 \right] \leq \|f\|_{W_1^1} \leq C_2 \left[\|\alpha_{L\cdot}\|_1 + \sum_{j \geq L} 2^{j/2} \|\beta_{j\cdot}\|_1 \right]$$

Proof: First note that $\|\psi_{jk}\|_{W_1^1} \leq 2^{j/2}\left(\|\psi\|_1 + \|\psi'\|_1\right)$

$$\Rightarrow \|f\|_{W_1^1} \leq 2^{L/2}\left(\|\varphi\|_1 + \|\varphi'\|_1\right) \sum_k |\alpha_{Lk}| + \left(\|\psi\|_1 + \|\psi'\|_1\right) \sum_{j \geq L} 2^{j/2} \sum_k |\beta_{jk}|$$

Take $C_2 = \max\left(2^{L/2} \|\varphi\|_{W_1^1} + \|\psi\|_{W_1^1} \right)$

For the reverse, suppose $\|f\|_{W_1^1} < \infty$. We can take $\int \psi = 0$ and assuming that for some interval I, support$(\psi) \subseteq I$, we have by integration by parts,

$$\left| \int_I f \psi \right| \leq \frac{1}{2} \|\psi\|_1 \int_I |Df|$$

For Daubechies family we can take $I = [-S+1, S]$ for $S \geq 2.5$ and applying the above bound for $\beta_{jk} = \int f \psi_{jk}$ implies $|\beta_{jk}| \leq \frac{1}{2}\|\psi\|_1 \, 2^{-j/2} \int_{I_{jk}} |Df|$ where $I_{jk} = 2^{-j}[k-S+1, k+S]$.

For j fixed as k varies, any given point x falls in at most $2S$ intervals I_{jk} $\Rightarrow$ adding over k we have for any j

$$2^{j/2} \sum_k |\beta_{jk}| \leq 2S \cdot \tfrac{1}{2} \|\psi\|_1 \int |Df|$$

A similar argument shows $\|\alpha_{L\cdot}\|_1 \leq 2^{L/2}\, 2S \|\varphi\|_\infty \|f\|_1$

This completes the proof of the proposition. $\blacksquare$

Besov Spaces: Comparing the three cases above we can now contrast how the coefficients at a given level j are weighted and combined over k.

Hölder $(p=\infty)$	$2^{(\alpha+\frac{1}{2})j} \|\beta_{j\cdot}\|_\infty$
Mean Square $(p=2)$	$2^{\alpha j} \|\beta_{j\cdot}\|_2$
Average $(p=1)$	$2^{(\alpha-\frac{1}{2})j} \|\beta_{j\cdot}\|_1$

(In the last two cases we are extrapolating).

Now introduce a new notation/index: $a = \alpha + \frac{1}{2} - \frac{1}{p}$

$\Rightarrow$ Each of the cases are $2^{aj} \|\beta_{j\cdot}\|_p := c_j$ (say). To combine c_j's across levels j, one uses l_q-norms $\left(\sum_{j\geq L} |c_j|^q\right)^{1/q}$.

Now use $\underset{\sim}{\theta} = \{\alpha_{Lk}, \beta_{jk}, j \geq L, k \in \mathbb{Z}\}$ and let

$$\|\theta\|_{b^\alpha_{p,q}} = \left(\sum_k |\alpha_{Lk}|^p\right)^{\frac{1}{p}} + \left(\sum_{j\geq L} 2^{(\alpha+\frac{1}{2}-\frac{1}{p})jq} \left(\sum_k |\beta_{jk}|^p\right)^{q/p}\right)^{\frac{1}{q}}$$

$$= \|\alpha_{L\cdot}\|_p + \left\| \left(2^{(\alpha+\frac{1}{2}-\frac{1}{p})L} \|\beta_{L\cdot}\|_p,\ 2^{(\alpha+\frac{1}{2}-\frac{1}{p})(L+1)} \|\beta_{L+1,\cdot}\|_p, \dots\right) \right\|_q$$

Thus Hölder $\longleftrightarrow$ $\|\theta\|_{b^{\alpha}_{\infty\infty}}$

$W_2^2 \longleftrightarrow \|\theta\|_{b^{\alpha}_{22}}$

$W_1^1 \longleftrightarrow \|\theta\|^{\alpha}_{11}$

The three parameters α, p, q may be interpreted as

$\alpha > 0$: Smoothness

$p \in (0,\infty]$: averaging (quasi)-norms over locations k.

$q \in (0,\infty]$: averaging (quasi)-norms over scales j.

One thereafter defines Besov space of smoothness α as follows. Let (φ, ψ) be of regularity $r \geq \lceil\alpha\rceil$ (has at least r vanishing moments and C^r). Then, let

$$B^{\alpha}_{pq,W}(\mathbb{R}) = \left\{ f \in L_p(\mathbb{R}) : \left(\sum_k |\langle f, \varphi_{0k}\rangle|^p\right)^{\frac{1}{p}} + \left(\sum_{j\geq 0} 2^{(\alpha+\frac{1}{2}-\frac{1}{p})jq} \left(\sum_k |\langle f, \psi_{jk}\rangle|^p\right)^{q/p}\right)^{\frac{1}{q}} < \infty \right\}$$

Call $\|f\|_{B^{\alpha}_{p,q,W}} = \|\theta\|_{b^{\alpha}_{pq}} + \|f\|_p$ where $\underset{\sim}{\theta}$ collects the wavelet coefficients of α. It turns out that B^{α}_{pq} has an equivalent characterization in ~~the~~ terms of the modulus of smoothness. In particular, let $\tau_h(f)(x) = f(x+h)$ and let $\Delta_h f(x) = \tau_h f(x) - f(x)$. Inductively now define for $r \in \mathbb{N}$, $\Delta_h^r(f) = \Delta_h(\Delta_h^{r-1}(f))$

$$\Rightarrow \Delta_h^r(f)(x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(x+kh)$$

Now for $f \in L_p(\mathbb{R})$ define the rth moduli of smoothness of f:

$$\omega_r(f,t) \equiv \omega_r(f,t,p) := \sup_{0<h\leq t} \|\Delta_h^r(f)\|_p, \quad t > 0.$$

Let

$$B^{\alpha}_{pq}(\mathbb{R}) = \left\{ f \in L_p(\mathbb{R}) : \|f\|_{B^{\alpha}_{pq}} \equiv \|f\|_p + |f|_{B^{\alpha}_{pq}} < \infty \right\} \quad \text{for } 1 \leq p < \infty$$

$$= \left\{ f \in C_u(\mathbb{R}) : \|f\|^{\alpha}_{pq} \equiv \|f\|_{\infty} + |f|_{B^{\alpha}_{pq}} < \infty \right\}$$

where $C_a(\mathbb{R})$ is set of all absolutely continuous functions on $\mathbb{R}$ and

$$|f|_{B^{\alpha}_{pq}} = \begin{cases} \int_0^{\infty} \left(\frac{\omega_r(f, t, p)}{t^{\alpha}} \right)^q \frac{dt}{t} & 1 \le q < \infty \\ \sup_{t>0} \frac{\omega_r(f, t, p)}{t^{\alpha}} & q = \infty \end{cases}$$

It is a deep result in function space theory that $B^{\alpha}_{pq,W} = B^{\alpha}_{pq}$ with equivalence of norms i.e. $\exists$ constants C_1 and C_2 such that

$$C_1 \|f\|_{B^{\alpha}_{pq}} \le \|f\|_{B^{\alpha}_{p,q,W}} \le C_2 \|f\|_{B^{\alpha}_{pq}}$$

(C_i may depend on $\varphi, \psi, \alpha, p, q$, but not on f).

Remark: check that for $0<\alpha<1$, $B^{\alpha}_{\infty\infty}$ actually corresponds to our usual Hölder space definition using the modulus of continuity definition.

On $[0,1]$: It can similarly be developed for functions on $[0,1]$, that if one use CDV type orthonormal wavelet basis of $L_2[0,1]$ with (φ, ψ) in C^r and having compact support, then once again the sequence norm $\|f\|_{B^{\alpha}_{pq,W}}$ and moduli of smoothness norm $\|f\|_{B^{\alpha}_{pq}}$ are equivalent whenever $[\alpha] \le r$. We will in general work with Besov Balls $B^{\alpha}_{pq}(M) = \{f : \|f\|_{B^{\alpha}_{pq}} \le M\}$ which is "similar" to the space $B^{\alpha}_{pq,W} = \{f : \|f\|_{B^{\alpha}_{pq,W}} \le M\}$.